The Intersection of Incidental Bidirectional Naming and Behavior Analytic Instructional Design Tactics

Ellis Smith

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2024

Abstract

The Intersection of Incidental Bidirectional Naming and Behavior Analytic Instructional Design Tactics

Ellis Smith

All individuals develop new skills over time and one important developmental stage is when individuals can learn in the absence of direct reinforcement. Behavior Analytic theories have stressed the importance of these emergent response repertoires and have consistently studied them across fields (i.e. Incidental Bidirectional Naming, Arbitrarily Applicable Relational Responding, Equivalence). Emergent responses are responses that are not explicitly taught during instruction. Recently, behavior analytic perspectives have studied their similarities across perspectives, one area of interest being Incidental Bidirectional Naming (Inc-BiN). Researchers have identified multiple types of Naming, including Pre-Unidirectional Naming (Pre-UniN), Incidental Unidirectional Naming (Inc-UniN), and Incidental Bidirectional Naming (Inc-BiN). Study 1 investigated the basic question about whether verbal development – in this case degree of Inc-BiN – is related to academic performance (mathematics and reading). The researcher correlated participants' listener and speaker responses for two Brief Inc-BiN Probes with academic performance (*iReady*® Diagnostic Mathematic and Reading scores) for 41 participants. Results demonstrated significant positive correlations between degree of Inc-BiN and reading and mathematics performance. Given this correlation, two experiments examined the relationship between verbal development and two specific instructional design tactics that focus on the emergence of novel behavior. Experiment 2 investigated the effectiveness and appropriateness of a relational training procedure (matrix instruction) on participants' emergent intraverbal responses. The researcher used learn units to teach intraverbal atomic units and measured generalized operant responding. The researcher used a matrix with five prefixes and five root words, which combined to make 25-word combinations. The researcher taught five combinations (1 prefix plus 1 root) and tested the other 20-word combinations. The researcher implemented a multiple probe design and measured emergent intraverbals and degree of Inc-BiN prior to and following intervention. Results demonstrated that emergent intraverbal responses varied in accordance with participant degree of Inc-BiN. Experiment 3 investigated the effects of another relational training procedure (Equivalence-Based Instruction; EBI), across participants with ranging degrees of Inc-BiN. During teaching, the researcher taught two of six relations in a class to mastery using learn unit instruction across 13 participants. Results in Experiment 3 demonstrated that, after EBI, participants emitted a greater number of emergent intraverbals when they had higher degrees of Inc-BiN and lower emergent intraverbals when they had lower degrees of Inc-BiN. Experiment 2 and 3 demonstrate similar results that participants with Inc-UniN and Inc-BiN emitted a high percentage of emergent relations following learn unit instruction across the three studies. Data highlight the importance of Inc-BiN's relation to academic performance in general (Experiment 1) and in predicting participants' success with behavior analytic instructional design tactics (Experiment 2 and 3).

Keywords: Academic Performance, Combinatorially Entailed Relations, Emergent Intraverbals, Incidental Bidirectional Naming, Mutually Entailed Relations

Table of Contents

List of Tables

List of Figures

Acknowledgments

Thank you to everyone who inspired me, believed in me, and taught me.

Thank you to my dissertation committee, Dr. Fienup, Dr. Dudek, Dr. Sivaraman, Dr. Gautreaux, and Dr. Lovett, for your time, intellect, and patience. I appreciate all your thoughtful questions and advice. Thank you to my professors, Dr. Fienup, Dr. Dudek, Dr. Delgado, Dr. Weber, and Dr. Greer, for teaching me the fundamentals and inspiring my love of learning and research along the way.

Dr. Fienup, you guided me during every step in my development as a researcher and a writer. I greatly appreciate your wisdom and perspective on everything. Thank you for showing me how to balance life with being a behavior analyst.

Dr. Weber…I have struggled to find the words to express my gratitude for you since I began this program. I look up to as a practitioner, researcher, teacher trainer, professor, and person so much. The way you love the kids and support your teachers is inspiring and I truly hope to support others the way you have supported me.

Dr. Gautreaux, thank you for taking me into the Louisiana CABAS family. You have made me feel so valued. I cannot thank you enough for the opportunities you have provided me with. I am so lucky to work and learn with you and I'm so excited for what our future research holds.

Thank you to my mentors, Hung, Leah, and Yifei, for training me, being patient with me, and checking up on me. Hung, you taught me so much and I am so grateful I was placed in class 6 first semester! Yifei, thank you for inspiring my love of research and helping be become a better writer. I was originally terrified to teach 5th graders, but your wisdom and support led me to love it!

Thank you to my mentees, Blanca, Rachel, and Gianna, for helping me become a better supervisor and loving our 4th graders. Rachel, thank you for putting your time and effort to help me start the first studies of my dissertation. Gianna, thank you for contributing your passion and thoughtful questions about this research and our students. Each of our students were better because they had you as their teacher, curriculum designer, and cheerleader. A special thank you to Renee, who seamlessly conducted research in her classroom and knows way more about behavior analysis than she lets on!

To the AH community, thank you for teaching me how to be a teacher. Although our time was cut short, the impact you all have had on me is unmatched. Thank you to my 4th grade team, Kristen, Raj, Patricia, Christine, and Emily, who took me in as their CABAS youngster. Thank you to Ed and Sue, for trusting me and always backing me. Thank you to Mary-Kate and Jen, for getting me through being a first-year teacher and making me laugh the whole way through. AH will always be a special place to me that I value very much.

To my cohort (all three of us!), Lauren and Han, we made it! After five years of growing pains, we have reached the finish line. Lauren, when we started at Keller together, I had no idea the impact you would have on me as a behavior analyst and as a person. Without you, I really don't know if I could have gotten to this point. Our "car talks" inspired most of this research and having you as one of my best friends got me through all the not so fun times. I think it is safe to say we have gone through something together that only we will ever truly understand.

Thank you to my family, Susan, Dave, Kylie, and Robbie, for keeping me grounded throughout this process. Thank you to Robbie for making sure I took breaks, encouraging me to push through, and keeping me smiling the whole time. Thank you Kylie for challenging me from day 1. I admire your intelligence and confidence in who you are and what you do. Thank you to my parents for supporting anything and everything I ever wanted to do. Mom and Dad, seeing you two love what you do and work so hard has had such a big impact on me and I appreciate all the sacrifices you made for me.

Dedication

To each of my students. Thank you for trusting me and inspiring this research.

Chapter 1: Introduction and Literature Review

The goal of education is to maximize student learning and achieve curricular goals. Whether educators are curriculum design professionals working to create a comprehensive and systematic curriculum for teachers and students, or teachers who are expected to motivate students and implement curricula, educators are consistently challenged by their students' needs and administrative demands to deliver effective and efficient instruction. Many teachers across the United States are required to teach their students material based on the Common Core State Standards (CCSS). The CCSS are a set of academic standards for English Language Arts (ELA) and mathematics that are organized by each grade level. The CCSS is "designed to prepare all students for success in college, career, and life by the time they graduate from high school" (Common Core State Standards Initiative). Curriculum companies design their curriculum to align with the CCSS and teachers are expected to implement the curriculum to teach the CCSS. The CCSS began in 2010, however, teachers have reported a growing need for time and resources, especially since the COVID-19 pandemic (Rosenberg & Anderson, 2021). For example, a 4th grade inclusion teacher may be responsible for a class of students who range in academic performance from 2nd grade to 6th grade. This teacher may be expected to teach prerequisite skills to the below-grade-level students prior to teaching them the 4th grade standards, while teaching the grade-level students the 4th grade standards and simultaneously engaging the above-grade-level students. Although one example, teachers across the country need resources that maximize their existing instructional time to benefit all their students. Two possible methods for maximizing instructional time are through teaching tactics and curricular design tactics.

1.1 Teaching Tactics and Curricular Design Tactics

Educators may use teacher-implemented tactics or curricular design tactics to help their students learn. First, teacher-implemented tactics, or teaching tactics, include strategies teachers implement within their own classroom to accomplish student learning objectives. These can be further categorized into *motivational tactics* (Greer, 2002, Ch.1; Hart & Risley, 1980; Michael, 1983*)* and *learning tactics* (Albers & Greer, 1991). Motivational tactics are tactics that are used to change the frequency of a performance behavior, or a behavior that is already in the students' repertoire (e.g., hand raising, sitting in their seat). Motivational tactics may include common strategies like vocal praise (e.g., "Lucy you are doing such a great job following directions"), specific contingencies (e.g., group contingency or peer-yoked contingency gameboard), or token economies (e.g., providing points/tokens for target behavior that they can exchange to access a reinforcer later) (Gunersel et al., 2023; Kim et al., 2021; Page et al., 2023). Motivational tactics can help improve the rate of work completion and rule following within a class to help maximize teachers' time spent teaching as opposed to managing classroom behavior.

Teachers also implement learning tactics, which are used to help students acquire novel behavior, such as CCSS objectives. Learning tactics may include modeling (e.g., showing an example of how to solve a problem), providing opportunities to respond (e.g., programming for multiple opportunities to practice a given skill), providing performance feedback (e.g., telling the student they got the question right or showing how to solve the problem correctly; Hranchuk, 2016; Menzies, et al., 2017; Van Den Bergh et al., 2013), or strategically arranging the antecedent[1], response[2] and consequence[3] contingencies (Albers & Greer, 1991). Learning tactics

---

[1] The antecedent is the stimulus that occurs immediately before the behavior (Cooper et al., 2019).
[2] The behavior is the student's response that occurs after the antecedent (Cooper et al., 2019).
[3] The consequence is the response to the behavior, likely, the reinforcement delivered after a correct response or correction after an incorrect response (Cooper et al., 2019).

help students acquire novel behavior/sets of behavior that they could not previously perform. When implemented correctly, learning tactics help maximize instructional time by decreasing the time it takes for the students to learn a new skill.

Second, curricula can be designed to promote effective and efficient student learning. Curriculum tactics can be further broken down into opportunities to respond and curriculum sequencing. Many researchers have investigated the relation between opportunities to respond and rate of learning and found the more learning opportunities an individual has, the faster they will acquire that given behavior (Albers & Greer; 1991; Bahadourian et al, 2006; Dorow et al., 1989). This suggests that curriculum can aid students' learning by providing ample opportunities for each skill targeted. For example, during a lesson about fraction vocabulary, curriculum designers may need to program their curriculum to allow for opportunities to select a picture of a word, select the definition of a word, draw a picture of a word, and write the definition of the word. The example above highlights multiple types of opportunities but opportunities to respond can be simply creating a chance for an individual to attempt an answer. Although teachers can sometimes provide these opportunities to learn, curricula can increase these opportunities by programming them into the curriculum. These opportunities can be *direct learning opportunities* (Engelmann & Carnine, 1991; Greer et al., 2006) or *emergent learning opportunities* (Chomsky, 1959; Critchfield & Twyman, 2014).

Direct learning opportunities mean the student is required to answer a question with a direct answer, or one that is directly stated (Cooper et al., 2019). For example, given explicit instruction on how to solve a single digit by single digit addition problem, a direct learning math question may ask students to solve the equation "2 + 5 = __." This was a direct learning opportunity because the addition response was directly taught (prompted and reinforced).

Emergent learning opportunities mean the student is required to produce a response that was not explicitly taught. This means the answer is implicit, and not directly stated (Cooper et al., 2019). For example, if a student is taught that "2 + 5 = 7," they may emit an emergent response that "5 + 2 = 7." This is an emergent learning opportunity because the responses with the "5" at the front of the equation with "2" as the second number, was not directly taught. Students are expected to be able to answer both direct and emergent questions throughout each grade level, subject area, as well as in social settings. Educators can design curriculum to include both direct and emergent learning opportunities. The precision with which they do this may have a significant impact on student learning. Researchers recognized a correlation between higher intelligence scores and emergent responses (Dixon et al., 2018; Hayes et al., 2021).

A second way curriculum design can influence student learning is how individuals analyze and sequence target objectives. The scope and sequence, sometimes called curriculum mapping, has an impact on student learning. Curriculum mapping is a systematic and data-informed analysis that outlines short and long-term objectives for students to achieve (Arafeh, 2015). Curriculum mapping is commonly used to accredit curricula as it maps on to Common Core Standards (DeLuca & Bellara, 2013; Oliver et al., 2010). Strategic curriculum mapping may increase student learning by maximizing instructional time. When curriculum designers are creating the scope and sequence for addition, they may choose to teach prerequisite skills prior to more complex skills. For example, curriculum may be organized to teach single-digit by single-digit addition first, followed by single-digit by double-digit addition, and then double-digit by double-digit addition. This example can be applied across almost, if not all, target objectives from the CCSS. Strategic curriculum design would teach prerequisite skills first while creating a dynamic sequence that gradually builds off the previous objectives.

1.2 Applied Behavior Analysis in Education

Applied behavior analysis (ABA) is a scientific discipline that seeks to understand and benefit human behavior in a meaningful way by examining the relation between environment and behavior. In many cases, behavior analysts exist in the field of education because the fields have similar goals: to improve educational outcomes. Education outcomes may include ELA, mathematics, science, etc. Educational outcomes can be improved through teacher-student interactions and curriculum design/implementation. ABA is a "scientific approach for discovering environmental variables that reliably influence socially significant behavior toward developing a technology of behavior change that takes practical advantage of those discoveries" (Cooper et al., 2019, p. 2). That is, ABA seeks to understand the causal relationship between teacher-student interactions (or curriculum) and socially significant student achievement scores such as reading and mathematics performance. ABA components are often embedded into curriculum design, education systems, and classroom practices. Examples include but are not limited to data collection, data analysis procedures, decision protocols, operational definitions, and explicit instruction. (Greer, 2002).

ABA brings specific scientific tools to education. ABA research seeks to understand human behavior across the lifespan, which includes infants, elementary students, college students, adults, and individuals with disabilities (Greer & Ross, 2008; Ho, et al., 2021; Irie et al., 2019; Peiris et al., 2022; Unholz-Bowden et al., 2020). While many educators learn to teach effectively and efficiently, what that instruction looks like and what those expectations are vary greatly. This can create inconsistent teaching practices, which is a problem that may lead to a lack of student learning and a lack of knowledge regarding the students' learning. Behavior analysts sought an objective and measurable way to improve instruction.

Teaching as a science was a response to inconsistent student outcomes yielded from classrooms that implemented teaching as an art (Greer, 1991, Greer, 2002). Teachers' knowledge and implementation of the science is a core component of ABA in education (Greer, 2002). Teachers have a direct impact on student learning and teachers are typically trained in teaching as an art instead of teaching as a science. Teaching as an art was part of the problem in the education system (Greer, 1991). The students in the educational system were failing because the teachers were failing, and the teachers were failing because their supervisors were failing. Greer (1991) proposed that training teachers with the science of behavior analysis was the solution to saving our schools and making educational gains.

To understand human behavior, behavior analysts operationally define terms that constitute both teaching operations and student outcome variables. Operational definitions are definitions that explain a concept concisely, concretely, and objectively such that they can be counted and quantified (Cooper et al., 2019). They also define contingency conditions, which specify the conditions and behaviors that belong to the antecedent, behavior, and consequence of a target behavior. An operational definition may include the student pointing to, or circling the numerator after the instructor says, "show me the numerator." The operational definition would specify that the researcher present the antecedent in print, the student respond within 5 s, and the researcher deliver an immediate consequence in print.

Behavior analysts arrange environmental contingencies to produce socially significant target outcomes for their students. Once instruction is effective and produces the intended educational benefit (e.g., percentage of accuracy or fluency) educators can focus on the design of more efficient instruction. Different kinds of methods can produce more efficient outcomes than others. A method may be considered efficient if a teacher's resources are conserved relative to another method. For example, if in one case a teacher simply models correct responses and in another case the teacher delivers explicit instruction for many learning opportunities, the modeling would be more efficient because it requires fewer teacher resources.

For example, Cengher et al. (2018) conducted a systematic review of literature that compared two or more prompt-fading procedures. First, the researchers determined the effectiveness of prompt-fading and then compared stimulus prompt procedures to response-prompt procedures. Cengher et al. (2018) found that stimulus prompts[4] were more efficient than response-prompting[5] procedures and that response prompt procedures demonstrated variable efficiency, when compared to each other. Another example is when a teacher strategically selects to explain a skill on a board so the whole class can observe, rather than show each person individually. This is faster than teaching each type of response individually. They chose this because their students only needed one example modeled to all at one time so they all could acquire the emergent learning responses. This method maximized their instructional time. This required less effort for the teacher and may require less time to teach. Designing instruction and implementing teaching practices to be effective and more efficient than some other practice requires educators to implement more of these behavior analytic tactics. It is increasingly important given teachers' limited teaching time, planning time, student learning gaps, and limited resources that are available to teachers (Crawford et al., 2010).

Educators choose to employ a variety of behavior analytic research-based tactics to teach effectively and efficiently. These tactics can be performance, learning tactics, or a combination of both. As stated above, these may range from simple tactics like positive reinforcement (i.e.

---

[4] Stimulus prompts are prompts that change the antecedent stimuli in some way (i.e., changing the materials)
[5] Response-prompting are prompts that help an individual respond to a given antecedent. (i.e., physical prompting, modeling)

providing praise and other stimuli to increase the likelihood of that behavior in the future) to comprehensive curriculum programs that incorporate multiple methodologies and tactics. These tactics and packages of these strategies are embedded in a multitude of instructional programs and designs. Among these effective and efficient strategies are those that are regularly implemented in general/special education that have applied behavior analysis concepts embedded within them. These may be implemented intentionally or naïvely. These strategies may have influences or roots in multiple fields of study but remain to have attributes in ABA. Some of these strategies include *short-term objectives* (Vargas, 2020, p.53), *opportunities to respond* (Vargas, 2020, p.138), and *positive reinforcement* (Cooper et al., 2019, p.36). Short-term objectives may be defined as smaller, more achievable goals than a given long-term objective/goal (Greer & Ross, 2008). In education, short-term objectives have a broad definition that can mean a variety of different learning strategies are being used. Opportunities to respond is a strategy used in education and applied behavior analysis practices in which the teacher introduces opportunities for the learner to engage in content (Greer & Ross, 2008). Opportunities to respond is a strategic chance that an instructor arranges for a student. For example, a teacher may provide opportunities to respond for solving math problems, writing sentences, reading text, and answering comprehension questions. Practitioners implement opportunities to respond with the intent of helping the learner make steps to achieve the learning goal (Menzies et al., 2017; Whitney et al., 2017). Positive reinforcement is another strategy in education that is also in the applied behavior analysis field. Positive reinforcement occurs when a stimulus is added to the environment that increases the likelihood that a behavior will occur in the future (Allen, 2019). For example, when a teacher provides praise, they are adding praise to the environment, to increase their students' behavior of completing their work. Teachers may report that they use short-term objectives, opportunities to respond, and positive reinforcement regularly in their classrooms. However, these strategies are implemented differently depending on the school, teacher, and students. For example, although short-term objectives may be implemented, the frequency of implementation and type of scaffolds implemented may be different. In a scientific field, this can create confusion around what short-term objectives mean and their corresponding effectiveness (Bishop et al., 2016; Bishop et al., 2017; Koutsoftas & Srivastava, 2020).

A specific strategic approach that is implemented in education and utilizes applied behavior analysis strategies, philosophies, methodologies, measures, and tactics is the Comprehensive Application of Behavior Analysis to Schooling (CABAS®). The Comprehensive Application of Behavior Analysis to Schooling (CABAS®) is a systematic behavior analytic approach to schooling that applies the foundational concepts of verbal behavior and extends beyond basic applied behavior analysis. CABAS® was developed as a potential solution to the national education crisis (Greer, 1991) and incorporates Verbal Behavior Development Theory (VBDT). CABAS® is based on behavior analysis and applies research-based practices and tactics to all elements of schooling which include students, teachers, parents, mentors, supervisors, the CABAS® Board, and the University Training Programs (Greer, 1997a, 1997b, 1992; Singer-Dudek et al., 2010). CABAS® researchers place the teacher in the role of the strategic scientist and proposed the application of this idea may be the solution to the educational crisis that was happening in our country (Greer, 1992). The first CABAS® school had only three students, however, by the end of the year they had many more and their students' performance had greatly improved (Greer et al., 2006). More recently, Greer, et al. (2006) discussed the quality and standards of a CABAS® accredited school by discussing the performance of two CABAS® accredited schools. Data demonstrated CABAS® schools implement more learn units with faster learning rates in the more recent 10 years than in the school's first 10 years of operation. Additionally, their analysis reported similar findings that CABAS® school yield 4-7 times more learning than non- CABAS® schools (Greer, et al., 2002; Greer, 1997a; Greer, 1997b; Greer, et al., 1989; Lamm & Greer, 1992; Selinske et al., 1991). There are now CABAS® schools that service both special education and general education that span across the globe. CABAS® schools use a plethora of behavior analysis-based tactics and interventions to support their students' learning, such as: learn units (Albers & Greer, 1991), mastery criterion[6] (Richling et al., 2019; Wong et al., 2022), Teacher Performance Rate and Accuracy observations[7] (TPRA; Ingham & Greer, 1991), token economy[8] (Ayllon & Azrin, 1990), multiple exemplar instruction (Gilic & Greer, 2011), and instructional models (Hranchuk, et al. 2019; Maurilus, 2018).

Educational tactics/strategies may be improved through systematically implementing ABA strategies. When researchers implement tactics/strategies systematically and scientifically they can start to determine what strategies are effective for whom and improve those strategies for those for whom it is not effective. Some specific tactics that behavior analysts use in a systematic manner are a) *learn units* (Albers & Greer, 1991), (b) *instructional models* (Hranchuk et al., 2019), and (c) *multiple exemplar instruction* (Vargas, 2020, p. 210). Behavior analysts have implemented these strategies systematically, or strategically, which is supported by published research (Archer & Hughes, 2010; Csillag, 2016). In education, terms like "multiple exemplars" can be vague, whereas a behavior analyst would specify the topography and frequency at which those "multiple exemplars" are presented. Multiple exemplars are

[6] Mastery Criterion is a predetermined performance standard for accuracy and/or fluency (Vargas, 2020)
[7] Teacher Performance Rate and Accuracy (TPRA) Scale is a tool that measures teachers accuracy and rate of instructional presentations (Greer & Ross, 2008)
[8] Token economy is a reinforcement system that is used to increase motivation and prolong the delivery of reinforcement (Kazdin, 2012)

stimuli/opportunities of a target stimulus that rotate non-essential components of the target stimulus for example, teachers may change the color and font of spelling words. Operational definitions are used by behavior analysts to support systematic implementation of strategies.

Learn Units. Behavior analysts have contributed tactics that are often used in education. Learn units are a tool used to measure student learning and teacher performance (Albers & Greer, 1991; Ingham & Greer, 1992; Singer-Dudek, et al., 2010). They consist of a three-term contingency for the student and at least two interlocking contingencies for the teacher (Greer & Ross, 2008). A single three-term contingency consists of an antecedent, behavior, and consequence. Instructors use learn units to teach both form and function of stimuli. Form is the topography of the response. The "function" of behavior refers to the variables that control a given behavior and the antecedent/consequence relations that give rise to it. For example, someone may vocally say "dog," or show a picture of a dog in which a listener may let them pet their dog. The form of each instance is different, but they serve the same function. Then, they allow the student to respond and provide reinforcement for correct responses and corrections for incorrect responses.

In two experiments, Albers and Greer (1991) tested the effects of learn units on participants' correct and incorrect responses during math instruction. In Experiment 1, researchers implemented learn units to teach two participants math skills. Researchers presented three times as many learn units during intervention than in baseline. Results demonstrated that presenting more learn units increased the number of correct responses. In the second experiment, researchers systematically replicated the first experiment. Researchers implemented learn units during intervention but systematically rotated vocal and written antecedents. Results from Experiment 2 replicated those from Experiment 1. Educators can apply learn units to benefit their own practices by ensuring the antecedent, behavior, consequence contingencies are in place to support the learning goal. For example, educators may choose to use small group instruction to apply learn units to consequate each students' responses. This would include providing reinforcement for correct responses and corrections for incorrect responses.

Researchers and practitioners also use learn units to calculate learn units-to-criterion as a measure of the speed of acquisition (Singer-Dudek et al., 2010). Learn units-to-criterion is the number of instructional units that occurred for an individual's behavior to meet a predetermined mastery criterion. Learn units-to-criterion is a measure of a student's rate of learning, or how many learn units are needed to master educational objectives. Researchers can calculate learn units-to-criterion for any objective that they are delivering instruction for. For example, a student learning to draw a triangle at a criterion of 100% accuracy across three consecutive responses may require five learn units to demonstrate mastery. In this case, the student's learn units-to-criterion for this objective is five.

To account for environmental variables that may influence student learning, researchers use the learn unit-in-context to account for motivating operations, phylogeny, antecedent, behavior, consequence, setting events, and instructional history (Greer & Ross, 2008). Practitioners, first, ensure that the learn unit is in place. This is evaluated by measuring the fidelity of instruction to ensure learn units are run accurately. If learn units are run with fidelity, behavior analysts analyze the other variables in the learn unit-in-context (i.e., correct curriculum and materials, motivating operations, setting events, instructional history, phylogeny) to identify the source of the problem. Once the practitioner identifies the source of the problem, they implement the appropriate tactic to ensure the student learns the target skill (Greer & Ross, 2008).

Instructional Models. Instructional models, which are called instructional demonstration learn units in behavior analytic practice, are examples of a target skill that a teacher models for their students. Instructional models do not require any response effort for students other than attending. For example, when an instructor teaches a student to draw a picture of a unit fraction, a teacher may draw their own example of a unit fraction on the board first, then ask the student to draw a picture of a unit fraction. However, applied behavior analytic research has suggested that instructional models be implemented in tandem with other forms of instruction, like learn units. Hranchuk et al. (2018) extended the literature by comparing the effects of learn units vs. instructional models to inform instructors' planning and implementation of instructional models. Hranchuk et al. (2018) implemented a repeated AB design to compare a standard learn unit (SLU) condition to an instructional demonstration learn unit (IDLU) condition. The SLU condition consisted of a researcher delivering academic instruction by presenting a clear antecedent, observing, and recording the participant's response, then delivering the corresponding consequence (i.e., explicit instruction). Researchers delivered one of two consequences: reinforcement or a correction. Researchers conducted the SLU condition using 20-trial sessions. The IDLU condition was identical to the SLU condition except researchers began each IDLU condition with two examples/instructional models of the target content prior to beginning the standard learn units. Researchers did not expect the participant to respond to the two examples at the beginning of the session. Researchers found that participants learned twice as fast in the IDLU condition as when they were in the SLU condition. These data support instructional models as an effective and efficient tool in teaching all four of the participants in the present study across reading and math objectives. However, these results only occurred with four preschool participants. Hranchuk et al.'s (2018) findings support that instructional models are an effective and efficient method of instruction for the participants in the study. However, the researchers did not test the effects of instructional models for students across various populations.

Multiple Exemplar Instruction and Multiple Exemplar Training. LaFrance and Tarbox (2019) call for uniformity of terms when referring to multiple exemplar instruction (MEI) and multiple exemplar training (MET). While Verbal Behavior Development Theorists tend to select MEI to teach the joining of listener and speaker topographies, Relational Frame Theorists utilize MET to establish emergent relations (LaFrance & Tarbox, 2019). MEI is a procedure in which instructors quickly randomize the presentation of various verbal operants (LaFrance & Tarbox, 2019). Verbal Behavior Development Theorists typically use MEI across the listener and speaker topography. For example, when a child is learns letter sounds "m, a, s, t", the instructor would rotate across selection (e.g., point to and match) and production responses.

Researchers implement multiple exemplar instruction (MEI) to "produce interdependence between speaker and listener repertoires or between verbal operants within the same repertoire" (LaFrance & Tarbox, 2019). Greer and Ross (2008) articulate two types of MEI. The first is general case teaching, and the second teaches initially independent responses to come under joint stimulus control (Greer & Ross, 2008). The first type teaches *abstractions*[9] by rotating the non-essential components of the stimuli, while the second rotates multiple forms of the same stimulus so that independent responses come under the same stimulus control. This form of MEI aligns with Skinnerian principles and teaches the formation of concepts (Verdun et al. 2020, 2021).

---

[9] Abstractions are a "property or collection of properties that exists only in objects or situations, never alone" (Vargas, 2020). Abstractions do not have concrete examples.

MET is a procedure in which multiple exemplars of the same stimuli are rotated, which can include properties that share some correspondence or no correspondence at all (LaFrance & Tarbox, 2019). For example, an instructor may ask the student to respond when asked "how many" where the instructor rotates the quantity of a number (five dots) and the Arabic number five (5). It is important to use concise terminology to accurately study, replicate, and disseminate research that studies MEI and MET.

1.3 One Size Does Not Fit All

Educators may need to employ behavior analytic strategies to promote individualized instruction. Additionally, researchers found that having consistent emergent learning opportunities is related to increases in intelligence scores, with both typical students and students with disabilities (Cassidy et al., 2016). These findings have implications for curriculum designers, who may program curriculum to allow for ample emergent learning opportunities.

Learn units, instructional models, and multiple exemplar instruction are examples of educational practices that employ behavior analytic research-based tactics (Wasik & Iannone-Campbell, 2012). Although teachers may choose research-based strategies to implement in their classroom across one or many students, not every strategy works. Strategies do not always work in the same way for each student either. The effectiveness and efficiency of a tactic is determined by the individual's performance. However, performance often varies between individuals. Curriculum tends to target grade-level standards, which may not account for a student's preparation for that content or prerequisite repertoires. Teachers are often required to teach students with a wide range of abilities/repertoires without knowledge on how to select appropriate tactics (Belmonte-Mulhall & Harrison, 2022). To accommodate for more students, curriculum design and teacher-delivered instruction needs to be flexible and account for all factors that influence student success.

Effective teachers may incorporate applied behavior analytic tactics and apply them to their students' prerequisite skills. Individualized instruction and strategic curriculum design are two of these factors that may influence student success, both of which must incorporate their students' prerequisite repertoires. Prerequisites are the necessary skills needed to learn a new objective. For example, a student must first learn how to add single-digit numbers accurately before they learn how to add double-digit numbers. Students' prerequisite skills play an important role in effectiveness and efficiency of strategic instructional design and are critical to creating individualized instruction (Sun, et al., 2016; Zodik & Zavlasky, 2008). A teacher may use students' prerequisite skills to select what instructional lesson and determine the number of short-term objectives that individuals may need prior to the lesson to learn the target skills. Designing or modifying instruction that promotes an individuals' application of a learned skill/s across different situations may be defined as individualized instruction. Teachers may implement various strategies that aim to promote individualized instruction including small group instruction or direct instruction. Designing or modifying curricula to produce desired outcomes and use minimal resources and time may be defined as strategic instructional design (Dunlosky et al., 2013; Roediger et al., 2012). Most instructional design concepts are vaguely defined in educational practice. This makes selecting and implementing strategies/tactics more difficult.

There is limited research that investigates the use of prerequisites to individualize instruction. Most published research demonstrates something that works well and should be used more. However, not every strategy is appropriate for everyone and measuring prerequisite repertoires in a more detailed way may help inform that instruction (Belmonte-Mulhall &

Harrison, 2022). How teachers “use” prerequisite skills to inform the selection and implementation of teaching strategies remains ambiguous. Prerequisite skills may need to be considered to select effective and efficient instruction. To be effective and efficient, teachers need to use as few resources as possible and their students still acquire the target skill. Advanced learners may require less instruction and other learners may require more instruction. There is not sufficient research investigating instructional design strategies based on prerequisite skills for general education practitioners to implement truly individualized instruction. Behavior analysts may use operational definitions to study different types of behavior, which include verbal behavior in general education and special education settings. Utilizing these operational definitions outlined by *Verbal Behavior Development Theory* (Greer, 2002; Greer & Ross, 2008) may help us understand the learning needs of our students as well as how to select appropriate strategies to teach them.

1.4 Verbal Behavior

Verbal behavior is the study of behavior that relates to communication between persons or socially mediated behaviors (Skinner, 1957). Skinner’s *Verbal Behavior* (1957) has led to decades of research and development in understanding human behavior, specifically, human learning (Greer, 2008; Greer & Keohane, 2005; Greer & Ross, 2008). Verbal behavior studies human communication, which allows for implications across a variety of fields like special education and general education. These six verbal operants; (a) *mand*, (b) *echoic*, (c) t*extual response,* (d) *tact*, (e) *autoclitics*, and (f) *intraverbals*; are studied as dependent variables in terms of their function, or effect on another person’s behavior (See Table 1 for definitions and examples of the six verbal operants

Table 1

Skinner's (1957) six basic verbal operants with their antecedent, behavior, and consequence, as well as examples.

Basic Verbal Operant		Antecedent	Behavior	Consequence
Mand	Definition	Establishing operation	Any behavior that is followed by the delivery of the reinforcer	Delivery of the reinforcer
	Example	Deprivation of water	Student says: "water"	Teacher: delivers water
Echoic	Definition	Auditory stimulus	Point-to-point vocal emulation	Generalized reinforcement
	Example	Teacher: "water"	Student: "water"	Teacher: "Yes, exactly"
Textual Response	Definition	Printed word	Point-to-point vocal response	Generalized reinforcement
	Example	Printed word "water"	Student vocally says "water"	Teacher: "Yes, good job!"
Tact	Definition	Nonverbal stimulus	Name stimulus	Generalized reinforcement
	Example	Teacher and student walking by a puddle	Student says "water"	Teacher says "Yes, it is"
Autoclitics	Definition	Presence of a listener	Any behavior that modifies the effect of the verbal behavior on the listener	The behavior of the listener changes as a function of the verbal behavior and modifier
	Example	Teacher and student walking together	Student says: "I have no water"	Teacher says: "Okay, are you thirsty?
Intraverbals	Definition	Vocal or printed stimulus	Verbal response that does not have point-to-point correspondence with antecedent	Generalized reinforcement
	Example	Teacher says, "What is that?"	Student says "water"	Teacher says "Yes, it is"

A mand is a verbal operant in which the antecedent is a form of deprivation or aversive stimulus and the behavior results in the delivery of the desired reinforcer. An example of a mand is a person who has not had food in two days says to another person “food,” in which the food is delivered to the original person. Mands are not defined by their formal properties but are emitted in multiple forms based on their function (e.g., words, gestures, or written sign language). An echoic is a verbal operant in which the antecedent is a vocal stimulus, and the behavior is a response that emulates the original antecedent with point-to-point correspondence for natural reinforcement. An example of an echoic would include if a parent said “dog,” and the child echoed by saying the word “dog” with point-to-point correspondence to recruit social reinforcement. This means every syllable was emitted the same as the vocal antecedent and occurred as either a mand or tact function. A textual response is a verbal operant that occurs when print stimuli evoke a vocal response that shares point-to-point correspondence with the printed word emitted under conditions to receive natural reinforcement. An example of a textual response is if a child sees the printed word “cat” and vocally says the word cat. A tact is a verbal operant in which the antecedent is a stimulus in the environment and the behavior of identifying that stimulus occurs to recruit social reinforcement. An example of a tact is when a child is walking with an adult and points to a bird and says “bird” in the presence of the adult in which the adult may say “yes that is a bird.” Autoclitics modify the effect of other verbal operants that include mands or tacts. Autoclitics function to qualify, affirm, negate, specify, or affect how another verbal operant acts on the listener. This differs from structural components of verbal operants. An intraverbal is a verbal operant in which the antecedent and response do not share point-to-point correspondence and occur under generalized reinforcement conditions (Skinner, 1957, p.71). Intraverbals account for a larger portion of our verbal behavior and range in complexity (Connie et al., 2023). An example of an intraverbal is when a teacher asks their student "What is 5 plus 5?" and the student responds. This response could be "10," "I do not know," "7," etc. and be considered an intraverbal. Teachers frequently use intraverbals to teach students. They may choose to utilize intraverbals in multiple ways which may include requiring the student to respond as a listener with a selection response (e.g., multiple-choice questions) or as a speaker with a production response (e.g., writing down the answer). Standardized assessments are filled with intraverbal responses which often alternate between multiple choice and written responses. Based on these expectations, it is critical for teachers to teach their students how to interact with the content accordingly.

1.5 Verbal Behavior Development Theory

Verbal Behavior Development Theory (VBDT) is the study of the development of language from a behavior analytic perspective (Greer, 2008; Greer & Keohane, 2005; Greer & Ross, 2008). VBDT is heavily influenced by Skinner's (1957) Verbal Behavior and focuses on the function of behavior. It is not about the form or topography of a response but the purpose for which the response is emitted. VBDT researchers have focused on language development. VBDT focuses primarily on the establishment of developmental cusps and identifying reinforcers within those cusps (Greer & Keohane, 2006; Greer & Ross, 2008; Pohl, et al., 2018). VBDT has identified the developmental progress that leads to behavioral cusps which include higher order emergent learning repertoires. A behavioral cusp is a developmental milestone in which an individual accesses the environment in new ways, learns faster than before the cusp was established, and, in some cases, learns in new ways (Fryling et al., 2020; Rosales-Ruiz & Baer, 1997). For example, generalized imitation is considered a preverbal cusp because the child can access the environment in new ways, learn faster, and learn in a new way (i.e., see-do). Early cusps within verbal behavior development fall under four categories: *preverbal*, *listener*, *speaker*, and *joining of listener and speaker* (Greer & Keohane, 2006). See Table 2 for a description of the four verbal behavior development categories.

Table 2

Displays verbal behavior development stages, descriptions, assessments with citations per stage (based on VBDA pinwheel)

Stage	Description	Cusps within Stage	Citations
Preverbal	Individuals who are preverbal rely on others for survival. These individuals have not entered the verbal community and may not have observing responses for environmental stimuli (Greer & Keohane, 2006).	Conditioned Reinforcement for 2D Stimuli	Dinsmoor, 1983; Longano & Greer, 2006; Tsai & Greer, 2006
		Conditioned Reinforcement for 3D Stimuli	Du et al., 2015
		Conditioned Reinforcement for Faces	Maffei et al., 2014
		Conditioned Reinforcement for Voices	Decasper & Spence, 1987; Greer et al., 2011
		Conditioned Reinforcement for Hear-Say	Ross & Greer, 2003; Tsiouri & Greer, 2003
		Conditioned Reinforcement for See-Do	Baer & Shermin, 1964; Ross & Greer, 2003
Listener	The behavior of individuals with listener status are mediated by speakers. Listeners can follow vocal directions and are less dependent on others for survival than preverbal learners (Greer & Keohane, 2006).	Listener Literacy	Rosales-Ruiz & Baer, 1996; Speckman-Collins et al., 2007
		Unidirectional Naming (UniN)	Greer et al., 2005; Greer & Ross, 2008; Horne & Lowe, 1996
Speaker	Individuals with speaker status can mediate the behavior of listeners in their environment. These individuals can access and contribute to their environment in many more ways than listeners by emitted mands, tacts, and echoics (Greer & Keohane, 2006).	Mands	Greer et al., 2017; Greer & Ross, 2008; Ross et al., 2006; Twyman, 1996a, 1996b
		Tacts	Greer & Du, 2010; Greenburg et al., 2014; Shmelzkopf et al., 2017; Twyman, 1996a, 1996b
		Echoics	Choi et al., 2015
		Transformation of MC Across Mands and Tacts	Greer & Ross, 2008; Nirgudkar, 2005; Nuzzolo-Gomez & Greer, 2004; Schmelzkopf et al., 2017
Joint Listener and Speaker	Individuals with joint listener and speaker status emit emergent relations across the listener and speaker topographies. These individuals can engage in conversational units with themselves and others. At this stage, an individual has entered the verbal community (Greer & Keohane, 2006)	Incidental Bidirectional Naming (Inc-BiN)	Hawkins et al., 2018; Horne & Lowe, 1996; Greer & Ross, 2008
		Conv. Units b/w an Adult and a Child	Lodhi & Greer, 1989; Donley & Greer, 1993; Greer & Keohane, 2006
		Conv. Units b/w children	Donley & Greer, 1993; Greer & Speckman, 2009; Lowe et al., 2002
		Say-Do Correspondence	Baer et al., 1988
		Self-Talk Conv. Units in Solitary Play	Donley & Greer, 1993; Greer & Speckman, 2009; Lowe et al., 2002
		Extensions of Inc-BiN	Cahill & Greer, 2014; Greer & Du, 2015a, 2015b; Sivaraman & Bhabu, 2018

Preverbal learners rely on others for survival (Greer & Keohane, 2006). Pre-listeners do not access the social community. Newborn babies are preverbal because they rely on others to feed them, change them, and take care of them. Typically, preverbal children may have some observing responses. Observing responses occur when an individual orients toward a given stimulus. For example, observing responses may include looking at someone moving across the room or turning their head toward a loud noise. Typically, an individual develops a listener repertoire before they develop a speaker repertoire (Greer, 2008; Greer & Keohane, 2005; Greer & Ross, 2008).

Listener cusps include selecting out voices, discriminating between positive and negative exemplars of words, the capacity for sameness for words with acoustic properties, and basic listener literacy (Greer & Ross, 2008). For example, an individual with listener skills selects out voices and may orient toward a person speaking. Once a learner has acquired basic listener behavior, they begin to socially interact with their community by performing verbally governed behavior. Listeners can follow instructions and learn from consequences (Greer & Keohane, 2006).

Once a learner has acquired speaker status, they can govern the behavior of listeners (Greer & Keohane, 2006). Speaker cusps include echoing words to recruit social attention and to access a reinforcer, and producing a response when asked a question. For example, an individual with speaker skills may learn to name an item in circumstances to recruit social attention and then begin to name that same item to gain access to that item. Up until this point, an individual may only acquire verbal operants through direct instruction. If an individual's listener and speaker repertoires continue to develop, they may acquire *higher-order learning repertoires*[10].

Once an individual's listener and speaker repertoires develop independently, they may develop further so that the listener and speaker repertoires become interdependent. This means individuals may acquire verbal operants without direct instruction or acquire language incidentally. Individuals may have speaker/listener exchanges and develop cusps: sequilics and conversational units (Greer & Keohane, 2006). For example, an individual with joined listener speaker repertoires will engage in a conversational unit if they speak to a listener and say, "How are you?," the listener responds and says "I'm good. How are you?," and the original speaker responds and says, "I'm good." Next, individuals may develop joint stimulus control across listener and speaker topographies. Hawkins et al., (2018) proposes categorizing Common Bidirectional Naming in two subcategories: Bidirectional Naming and Incidental Bidirectional Naming. Bidirectional Naming is when the listener and speaker repertoire are interdependent of each other after either listener or speaker instruction. Incidental Bidirectional Naming is when the listener and speaker repertoire are interdependent of each other without direct instruction in either (Gilmore et al., 2024; Hawkins et al, 2018). Inc-BiN is a developmental cusp in which and an individual's listener and speaker repertoires are interdependent of each other, and an individual acquires listener/speaker repertoires without direct instruction. For example, and individual who directly learns how to select a "ball" among other objects, would also vocally tact a "ball" independently without being directly taught. Individuals who do not have any incidental Naming repertoires are classified as preverbal.

[10] Higher order operants are overarching operants that occur following multiple exposures to experiences or stimuli. Higher-order operants occur when two independent response topographies become joined (i.e., Incidental-Bidirectional Naming; Greer & Ross, 2008)

When an individual’s listener and speaker repertoire are not yet developed (preverbal), an individual is considered to have Pre-Unidirectional Naming (Pre-UniN). Pre-UniN means an individual cannot emit emergent relations across listener/speaker topographies. Individuals with Pre-UniN learn at a slower rate than individuals with the more incidental bidirectional listener and speaker responses. A student who has a listener repertoire has entered the verbal community and may emit rule-governed behavior and attend to the behaviors of others in their community because the behavior of the individual has come under the control of spoken words (Choi et al., 2015). Once an individual has listener and speaker repertoires they may join, which is when a student may learn incidentally (Greer & Ross, 2008; Pohl et al., 2018). Once the listener repertoire is fully established this is called the listener half of Naming or Unidirectional Naming (UniN; Greer et al., 2005). Individuals with Inc-UniN can learn incidentally through observing models. Individuals with Inc-UniN can emit emergent listener responses after being directly taught speaker relations. For example, if a student is directly taught how to write their name, they can later select their name in a field of/or array of words without being directly taught. Individuals with Inc-UniN learn at a faster rate than those with Pre-UniN but at a slower rate than those with the listener and speaker half of Naming (Hawkins et al., 2018; Greer et al., 2017). A student who has speaker cusps can affect the behavior of others through their own speaker behavior (Eby & Greer, 2017). VBDT researchers define the joining of the listener and speaker repertoires as Incidental-Bidirectional Naming (Inc-BiN; Horne & Lowe, 1996). Inc-BiN is a verbal behavior development repertoire that allows an individual to learn incidentally across listener and speaker topographies, (e.g., learn from a tact model without direct training or programmed reinforcement for either the listener or the speaker response). Individuals with Inc-BiN learn at a faster rate than those with Inc-UniN and Pre-UniN (Greer et al., 2011).

Chapter 2: Incidental Bidirectional Naming and Academic Performance

2.1 Brief Literature Review

Academic performance on standardized tests has been used to indicate what students have learned in relation to normative samples. These assessments can also serve as a baseline for evaluating how instruction changes academic performance. Academic assessments provide a point of reference of where individual students would fall in reference to grade-level standards. For example, academic assessments may indicate if an individual was meeting, exceeding, or behind grade-level standards. These scores are often based on proficiencies on the Common Core standard for each grade. Students' scores on academic assessments may be used to hold teachers, schools, and school districts accountable for teaching their students the necessary skills. Although important, these scores do not offer teachers strategies for how to teach students who perform below average. Incorporating individuals' degree of Inc-BiN may help teachers determine what instruction is appropriate for whom by understanding how their students acquire (or don't) incidental language.

Researchers have begun to explain individuals' learning according to their degree of Inc-BiN (Greer, et al., 2011; Morgan, et al., 2021). Researchers have induced Inc-BiN and observed that participants learn at a faster rate after Inc-BiN had been induced (Hotchkiss & Fienup, 2020). Beyond general increases in performance, researchers have examined effects of Inc-BiN on specific academic repertoires. In an unpublished dissertation, Baldonado (2021) extended this research by measuring the associations, differences, and predictive value between degree of Inc-BiN and reading comprehension. Researchers measured reading comprehension with two main variables: *iReady*® Reading literature comprehension scores and the Woodcock Johnson

Inventory for Vocabulary (WJIV). Researchers measured degree of Inc-BiN using Inc-BiN probes. These probes had three components: Naming experience, listener responses, and speaker responses. Researchers measured degree of Inc-BiN by reporting the sum of the listener responses and sum of the speaker responses. Researchers classified individuals with less than 80% correct listener responses as Pre-listener/Pre-UniN. Researchers classified individuals with greater than 80% correct listener responses as Inc-UniN. Lastly, researchers classified individuals with greater than 80% correct responses on both listener and speaker components as Inc-BiN. They measured degree of Inc-BiN for both *familiar* and *unfamiliar stimuli*. Familiar stimuli includes stimuli that may be found in the participants' environment. For example, Baldonado (2021) used cartoon characters. Unfamiliar stimuli includes stimuli that the participants, likely, had not encountered. For example, unfamiliar stimuli for American participants may include Chinese written characters. Baldonado (2021) found that the strongest positive correlations were between reading comprehension and Inc-BiN for familiar stimuli and reading comprehension as well as Inc-UniN and unfamiliar stimuli. Baldonado's (2021) findings uniquely contribute to the literature by measuring degree of Inc-BiN for both familiar and unfamiliar stimuli. This line of research could be extended by measuring reading and math performance simultaneously, as opposed to just reading.

In an unpublished dissertation, Abdool-Ghany and& Fienup (in press) implemented two conditions across 14 preschool participants with varying degrees of Naming (Pre-Unidirectional Naming, Unidirectional Naming, and Bidirectional Naming). Researchers studied the effects of directly reinforcing either listener (point to) or speaker (tact) responses and testing if participants acquired the untaught topography (listener or speaker). Abdool-Ghany and Fienup (in press) concluded that participants with Inc-BiN emit emergent speaker responses (after direct consequences for listener responses), participants with UniN emit emergent listener responses (after direct consequences for speaker responses), and participants with Pre-UniN did not consistently emit emergent listener or speaker relations. This research added applied research that was consistent with Naming theory (Horne & Lowe, 1996; Hawkins et al., 2018). Researchers call for future research to account for the low sample size and, potentially, replicate their findings. Additionally, researchers used Naming assessment data but did not use other measures to test emergent responses across other academic responses.

In another unpublished dissertation, Friedman (2021) investigated the relation between Bidirectional Naming and derived relational responding. In multiple experiments, researchers analyzed the relation between production of echoics on Naming probes and their relation to untaught listener/speaker responses on Naming probes. Additionally, researchers also investigated relations between auditory and visual stimuli presentations with a 1s delay between auditory and visual stimuli. These data did not demonstrate significant correlations. In Experiment 4, researchers tested if there was a functional relation between Naming and arbitrarily applicable derived relations. Once a higher degree of Naming was established, four out of six participants demonstrated increases in mutually entailed relations while 2 out of the 6 participants demonstrated increases in combinatorally entailed relations. These data are significant in demonstrating a positive correlation between Naming and emergent responses. Like, Abdool-Ghany and Fienup (in press) this research contributes to the field of verbal behavior and Naming theory but could increase generalizability by incorporating other academic variables that have some social validity.

Incidental Bidirectional Naming (Inc-BiN) appears to predict general increases in learning rate, emergent relations, and reading comprehension (Abdool-Ghany & Fienup, in press;

Baldonado, 2020; Friedman, 2021; Morgan et al., 2021). In unpublished dissertations, Abdool-Ghany and Fienup (in press) and Friedman (2021) found that preschool participants with Inc-BiN learn at a faster rate and emit greater emergent relations than participants with lower degrees of Inc-BiN. Baldonado (2021) found a significant positive relation between Inc-BiN and reading comprehension scores, while Morgan et al. (2021) found a significant correlation between Inc-BiN and other emergent responses. Although these researchers investigated the relation between Inc-BiN and other variables in different ways, each summarized the positive correlation between Inc-BiN and the other variables in question.

Previous research has investigated instructional methods that will help students learn specific skills and investigated learning repertoires that will help students learn how to learn (Corwin & Greer, 2017; Fiorile & Greer, 2007; Gilic & Greer, 2011). One of the branches of research has focused on Inc-BiN, which is a learning repertoire, cusp, that has been linked to rate of learning. Researchers have induced Inc-UniN and Inc-BiN and noticed increased learning as a function of the newly acquired Inc-BiN repertoire (Corwin & Greer, 2017; Fiorile & Greer, 2007; Gilic & Greer, 2011; Greer, et al., 2011; Greer, et al., 2005). Although this research is critical to understanding Inc-BiN, to date, limited research has investigated the relation between degree of Inc-BiN and academic performance (Baldonado, 2020). Most of the relevant research has studied Bidirectional Naming, with less focus on Incidental Bidirectional Naming (Inc-BiN) and focused less on incorporating academic performance measures in a small group setting.

2.1.1 Rationale for Experiment 1

The purpose of Experiment 1 was to expand the literature and start to fill a gap by investigating the relation between academic performance (*iReady*® Diagnostics) and degree of Inc-UniN and Inc-BiN for unfamiliar stimuli. Researchers reported participants' *iReady*®

Diagnostic scores for mathematics and reading and conducted two unfamiliar brief Inc-BiN probes. Scores were collected for 41 4^{th}- and 5^{th}-grade participants who attended either a public school or charter school outside a major metropolitan area.

Experiment 1

2.2 Methods

2.2.1 Participants

The researcher selected 41 fourth- and fifth-grade participants from two schools outside major metropolitan areas. The researcher selected fourth- and fifth-grade participants who demonstrated instructional readiness. Individuals who were frequently pulled to go to the principal's office were excluded from the study. One school was a public school and one was a charter school. Participants were in a general education classroom that implemented components of Comprehensive Application of Behavior Analysis to Schooling (CABAS®) and Accelerated Independent Learner (AIL®; Greer, 1994) educational models (Greer, 2002). In this model, instructors implemented applied behavior analysis methodology to individualize the Common Core Standards to fit the needs of each participant. Participants ranged from 9-year-olds to 11-year-olds. Eighteen participants were female and twenty-three were male. Participants' parents/guardians self-reported their race/ethnicity to the participants' school: 10 participants identified as Hispanic, four identified as Mixed, twenty-four identified as White, three identified as black, and one identified as Asian. As part of the participants' school-wide assessments, each participant completed an *i-Ready® K-12 Adaptive* Reading and Math Diagnostic, a computer-based, adaptive learning and assessment tool (Curriculum Associates, 2014), in both reading and math. On the *i-Ready®* reading diagnostic participants' scores ranged from a 409 to 659, with a mean of 526. These scores mean that participants ranged from a grade K to a grade 7, with a mean of grade 3.1. On the *i-Ready®* math diagnostic participants ranged from a 397 to 532, with a mean of 446. These scores mean that participants ranged from a grade K to a grade 5, with a mean of grade level 2.9. The researcher measured participants' Inc-BiN repertoires for unfamiliar stimuli using researcher-created assessments. Twenty-seven participants demonstrated Pre-Unidirectional Naming (Pre-UniN), ten participants demonstrated Incidental-Unidirectional Naming (Inc-UniN), and four participants demonstrated Incidental Bidirectional Naming (Inc-BiN). Summarized participant information is displayed in Table 3 and individual participant data are displayed in Table 4.

Table 3

Description of Participants for Experiment 1 (sex, race, learning classification, Inc-BiN classification)

Variable	*N*	Percent
Sex	M = 23	M = 56%
	F = 18	F = 43%
Race	W = 24	W = 58%
	M = 4	M = 9%
	H = 10	H = 24%
	B = 3	B = 7%
Learning Classification	IEP = 11	IEP = 26%
	504 = 10	504 = 24%
	ESL = 1	ESL = 2%
	I & RS = 1	I & RS = 2%
Inc-BiN Classification	Pre-UniN = 28	Pre-UniN = 68%
	Inc-UniN = 9	Inc-UniN = 21%
	Inc-BiN = 4	Inc-BiN = 9%

Note. W = white, M = Mixed, H = Hispanic, B = Black, A = Asian, IEP = Individualized Education Plan, ESL, English as Second Language, I&RS = Intervention and Referral service

Table 4

Individual Participant Information

Participant	Age	Gender	Race	Learning Classification	Degree of Naming: Unfamiliar Classification	Total Listener (20)	Total Speaker (20)	*iReady*® Math Diagnostic Scaled Score	Grade Level	*iReady*® Reading Diagnostic Scaled Score	Grade Level
1	9	M	B	I&RS	Pre-UniN	7	0	424	2	514	3
2	10	M	W		Pre-UniN	15	0	433	2	440	1
3	11	M	W	504	Pre-UniN	14	0	464	4	573	4
4	10	F	B	504	Pre-UniN	5	0	417	2	457	1
5	10	F	W	IEP	Pre-UniN	12	0	410	1	460	1
6	11	F	W	IEP	Pre-UniN	8	0	403	1	437	1
7	10	M	W		Pre-UniN	5	0	444	3	511	3
8	9	F	M	504	Pre-UniN	9	1	397	1	409	K
9	9	F	H		Pre-UniN	9	2	461	3	559	4
10	9	F	H	IEP, ESL	Pre-UniN	11	2	465	4	542	3
11	9	M	H	IEP	Pre-UniN	11	2	414	2	509	3
12	10	M	W	IEP	Pre-UniN	10	2	415	2	422	1
13	9	M	W		Pre-UniN	10	2	423	2	409	K
14	9	F	W	504	Pre-UniN	9	2	422	2	529	3
15	9	F	B	IEP	Pre-UniN	15	2	407	1	474	2
16	11	F	W	504	Pre-UniN	11	3	452	4	515	3
17	11	M	W		Inc-UniN	17	3	449	3	488	2
18	11	F	W	504	Pre-UniN	8	3	428	2	513	3
19	11	M	W	IEP	Pre-UniN	11	4	439	2	542	4
20	10	F	W	IEP	Pre-UniN	6	5	436	3	521	3

21	9	F	W	504	Pre-UniN	9	5	426	2	430	1
22	10	M	H	IEP	Pre-UniN	10	5	412	1	449	1
23	9	M	W		Inc-UniN	19	6	481	4	555	4
24	10	M	M		Pre-UniN	12	6	434	3	505	3
25	9	M	W	IEP	Inc-UniN	17	6	435	3	570	4
26	9	F	H		Pre-UniN	12	6	439	2	572	4
27	10	M	H	IEP	Pre-UniN	15	6	425	2	475	2
28	10	F	W		Pre-UniN	7	7	477	4	644	6
29	9	M	W	504	Inc-UniN	19	7	434	3	436	1
30	9	F	H		Inc-UniN	17	7	418	2	561	4
31	9	F	H		Pre-UniN	14	9	492	4	623	4
32	10	M	M		Pre-UniN	13	9	481	4	604	4
33	9	M	W		Inc-UniN	17	10	479	4	583	4
34	10	M	H		Inc-UniN	19	10	486	4	587	4
35	9	F	H	504	Pre-UniN	14	11	474	4	529	3
36	11	M	W		Inc-UniN	16	11	379	K	498	3
37	10	M	M	504	Inc-UniN	16	11	499	5	613	4
38	9	F	W		Inc-BiN	20	17	474	4	616	4
39	9	M	A		Inc-BiN	20	18	532	5	659	7
40	10	M	W		Inc-BiN	20	18	489	5	613	4
41	9	M	W		Inc-BiN	20	20	516	5	648	6

2.2.2 Setting and Materials

Brief Incidental-Bidirectional Naming Assessment. The researcher conducted Inc-BiN assessments and academic assessments in the participants' typical school 4th or 5th grade classroom. Participants completed assessments independently either at their individual desks, a U-table, or at the researcher's desk. The researcher ensured that the participant completed assessments with no other participants within three feet of them.

The researcher conducted Inc-BiN assessments for unfamiliar stimuli, which can be defined as arbitrary stimuli that did not resemble other stimuli in the natural environment (i.e., Greek symbols, Chinese characters). All stimuli were presented in black and white. Inc-BiN assessments were identical other than the selection of the stimuli (See Table 5 for the stimuli used).

Each Naming experience was displayed on a laptop with a 13.3-inch (33.8 cm) screen. Stimuli on the screen measured 4-inches by 4-inches (10.2 cm by 10.2 cm). Stimuli were presented in isolation on a slide of GoogleSlides© for a total of 20 slides in the Naming experience. Listener probes consisted of 10 total slides. Each slide had three stimuli presented horizontally across the slide. Speaker probes consisted of 10 total slides. The researcher interspersed all stimuli so that no stimulus was repeated in back-to-back slides.

Table 5

Unfamiliar Inc-BiN stimuli

Unfamiliar			
Set 1	Set 2	Set 3	Set 4

i-Ready® K-12 Adaptive Reading and Math Diagnostic. The *i-Ready*® Diagnostic includes multiple choice and production questions. It measures participants' academic performance across multiple domains for reading (i.e. phonics, phonemic awareness, high-frequency words, vocabulary, comprehension of literature, and comprehension of informational texts) and math (i.e. numbers and operations, algebra and algebraic thinking, measurement and data, and geometry). For example, a 4th-grade reading question may include the participant reading a fiction passage and answering wh- questions about the text such as "What is the name of the main character?" An example of a math question may include a long division word problem such as "Olivia has 4,683 pencils. Each pencil case can hold 100 pencils. How many pencil cases does she need?"

2.2.3 Measurement

Brief Incidental-Bidirectional Naming Assessment. The researcher used Incidental Bidirectional Naming (Inc-BiN) Assessments to measure participants' degree of Inc-UniN and Inc-BiN. Each Inc-BiN assessments consisted of a Naming experience, listener probe, and speaker probe.

During each listener probe, the researcher presented 10 opportunities for the participant to respond to a field of three to a vocal antecedent. The researcher recorded a correct response (plus) if the participant pointed to the stimulus that corresponded to the researcher's vocal antecedent, "point to ___," within 5 s. For example, if the researcher said, "point to chala" and the participant pointed the picture of the emu within 5 s, the researcher recorded a correct response. The researcher recorded an incorrect response (minus) if the participant pointed to a stimulus that did not correspond to the vocal antecedent, responded after 5 s, or did not respond.

The researcher summed the number of correct listener responses and reported that as the listener score for that Inc-BiN assessment.

During each speaker probe, the researcher presented 10 total opportunities for the participant to respond to a field of three with a vocal antecedent. The researcher recorded a correct response (plus) if the participant vocally named the single stimulus within 5 s. For example, if the researcher presented a picture of an emu and the participant vocally said "emu," the researcher recorded a correct response. The researcher recorded an incorrect response (minus) if the participant emitted a vocal response that did not correspond to the given picture, responded after 5 s, or did not respond. The researcher summed the number of correct speaker responses and reported that as the speaker score for that Inc-BiN assessment.

If the participant emitted less than 80% correct listener responses and less than 80% correct speaker responses, they were placed in the Pre-Unidirectional Naming (Pre-UniN) category. If the participant emitted greater than 79% listener responses, but less than 80% speaker responses, they were placed in the Incidental-Unidirectional Naming category (Inc-UniN). If the participants demonstrated 80% or more correct listener responses and 80% or more correct speaker responses, they were placed in the Incidental Bidirectional Naming (Inc-BiN) category.

i-Ready® K-12 Adaptive Reading and Math Diagnostic. The *i-Ready® K-12 Adaptive Reading and Math Diagnostic* is an assessment tool that measures individuals' performance on the Common Core Standards. Participants complete the *i-Ready®* Diagnostic at the beginning, middle and end of each school year. Once participants complete each Diagnostic, *i-Ready®* uses the data from the participants' responses to generate a scaled score. Each score on a scaled score is on a single continuum, meaning, the researcher can compare scores across grade levels.

2.2.4 Procedure

Brief Incidental-Bidirectional Naming Assessment. The researcher sat within 3-feet (.91 m) of each participant and presented Inc-BiN assessments with their computer placed directly in front of the participant. The researcher conducted Brief Inc-BiN assessments with one participant at a time. The researcher began Inc-BiN assessments by asking the participant to come and sit next to them. The researcher conducted the Naming experience first. The Naming experience consisted of the researcher presenting the five stimuli, four times each in a random order, for a total of 20 trials. The researcher ensured the participant was attending and presented a single stimulus on their computer screen and vocally tacted the symbol. The researcher remained on each slide for up to 5 s. During the Naming Experience, participants were not required to respond. Once the researcher and participant completed the 20 trials, the researcher instructed the participant to return to their desk. After a minimum of two hours and a maximum of 24 hours, the researcher asked the participants to come sit at a teacher's table with them for the listener and speaker probes.

During the listener probes the researcher displayed one target stimulus with two non-target stimuli in a horizonal row on a slide. The non-target stimuli were stimuli within the given Inc-BiN set. The researcher also presented a vocal antecedent of "Point to ____." Participants pointed to one of the three stimuli, the researcher recorded a correct or incorrect response and moved to the next slide. The researcher did not consequate any responses. The researcher continued this procedure for each stimulus, two times each, for a total of ten trials. Next, the researcher immediately began the speaker probes. During the speaker probes, the researcher displayed one target stimulus centered on the screen. There was no vocal antecedent. The researcher waited up to 5 s for the participants to tact the stimulus or say, "I don't know." Next, the researcher recorded if the participant emitted a correct or incorrect response and moved to the next slide. The researcher summed the number of correct listener and speaker responses and reported both for the corresponding probe. Three such identical Inc-BiN assessments were conducted with a novel set of stimuli each time.

i-Ready® K-12 Adaptive Reading and Math Diagnostic. Each participant completed two *i-Ready*® Diagnostic assessments; one for Reading and one for Mathematics. Each *i-Ready*® Diagnostic required participants to complete multiple types of questions that corresponded to the various domains and Common Core Standards. Participants sat at one of the following locations: their individual desks, teacher U-table, or teacher's desk. No participant was within 3 ft of each other. Participants took frequent breaks and completed each of the *i-Ready*® Diagnostics within 1-2 school days. The *i-Ready*® Diagnostic is a standardized assessment tool that requires participants to answer questions in different academic domains.

2.2.5 Statistical Analysis

The researcher analyzed the data by running spearman correlations that analyzed the relation between participants' degree of Inc-BiN and academic performance. The researcher measured the degree of Inc-BiN for unfamiliar stimuli for each participant by conducting two brief Inc-BiN assessments and reported scores for their listener and speaker repertoires. The researcher totaled the number of listener responses and called that their degree of Inc-UniN. The researcher totaled the number of speaker responses and called that their degree of Inc-BiN. The researcher measured academic performance by reporting the participants' overall scaled *i-Ready*® Diagnostic scores for reading and mathematics. The researcher conducted nonparametric spearman correlations to test the association between math scores and degree of Inc-UniN, math scores and degree of Inc-BiN, reading scores and Inc-UniN, and reading scores and Inc-BiN.

2.2.6 Interobserver Agreement

Two researchers independently collected data on participants' Brief Inc-BiN probe responses. The researcher calculated Interobserver Agreement (IOA) for responses to listener and speaker responses to Inc-BiN probes. The researcher used trial-by-trial agreement and divided the number of agreed upon trials by the total number of trials. They divided the number of agreed upon trials by the total number of opportunities to respond and then multiplied the result by one hundred to make a percentage. The researcher conducted IOA for 50% of Inc-BiN assessments, with 99% agreement with a range of 95% and 100% agreement.

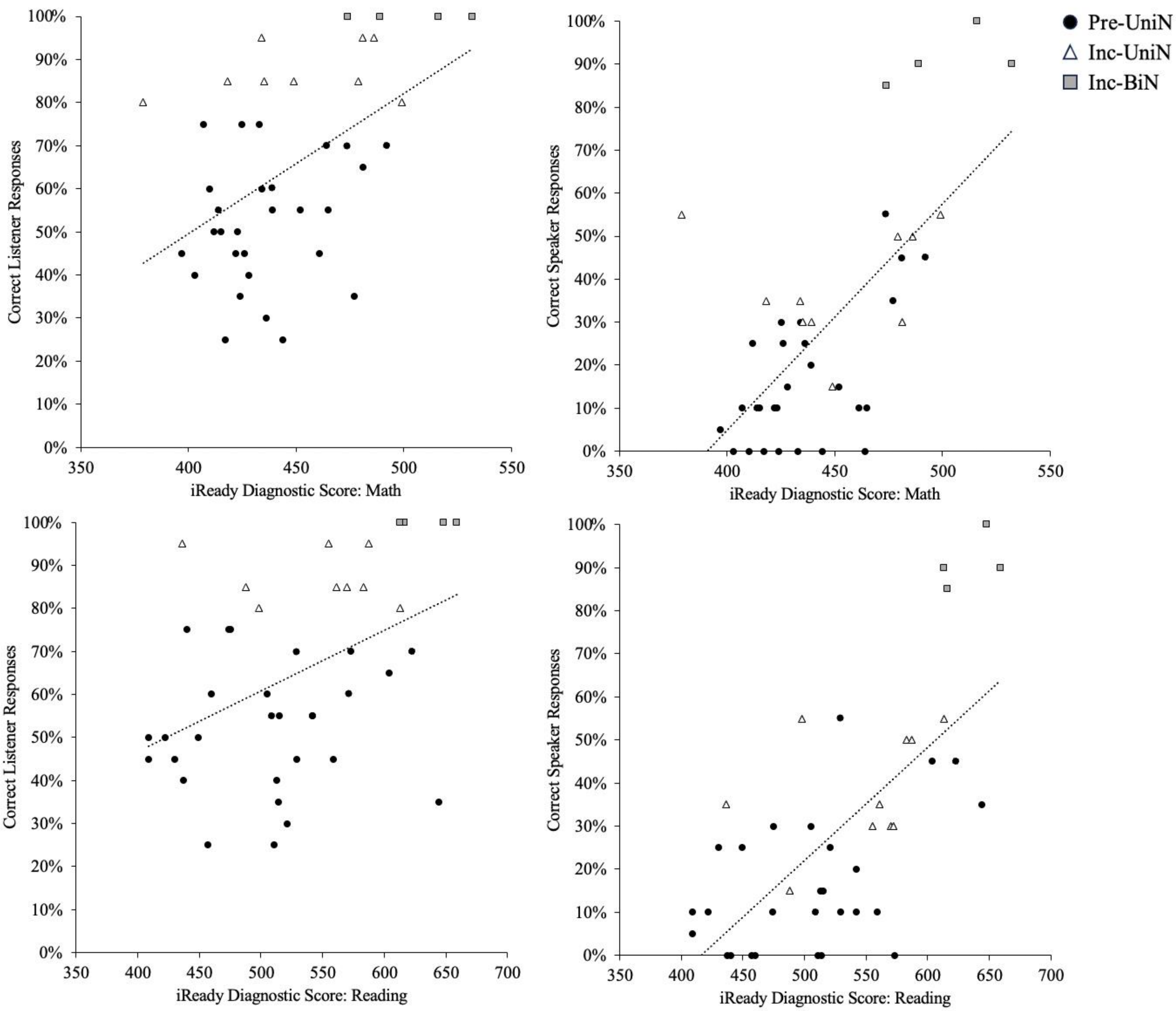
2.3 Results

Figure 1 displays participants' number of correct listener/speaker responses on Inc-BiN probes with *iReady*® Diagnostic scores for math and reading. The left graphs display the correlation between participants' percentage of correct listener responses with number of sets to mastery. The right graphs display the correlation between percentage of correct speaker responses with *iReady*® Diagnostic scores. The top graphs display correlations between listener/speaker responses on unfamiliar Inc-BiN probes with *iReady*® Diagnostic scores for math. The bottom graphs display correlations between listener/speaker responses on unfamiliar Inc-BiN probes with *iReady*® Diagnostic scores for reading.

The top left graph displays the relation between participants' percentage of correct listener responses on Inc-BiN probes and *i-Ready*® Mathematic Diagnostic scores. The researcher ran a nonparametric spearman correlation and found a significant positive relation between listener responses and *i-Ready*® Mathematics Diagnostic scores, $r = .465$, $p < .001$. The top right graph displays the relation between participants' percentage of correct speaker responses on Inc-BiN probes and *i-Ready*® Mathematic Diagnostic scores. The researcher ran a nonparametric spearman correlation and found a significant positive relation between speaker responses and *i-Ready*® Mathematics Diagnostic scores $r = .607$, $p < .001$. The bottom left graph displays the relation between participants' percentage of correct listener responses on Inc-BiN probes and *i-Ready*® Reading Diagnostic scores. The researcher ran a nonparametric spearman correlation and found a significant positive correlation between the variables, $r = .441$, $p < .001$. The bottom right graph displays the relation between participants' percentage of correct speaker responses on Inc-BiN probes and *i-Ready*® Reading Diagnostic scores. The researcher ran a nonparametric spearman correlation and found a significant positive correlation between the variables, $r =.634$, $p < .001$.

Figure 1

Correlation between participants' percentage of correct listener/speaker responses on Inc-BiN probes and iReady® diagnostic score for math and reading.



2.4 Discussion

The findings suggest there are significant correlations between degree of Inc-BiN and academic performance. Specifically, Inc-BiN has a stronger relation to academic performance than Inc-UniN (i.e., higher correlation coefficients). Participants' listener responses on Inc-BiN probes had a significant relation with academic performance, but not as strong as the relation between speaker responses and academic responses. These correlations were similar across reading and mathematics performance.

The researcher speculates there is a stronger relation between Inc-BiN and academic performance, as opposed to Inc-UniN and academic performance, because of the frequency of speaker responses in academic assessments. For example, participants were often asked to produce their own responses, as opposed to selecting an answer with multiple choices. When given a multiple-choice question, participants may guess and still have a chance to select the correct answer. However, when participants produce their own responses there is more room for error. The difference in difficulty is like the listener and speaker responses on the Inc-BiN probes, where the speaker responses have more room for error and are more effortful than listener responses. This may explain why Inc-BiN (speaker responses) scores are more strongly correlated to academic performance than Inc-UniN (listener responses) scores.

These correlations mean that participants with higher degrees of Inc-BiN tend to have higher academic scores. This aligns with previous research and Verbal Behavior Development Theory (Baldonado, 2021; Gilic & Greer, 2011; Greer & Ross, 2008; Morgan et al., 2020). Verbal Behavior Development Theory researchers may explain these correlations occur because individuals with Inc-BiN learn at a faster rate than those who do not (Abdool-Ghany & Fienup, in press; Friedman, 2021; Hranchuk, 2018). Inc-BiN allows an individual to incidentally acquire operants across the listener and speaker topographies. The presence of Inc-BiN may allow individuals to emit more emergent relations on the academic assessment to perform better than those who do not have Inc-BiN.

Previous research has more narrowly studied the relation between Inc-BiN and Academic Performance (Baldonado, 2021). To some degree, research has studied the impact of participants' degree of Inc-BiN on specific academic variables, such as tacts and derived relational responding (Abdool-Ghany & Fienup, in press; Friedman, 2021; Hranchuk, 2018; Morgan et al., 2021). In general, Inc-BiN research has focused on methods that induce Inc-BiN rather than utilizing Inc-BiN as a measurement tool to predict the appropriate intervention or to predict academic performance. For example, researchers have implemented MEI across listener and speaker topographies as a method to establish the stimulus control for Inc-BiN (Corwin & Greer, 2017; Fiorile & Greer, 2007; Gilic & Greer, 2011; Greer et al., 2011; Greer, et al., 2005). Until now, the relation between participants' degree of Inc-BiN and Academic performance has not been studied extensively in published works (Baldonado, 2021; Hranchuk et al., 2016). Morgan et al. (2020) analyzed the relation between Inc-BiN and emergent relations, where the present study analyzed the relation between Inc-BiN and academic performance on Common Core Standards. Students are often asked to emit responses that they have not explicitly learned. Their correct responses, emergent relations, may be an important variable in their individual academic performance. The present study extends literature by including a more commonly used academic measurement tool (i.e., *i-Ready*®).

Although the present study fills a gap in current literature, there are several limitations. Primarily there is a need for replication of the current study with more participants to fully understand the relation between Inc-BiN and academic performance. Additionally, these participants were all 4th and 5th grade participants in CABAS® classrooms. These procedures should be replicated to include multiple grade levels as well as different types of classrooms. These data demonstrated a clear trend and significant correlation; however, some participants stray from the mean. These differences may be better understood by including more participants. Future research should also investigate the relation between Inc-BiN for familiar stimuli and academic performance, as opposed to only unfamiliar stimuli. Morgan et al. (2020) explained that individuals perform better on familiar Inc-BiN probes than unfamiliar Inc-BiN probes because these stimuli are more like their immediate environment. Future researchers should also investigate the predictive value of Inc-BiN and academic assessments. The present study extended the literature by including a measure of academic performance across reading and math, but causality cannot be attributed based on these data alone.

Chapter 3: Incidental Bidirectional Naming and Matrix Instruction

3.1 Brief Literature Review

When teaching, it is crucial to maximize instructional time and resources. The results from Experiment 1 demonstrated a correlation between degree of Inc-BiN and academic performance. Specifically, speaker responses were the most highly correlated with academic performance in reading and in math. Often, teachers may not have time to directly teach each speaker response that their students could be required to produce both academically and socially. Here, emergent learning repertoires become important for maximizing instructional time. Emergent learning is a critical component of academic success (Critchfield & Twyman, 2014; Hawkins, et al., 2018). Emergent learning is a broad category that occurs when an individual responds accurately in a way that was not directly taught. As students become older, the curriculum becomes more and more complex. As the complexity and amount of content increases, there are fewer opportunities for teachers to directly teach each part of the curriculum. For example, first-grade students begin by learning how to textually respond to words, whereas fourth-grade students may be asked to produce the meaning of a vocabulary words. However, all these responses cannot be directly taught, which means individuals who do not have emergent learning repertoires will not learn as much as the individuals who do.

Researchers have also discussed the practical implications of measuring degree of Inc-BiN as it relates to academic instruction (Hawkins et al., 2018). Previous research has induced Inc-BiN in participants which caused participants to learn specific skills at a faster rate (Greer et al., 2011; Hotchkiss & Fienup, 2020). Teachers may benefit by classifying individuals by their degree of Inc-BiN to inform how instructional antecedent-response topography contingencies should be arranged to maximize instruction. For example, Hranchuk et al. (2016) investigated the effectiveness of instructional models with learn units vs. only learn units for participants with Inc-BiN and found that instructional models with learn units were more appropriate for participants with Incidental Bidirectional Naming (Inc-BiN) than only learn units. However, this is one of the only experiments that specifically compares an instructional design strategy (instructional models) with Inc-BiN. Results from Experiment 1 led the researcher to ask questions regarding the relation between specific instructional design strategies (matrix instruction and Equivalence-Based Instruction) and degree of Inc-BiN.

3.1.1 Matrix Instruction

Researchers have used matrix training to plan minimal exemplars within a multiple exemplar instruction context. Matrix instruction is a form of relational training that is an example of multiple exemplar instruction for rotating across stimulus dimensions. Matrix Instruction is a strategic instructional design approach that has been used to assess the number of exemplars one should use during teaching to promote emergent learning (Curiel, et al. 2020; Kemmerer, et al., 2021). Researchers can use matrix instruction to teach a variety of different skills, which fall under different categories of verbal behavior. For example, researchers have used matrix instruction to teach selection, tact, and textual responses (Curiel, et al. 2020; Kemmerer, et al., 2021). Often, researchers use matrix instruction to select essential atomic units to teach and establish generalized operant responding. Atomic units are individual components that can be used in a variety of context (Palmer, 2012). For example, the number five could serve as an atomic unit and may appear at the end of the equation “2 + 3 = _”, “11 – 6 = _”, “25 / 5 = _”, as well as many other antecedent variations. The individual may write the number “5” but has done so across multiple contexts. Generalized operant responding is a repertoire in which an individual functionally uses an atomic unit across a variety of contexts (Palmer, 2012). For example, when an individual acquires generalized operant responding for the word “not” they recognize that when “not” is presented before a word, it means the opposite of that word (i.e. not good = bad), regardless of the specific word it precedes. Researchers frequently use matrix instruction to test and establish generalized operant responding.

To implement matrix training, a teacher begins by considering the multiple antecedent stimuli they want to control a student’s academic responses and placing those in a 2-dimensional matrix (see Ch. 4 Figure 2). For example, if a teacher were planning noun-action correspondence, they could arrange the different nouns on the horizontal axis of a matrix and the relevant actions on the vertical axis of the matrix. Rather than teach all possible combinations of nouns and actions – which would require many teaching resources – the teacher could teach specific noun-verb pairs along the centerline of the matrix, and one might expect that the rest of the noun-verb pairs will be learned as emergent responses (Curiel et al., 2020; Frampton et al., 2016; Kemmerer et al., 2021).

Matrix training teaches that one stimulus (an atomic unit) has the same meaning/function in some other context (Alter & Borrero, 2015; Critchfield & Twyman, 2014; Fienup et al., 2011; Fienup et al., 2016 Ninness et al., 2006; Rehfeldt, 2011). Mueller et al., (2000) investigated the effects of matrix training of within-syllable units on participants’ emergent responses of selecting “novel consonant-vowel-consonant words” that have overlapping qualities. Research also looked at the effects of matrix training on vocal reading and reading comprehension. They implemented pre-intervention tests, then taught atomic units (word sets) with match-to-sample (MTS) training across five sets, followed by posttests. The general conclusion is that participants demonstrated low in-syllable responses until they completed the matrix training intervention, when their responses increased on the post-intervention responses. Three out of the five participants demonstrated generalized operant responding with substantial increases in emergent responses for within in-syllable units.

Frampton et al. (2016) used matrix instruction to arrange instruction to teach three nouns (atomic units) and three actions (atomic units) to participants with autism spectrum disorder (ASD). Researchers arranged the nouns (cat, dog, bear) on the vertical axis and actions (eating, drinking, jumping) on the horizontal axis. After baseline, researchers taught participants only the three noun-verb combinations (cat eating, dog drinking, bear jumping) that fell on the diagonal centerline of the matrix. The noun-verb combinations that did not fall on the diagonal were not directly taught (e.g., dog eating, bear drinking, cat jumping); however, the participants learned these noun-verb combinations following learning specific noun-action combinations. The researchers considered the non-directly taught responses *emergent relations*.

Researchers and educators have used matrix instruction to teach participants effectively and efficiently (Curiel et al., 2020; Frampton et al., 2016; Kemmerer, et al., 2021). A primary goal of matrix instruction is to expose the participants to the components and combinations of the matrix stimuli as effectively and efficiently as possible. Teachers use matrix instruction as an instructional tool to arrange the minimal number of multiple exemplars needed to learn a larger set of responses (Mueller, 2000; Curiel et al., 2020; Kemmerer et al., 2021).

Matrix Instruction Outcomes. Mueller et al. (2000) investigated the effects of matrix instruction of within-syllable units on participants' emergent responses of selecting "novel consonant-vowel-consonant words" that have overlapping qualities. Researchers taught four words at a time across six sets. The words in each set had overlapping letters. For example, set one consisted of researchers teaching words mat, sat, sop, and sug and then tested words, mop, mug, sop, and sug. This pattern was replicated across six sets of words. Two out of three participants demonstrated increases in emergent responses after one set and the third participant demonstrated increases in emergent responses after two sets. The general conclusion is that participants demonstrated low within-syllable responses until they completed the matrix instruction intervention.

Many researchers have extended Mueller et al.'s (2000) work and studied the effects of matrix instruction with multiple types of stimuli, participants, and variations of matrix interventions (Curiel et al., 2020; Frampton et al. 2016; Kemmerer et al. 2021). Rickoski et al. (2023) implemented an adapted alternating treatment design embedded within a multiple baseline across participants design to assess the effects of matrix instruction effectiveness and efficiency in three different conditions: overlapping matrix training, non-overlapping matrix training, and non-overlapping matrix training plus instructive feedback. Researchers taught participants to label common characters and cities. For example, some of the targets were Daredevil in Phnom Penh, Beast in Vienna, and Thing in Istanbul. In the overlapping matrix condition researchers taught eight targets and tested for eight untaught targets. In the non-overlapping matrix training, researchers taught four targets and tested twelve targets. In the non-overlapping matrix training plus instructive feedback condition researchers trained four targets and tested twelve targets. Researchers taught three matrices to three participants to mastery. All three participants acquired the directly taught and emergent responses across all three of the matrices. Researchers concluded that matrix training is an efficient and effective tool for teaching emergent responses.

Curiel et al. (2020) and Kemmerer et al. (2021) summarized the growing body of matrix instruction literature since Mueller et al. (2000). Curiel et al. (2020) reviewed matrix instruction studies that tested their interventions on participants with autism spectrum disorder (ASD). They analyzed 12 studies that implemented a matrix instruction procedure across language development skills, play skills, and sentence construction and spelling. Most of the matrix instruction procedures utilized discrete trial training and targeted language skills. On average, researchers reported that participants emitted a mean of 69% of emergent relations as a function of matrix instruction. Kemmerer et al. (2021) extended this work in a systematic review of the matrix instruction literature on a broader population with 35 studies. These studies implemented matrix instruction procedures across 230 participants who varied in age and diagnosis. Kemmerer et al. (2021) reported that researchers typically implemented matrix instruction to teach language with multiple probe designs. They reported variability in participants' acquisition of emergent relations through matrix instruction and stated these differences may be a function of participant prerequisite skills as opposed to the matrix intervention itself. Kemmerer et al. (2021) broadened the scope of researchers' understanding of matrix instruction and who it may be best for by including participants with and without ASD.

Matrix Instruction Limitations. Depending on the student outcomes, matrix instruction may be considered efficient instruction when the total amount of learning exceeds that which was specifically taught (Curiel et al., 2020). However, the participants who completed the matrix instruction interventions have a wide range of success in the number of emergent responses emitted (Curiel, et al. 2020; Kemmerer, et al., 2021). Researchers hypothesize different reasons for these mixed outcomes. Some reasons include setting events during intervention, missing prerequisites, and inconsistencies in dependent variable definitions and measurements (Curiel, et al. 2020; Kemmerer, et al., 2021).

Another limitation of matrix instruction research is that most studies use listener responses as their primary dependent variable. Listener responses include matching and selection responses so that the participant does not have to independently produce a response. Although important to study, there seem to be a much smaller number of studies that measure speaker responses, specifically emergent intraverbals (Conine et al., 2023). Individuals are constantly required to emit intraverbal responses in social and academic settings. Future research should investigate matrix instruction and emergent intraverbal responses to fill this gap. Emergent behavior is behavior that is not explicitly taught (Connie et al., 2023; Fienup, 2018).

3.1.2 Emergent Intraverbal Responses

Emergent behavior may be further categorized by what antecedent topographies are used, what response topographies are used, what generalization responses are tested, and what stimulus relations are tested (Connie et al., 2023). For example, a researcher may choose to present antecedents or require responses through a vocal or written topography. Researchers acknowledge there is a substantial need for more research on emergent relations to fully understand the contributing variables needed to occasion emergent relations as well as the implications emergent relational repertoires have on student learning outcomes (Connie et al., 2023).

Emergent intraverbal responses are one area of emergent behavior that may warrant more attention due to the importance of understanding emergent behavior and the frequency of intraverbal responses in our verbal repertoires (Connie et al., 2023; Jennings & Miguel, 2017). Skinner (1957) stated intraverbals are one of the five basic verbal operants and are defined by their function. Intraverbals are responses with a verbal antecedent that do not share point-to-point correspondence with the original antecedent and occur under generalized reinforcement conditions. A common example of an intraverbal response is answering questions. The question is a verbal antecedent, the response to the question is the intraverbal behavior, and the consequence involves some feedback on one's response or social interaction (e.g., Antecedent: "How was your evening?;" Intraverbal: "It was great, I made spaghetti," Consequence: "Nice, I made chicken parmesan.)" Intraverbals are particularly relevant because of the high frequency in which we use intraverbals in our verbal repertoires which includes social and academic interactions (Palmer, 2016; Skinner, 1957; Sundberg & Sundberg, 2011).

Researchers have studied emergent intraverbals in different ways. These include different types of intraverbal responses and different methods used to teach intraverbal responses. Emergent intraverbals differ from intraverbals in that emergent intraverbals are intraverbal responses are not directly taught, whereas intraverbal responses are directly taught. Different types of intraverbal responses include fill-in responses (e.g., 1+ _ = 3 (2)), chained intraverbals (e.g., A, B, C, _ (D)), intraverbal associations (e.g., Mommy and _____ (Daddy)), to conversational intraverbals (e.g., How are you? – I am good). Furthermore, intraverbal response topography can vary between vocal, written, typed, sign, and picture-based communication (Connie et al., 2023). Over the last 30 years there has been a growing amount of published work that studies emergent intraverbal responses (Connie et al., 2023). Connie et al., (2023) conducted a research review on emergent intraverbal research that included Naming Theory research, Relational Frame Theory research, and Stimulus Equivalence research.

There are different types of emergent intraverbal responses that are studied at different frequencies. Some examples of emergent intraverbals are written, typed, and selection. Emergent written intraverbals are responses that are not directly taught and emitted in a written topography. For example, an emergent written intraverbal would occur if the instructor modeled how to draw a square, then, wrote on the board "draw a square," and the student independently drew a square. Emergent typed intraverbals are not directly taught and emitted in a typed topography. Connie et al. (2023) demonstrated that 89% of emergent intraverbal research measured their dependent variables with vocal emergent intraverbal responses, as opposed to written, typed, signed, picture exchange, or text selection responses. Written responses accounted for 5% of the emergent intraverbal research, which seems concerning when considering the prevalence of written intraverbal responses in general and special education. Researchers call for more research on emergent intraverbals as dependent variables, specifically, written emergent intraverbals.

3.1.3 Rationale for Experiment 2

Since speaker responses demonstrated the most significant positive correlation with academic performance in Experiment 1, Experiment 2 investigated if degree of Inc-BiN predicted success in matrix instruction. Specifically, the researcher measured degree of Inc-BiN and emergent intraverbal responses to academic stimuli. Experiment 2 investigated the effects of matrix instruction on emergent intraverbal responses based on 14 participants' degree of Inc-BiN. Researchers implemented a relational training procedure (matrix instruction) that taught intraverbal atomic units and tested generalized operant responding. Experimenters implemented a multiple probe design across participants/groups. Researchers delivered learn units from the matrix and taught five vocabulary words during intervention and assessed emergent intraverbal responses pre and post intervention.

Experiment 2

3.2 Methods

3.2.1 Participants

The researcher selected 14 fourth-grade participants from a public school outside a major metropolitan area. Researchers selected fourth-grade participants who demonstrated instructional readiness. Individuals who were frequently pulled to go to the principal's office were excluded from the study. Table 6 displays summarized participant data and Table 7 displays individual participant data. Participants were in the same setting (CABAS® AIL® 4th-grade classroom) and were the same age as in Study 1. The researcher selected the participants for this study because the participants demonstrated a need for vocabulary instruction.

Five participants had Individualized Education Programs (IEPs). Four out of the five participants with IEPs were male, while the other participant with an IEPs was female. The remaining nine participants were general education participants. The general education participants included six males and three females. All fourteen participants functioned as readers and writers, which is a verbal behavior development classification of the participants' academic repertoires. As readers and writers, participants could textually respond to both literature and informational texts and write to inform, persuade, and entertain. Eight participants were on grade level for reading, based on schoolwide reading diagnostic results. Six participants' diagnostic results indicated that they were functioning one or more grade levels below their enrolled grade for reading (Table 7). The *i-Ready*® Diagnostic also reported participants' vocabulary scores. Eight participants were on or above grade level in vocabulary, three participants were one grade below in vocabulary, and three participants were two or more grade levels below in vocabulary.

Seven participants demonstrated Pre-Unidirectional Naming (Pre-UniN), five participants demonstrated Incidental-Unidirectional Naming (Inc-UniN), and two participants demonstrated Incidental Bidirectional Naming (Inc-BiN). Participant information is displayed on Table 6 and 7. Researchers selected experimental groups based on academic performance and rate of learning that reflected their typical classroom groupings.

Table 6

Participant information (age, gender, race, classification degree of Inc-BiN listener, and speaker scores.

Variable	*N*	Percent
Sex	M = 10	M = 71%
	F = 4	F = 28%
Race	W = 6	W = 42%
	M = 3	M = 21%
	H = 4	H = 28%
	B = 1	B = 7%
Learning Classification	IEP = 5	IEP = 35%
	504 = 1	504 = 7%
Inc-BiN Classification	Pre-UniN = 7	Pre-UniN = 50%
	Inc-UniN = 5	Inc-UniN = 35%
	Inc-BiN = 2	Inc-BiN = 14%

Note. W = white, M = Mixed, H = Hispanic, B = Black, A = Asian, IEP = Individualized Education Plan, ESL, English as Second Language, I&RS = Intervention and Referral service

Table 7

Individual Participant information (age, gender, race, classification degree of Inc-BiN listener, and speaker scores).

					Degree of Naming: Unfamiliar			*iReady*® Reading Diagnostic	
Participant	Age	Gender	Race	Learning Classification	Classification	Total Listener (30)	Total Speaker (30)	Scaled Score	Grade Level
1	9	F	B	IEP	Pre-UniN	19	3	474	2
2	10	M	H	IEP	Pre-UniN	17	7	449	1
3	11	M	W	IEP	Pre-UniN	15	7	542	4
4	10	M	M		Pre-UniN	17	8	505	3
5	9	M	W		Inc-UniN	27	9	555	4
6	10	M	H	IEP	Pre-UniN	24	9	475	2
7	9	M	W	IEP	Inc-UniN	27	10	570	4
8	9	F	H		Pre-UniN	20	11	572	4
9	10	M	M	504	Pre-UniN	23	11	613	4
10	9	F	H		Inc-UniN	24	13	561	4
11	10	M	M		Pre-UniN	23	14	604	4
12	9	F	W		Inc-UniN	28	20	616	4
13	10	M	W		Inc-BiN	30	28	613	4
14	9	M	W		Inc-BiN	30	30	648	6

3.2.2 Setting and Materials

The researcher created their own vocabulary matrix with five prefixes and word roots (Figure 2). The researcher color-coded Figure 2 to indicate which words were directly taught and which words were used in the pre and post assessments. The researcher directly taught the yellow words in intervention. The researcher selected words based on 4th-grade CCSS reading standards. The researcher also ensured that the prefixes and word roots combined with each other logically on the matrix. The researcher collected data within the participants' classroom. Participants completed two researcher-designed probes, with a paper and pencil, while seated at their own desks. Participants completed intervention using their school-provided Chromebooks while seated at their own desks. The researcher and participants utilized an instructional application, called PearDeck©, to complete the intervention portion of the experiment. The researcher created their own instructional materials using GoogleSlides© that included one slide for the antecedent with a participant typing response, one slide for the correction/definition of the word, and one slide with the antecedent re-presented. The researcher created slides for each of the five operants and then randomized the order of the operants (see Figure 6). For Brief Inc-BiN Assessments, the researcher used materials identically as in Experiment 1.

Figure 2

Matrix operants used for intervention and probes

Set 1	Form -	Qualify	Fix	Cook	Graph
Dis-	Disform	Disqualify	Disfix	Discook	Disgraph
Un-	Unform	Unqualify	Unfix	Uncook	Ungraph
Pre-	Preform	Prequalify	Prefix	Precook	Pregraph
Re-	Reform	Requalify	Refix	Recook	Regraph
Semi	Semiform	Semi-qualify	Semifix	Semicook	Semigraph

Note. Yellow = directly taught words during intervention, white = emergent words tested pre/post intervention

3.2.3 Measurement

Emergent Intraverbal Probe. The researcher measured the number of correct emergent vocabulary responses using the researcher-designed, Emergent Intraverbal Probes. There were two parts to Emergent Intraverbal Probes; *Written Definition Probes* (see Figure 4*)* and *In-Context Probe (*see Figure 5). The Written Definition Probe consisted of all the untaught matrix words (20) being listed on a worksheet with a printed antecedent of "write the definition of each word." Participants could earn between zero and two points per question. A correct response was defined as the participant writing the definition of the full word. In other words, they included both the prefix's meaning and the root word's meaning in their definition. If the participant included both components of the definition, the researcher recorded two points for that question. If the participant included only the prefix's meaning, or only the root word's meaning, the researcher recorded one point for that question. An incorrect response was defined as the participant writing definitions that did not represent what the word, or word parts, meant. An incorrect response was also defined as the participant emitting no written response on a question. If the participant emitted an incorrect response, the researcher scored it as a zero. Once scored, the researcher summed the number of points earned out of 40 points. See Table 8 for the teacher script of correct and incorrect responses (See Table 8).

Table 8

Teacher script for intervention operants

Word	Definition	Other Correct Responses	Incorrect Definitions
Disform	Dis (take away) form (shape/take shape) → take away shape	-Take away shape -Lose form -Change form	-not order -lose order -deform
Unqualify	Un (not) qualify (having the skills) → not having the skills	-not fit for -lacking skills -does not meet requirements	-not knowing -pre-requisites -qualified
Prefix	Pre (before) fix (secure or fasten) → before secured/fastened	-before fixed -before attached -before being secure	-fixing -before changing -after secure
Recook	Re (again) cook (make food) → To make food again	-to make food once more -do over making food -make food again and again	-make again -making food -redoing cooking
Semigraph	Semi (half or part) graph (to write) → halfway or partly written	-half writing -partway written -partially written	-written down -half of a word -part of a picture

Figure 3

Slides used in intervention, antecedent and correction.

Antecedent　　　　　　　　　　　　Correction

Figure 4

Emergent intraverbal written definition probe.

Name:______________

Pre-1 -- Write the definition of each word

Disqualify

Unfix

Precook

Regraph

Unform

Prequalify

Refix

Semicook

Preform

Disfix

Figure 5

Emergent intraverbal in-context probe sample page

Name: ____________
Date: GD - Probe 1

Circle the root in the underlined word. Then write the meaning of each root below the sentence.

1. They may disqualify him for cheating in the game.

___

2. What temperature do you precook the turkey to?

___

3. The building began to unform as it got old.

___

4. I can't believe I had to refix that desk after the 5th time.

___

5. Angel wanted to preform the dough before putting it in the oven

___

6. Once I learned the right way to do it, I had to regraph everything.

___

7. I wish I could prequalify for my dream college.

___

8. Thankfully, my dog couldn't unfix the table this time.

___

9. To best make the chicken, you should semicook it the day before.

___

10. The baby loved to disfix the neat shelf every chance they got.

___

The researcher measured the number of correct In-Context Probe responses using the researcher-designed, In-Context Probes. In-Context Probes (see white words in Figure 2) consisted of the same untaught words (20) being presented, however, they were written within a sentence that gave a context clue regarding the meaning of the word. For example, with the word "disqualify," the antecedent was a written sentence "They may disqualify him for cheating in the game" and a written direction to "Circle the root in the underlined word. Then write the meaning of each root word." A correct response was defined as the participant circling the corresponding root word and writing the corresponding meaning of only the root word. If the participant emitted a correct response, the researcher scored that response with two points. Like the Written Definition Probe, partial points were also given. One point was recorded if the participant only circled the root word, or only wrote the correct meaning of the root word. An incorrect response, for the first instruction, was defined as the participant circling the whole word, circling more than just the root word, circling only a part of the root word, or not circling anything. An incorrect response, to the second part of the direction, was defined as the participant writing the definition for the whole word, writing a definition that did not correspond to the root word, or writing no response.

Brief Incidental-Bidirectional Naming Assessment. The researcher measured and classified participants' Inc-BiN identically to Experiment 1 except instead of two unfamiliar probes, the researcher conducted three unfamiliar probes. The researcher measured the participants' degree of Inc-BiN using three unfamiliar Inc-BiN probes.

Learn Units to Criterion. The researcher measured participants' learn units to criterion following intervention. Researchers counted one learn unit for each participant once the participants observed the antecedent, participants emitted a response, and the researcher consequated the response. Researchers calculated learn units-to-criterion by recording the total number of learn units delivered to each participant before they met the intervention mastery criterion. Researchers defined mastery criterion as five consecutive correct responses per word.

3.2.4 Procedure

Emergent Intraverbal Probes. Prior to and following intervention, the researcher systematically delivered the two emergent intraverbal probes (Written Definition Probe and In-Context Probes). The researcher gave the participant the probes and instructed them to "Do the best they can." The researcher did not provide any feedback on these probes, so they were unconsequeated and an independent response.

Matrix Instruction. Matrix instruction was conducted in small groups with one researcher and between three and five participants per group (see yellow words in Figure 2 for words taught during intervention). The researcher selected groups based on their academic performance and rate of work completion, meaning, the researcher placed participants who performed similarly in groups together. The researcher instructed the participants to open their Chromebook, open Peardeck© and type in the code needed for the session. The researcher implemented matrix instruction using learn units. Once, all the participants displayed the antecedent on their screen, the researcher presented the first slide/antecedent. The antecedent consisted of a written question, "What does _______ mean?" (One of the target five words). Participants typed their response on their own Chromebook. The researcher could see all the participant's responses, while each participant could only see their own. Once all the participants typed their responses, the researcher provided vocal praise to the participants who emitted a correct response and a correction to the participants who emitted an incorrect response. Vocal reinforcement consisted of the researcher saying, "Good job!" or "That's right!" A correction consisted of the researcher showing the next slide that had the complete definition of the entire word as well as the definition of each of the word parts. Once the researcher vocally read what was also written on the participant's screen, the researcher removed the prompt and re-presented the original antecedent to allow the participant an independent opportunity to respond. The researcher instructed the participants who already emitted the correct response to not respond again and wait quietly until the next question. The researcher repeated the correction procedure up to three times, or until the participant emitted a correct response.

The researcher randomized the order to rotate antecedent stimuli across all five operants until the participants met mastery criterion. The researcher set mastery criterion at five consecutive correct responses per operant. So, each participant emitted 100% accuracy across one session. Once the participant mastered an operant, the researcher did not require the participant to respond to that operant again. Once the participant mastered all five of the operants, participants completed another set of probes. After the participant completed those probes, the participant completed other schoolwork (i.e. silent reading, overdue assignments, extra credit work etc.) independently and no longer sat with their group during intervention. Intervention sessions lasted between 15 to 40 min per day. The number of learn units presented in each section varied between five to 15 learn units, based on the length of the session.

3.2.5 Experimental Design and Statistical Analysis

The researcher implemented a multiple probe design across groups to evaluate the effects of matrix instruction and degree of Inc-BiN on emergent intraverbals. All participants completed three Written Definition Probe and In-Context Probes prior to intervention and three times after intervention. All participants completed probe one at the same time on the same day. Next, Group 1 completed the second probes two and three, completed intervention, and then took the three more probes after intervention. Group 2 only completed probes two and three once all of Group 1 had completed their intervention. The same order continued for Group 3 and Group 4. Figure 6 displays the multiple probe design across groups. Figure 6 displays the multiple probe design across groups. Each row is one group and each bar (per row) represents one participant.

The researcher analyzed the data by running correlations that analyzed the relation between participants' degree of Inc-BiN and emergent intraverbal responses. The researcher measured degree of Inc-BiN for unfamiliar stimuli for each participant by conducting three brief Inc-BiN assessments and reported scores for their listener and speaker repertoires. The researcher calculated degree of Inc-UniN and Inc-BiN in the same way as Experiment 1. The researcher measured academic performance by reporting the participants' total emergent intraverbal responses. The researcher conducted nonparametric spearman correlations to test the association between emergent intraverbal responses and degree of Inc-UniN and emergent intraverbal responses and degree of Inc-BiN.

3.2.6. Interobserver Agreement and Treatment Fidelity

Two researchers independently collected data on participants' Brief Inc-BiN probe responses. The researcher calculated Interobserver Agreement (IOA) for responses to listener and speaker responses to Inc-BiN probes. The researcher used trial-by-trial agreement and divided the number of agreed-upon trials by the total number of trials. They divided the number of agreed-upon trials by the total number of opportunities to respond and then multiplied the result by one hundred to make a percentage. The researcher conducted IOA for 27% of Inc-BiN assessments, with 100% agreement.

The researcher also collected permanent products from all the unconsequeated Written Definition Probe In-Context Probes. Two researchers independently collected data on participants' responses. We divided the number of agreed-upon responses by the total number of opportunities to respond and then multiplied the result by one hundred. ISA was conducted for 100% of Written Definition Probe and In-Context Probes with 99% agreement.

Treatment fidelity was collected during intervention sessions to measure the accuracy of intervention implementation. Treatment fidelity was conducted with the use of Teacher Performance Rate and Accuracy forms (TPRA; Ingham & Greer, 1992). The researcher used TPRAs to measure the accuracy of an instructor's delivery of antecedents and consequences. The researcher conducted TPRAs by observing another individual deliver the intervention. TPRAs were conducted to ensure accurate implementation of the intervention. The researcher calculated the percentage of sessions with treatment fidelity by totaling the total number of trials observed with a TPRA and dividing that number by the total number of trials delivered. The researcher calculated treatment fidelity agreement by totaling the number of correctly delivered trials and dividing that number by the total number of trials observed. Treatment fidelity was conducted for 21% of the Intervention with 99% agreement.

3.3 Results

Table 9 displays the total number of correct responses on pre-intervention and post-intervention emergent intraverbal probes. The table also displays the difference between participants' pre-intervention and post-intervention scores. Participants are listed in order by their change scores from pre-intervention to post-intervention. Thirteen out of fourteen participants demonstrated increases in correct emergent intraverbal responses after intervention. Table 9 displays these raw scores as well as participants' Inc-BiN classification. Pre-UniN participants demonstrated a mean of 29.8 correct responses after intervention, Inc-UniN participants demonstrated a mean of 23.6 correct responses after intervention, and Inc-BiN participants demonstrated a mean of 80.5 correct response after intervention.

Figure 6 displays the multiple probe design across groups with participants' pre-intervention and post-intervention emergent intraverbal scores. In the first group, all three participants (Participants 8, 10, and 4) emitted 0 correct responses on pre-intervention probes. On the post-intervention probes, Participant 8 emitted 15%, 20%, then 20% correct responses. Participant 10 emitted 8%, 18%, and 13% correct responses. Participant 4 emitted 1%, 4%, and 4% correct responses.

The second group in Figure 6 had five participants. On pre-intervention probes, Participants 11,12, 14, 9, and 13 emitted a low level of correct responses (0-10% correct responses). On post-intervention probes, Participant 11 emitted 38% correct on all three of the probes. Participant 12 emitted 28%, 23%, and 25% correct responses. Participant 14 emitted 100%, 100%, and 98% correct responses. Participant 9 emitted 98%, 93%, and 98% correct responses. Participant 13 emitted 20%, 43%, and 43% correct responses.

The third group in Figure 6 had three participants. On pre-intervention probes participants emitted low levels of correct responses (0-5% correct responses). On post-intervention probes, Participant 5 emitted 8%, 8%, and 13% correct responses. Participant 7 emitted 10%, 10%, and 5% correct responses. Participant 3 emitted 3% correct responses on all three probes.

The fourth group in Figure 6 had three participants. On pre-intervention probes, participants emitted low levels of correct responses (0-3%). On post-intervention probes, Participant 6 emitted 43%, 43%, and 45% correct responses. Participant 1 emitted 0% correct responses on all three probes. Participant 2 emitted 10%, 20%, and 8% correct responses.

Figure 7 displays participants' listener/speaker responses as compared with their emergent intraverbal responses. The left graph in Figure 7 displays the relation between percentage of correct listener responses as compared with the total number of correct emergent intraverbal responses. The researcher ran a nonparametric spearman correlation and found a significant correlation between variables, $r = .560$, $p = .019$, meaning participants' percentage of correct listener responses was positively related to emergent intraverbal responses. The right graph in Figure 7 displays the relation between percentage of correct speaker responses as compared with the total number of emergent intraverbal responses. The researcher ran a nonparametric spearman correlation and found a significant correlation between variables, $r = .749$, $p = .00$, meaning participants' percentage of correct speaker responses was positively related to emergent intraverbal responses.

Figure 8 displays participants' number of correct responses on Inc-BiN probes with their learn units to criterion. The left graph displays the correlation between participants' percentage of correct listener responses with their learn units to criterion. The researcher ran a one-tailed nonparametric spearman correlation and found a significant negative correlation between variables, $r = -.482$, $p = .041$. The second graph displays the correlation between participants' percentage of correct speaker responses with their learn units to criterion. The researcher ran a one-tailed nonparametric spearman correlation and found a significant negative correlation between variables, $r = -.589$, $p = .013$.

Table 9

Displays individual participants' total number of correct responses on pre- and post-intervention emergent intraverbal scores, learn units to criterion, and the difference between their pre- and post- intervention scores.

Participant	Classification	Learn Units to Criterion	Pre-Intervention Probe Scores	Post-Intervention Probe Scores	Change from pre to post
1	Pre-UniN	40	1	0	-1
3	Pre-UniN	56	1	3	2
4	Pre-UniN	60	0	9	9
5	Inc-UniN	44	2	11	9
7	Inc-UniN	53	1	10	9
2	Pre-UniN	43	2	15	13
10	Inc-UniN	59	0	15	15
12	Inc-UniN	45	9	30	21
8	Pre-UniN	43	0	22	22
13	Inc-BiN	41	4	42	38
11	Pre-UniN	52	2	45	43
6	Inc-UniN	52	1	52	51
9	Pre-UniN	46	4	115	111
14	Inc-BiN	42	6	119	113

Figure 6

Multiple probe design graph depicting emergent intraverbal responses across sessions across groups

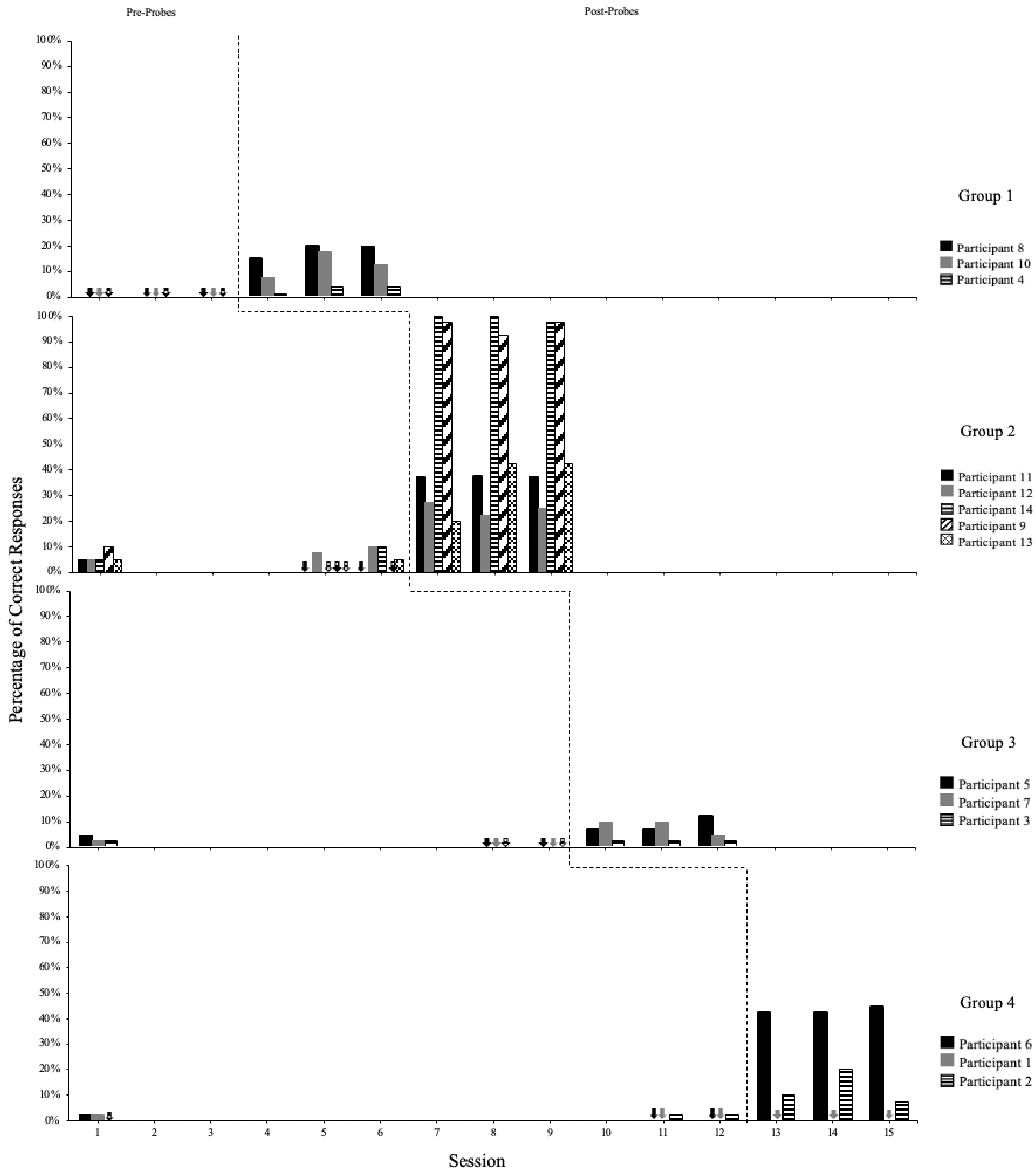


Figure 7

Percentage of correct listener/speaker responses and total number of correct emergent intraverbal responses.

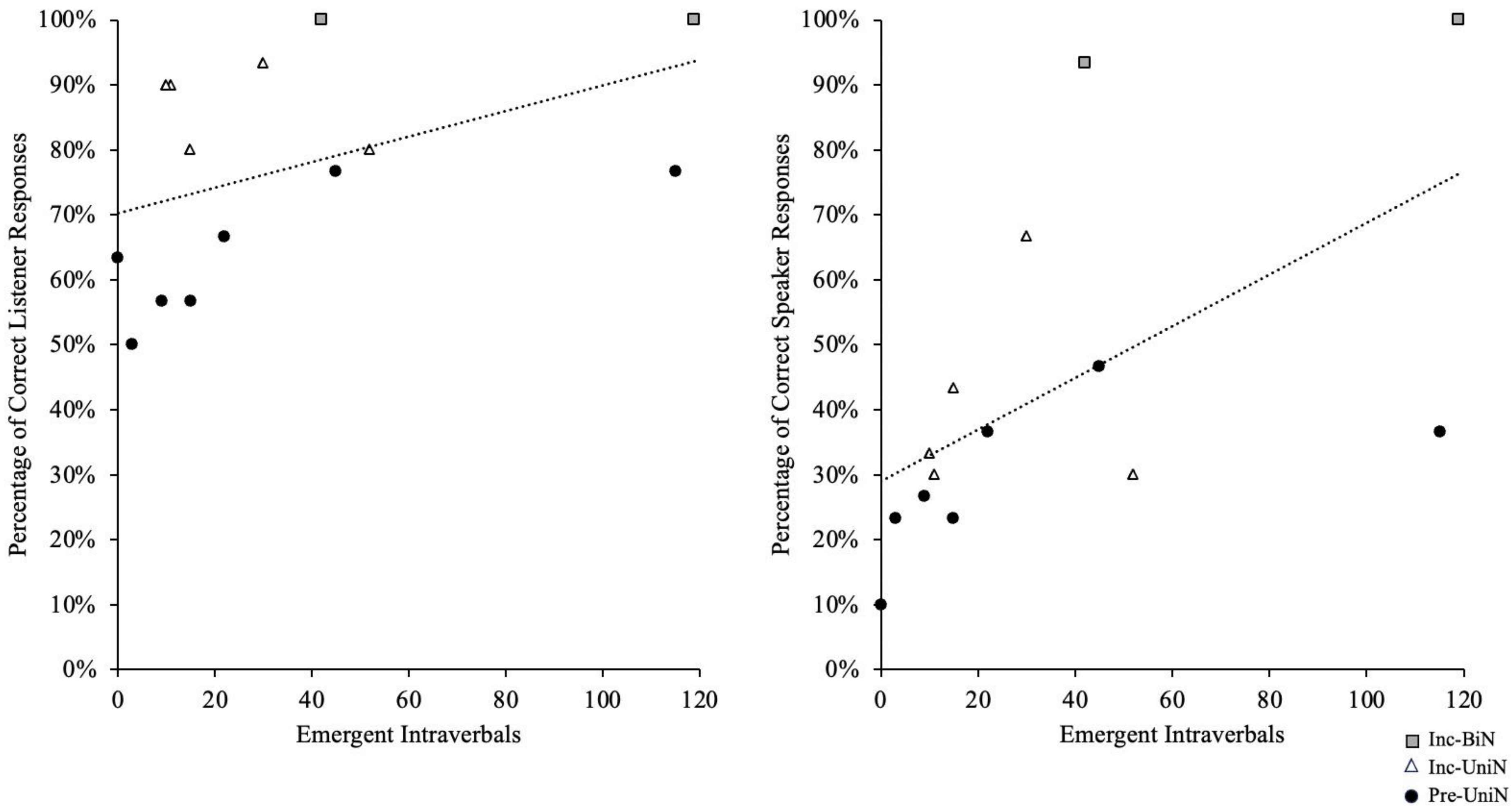
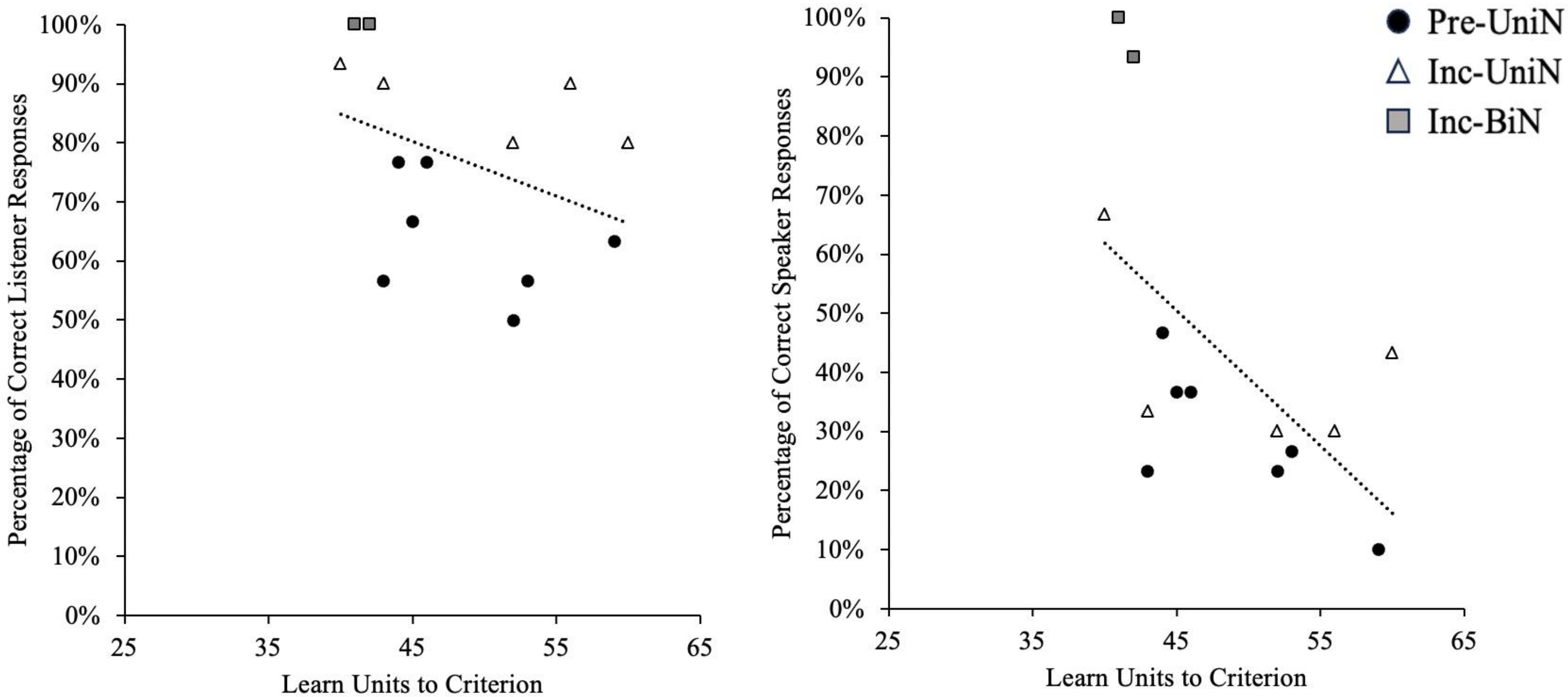


Figure 8

Percentage of correct listener/speaker responses and learn units to criterion



3.4 Discussion

In this experiment, these data show that Inc-BiN is positively correlated with generalized operant responses, in this case, emergent intraverbals. The Pre-UniN and Inc-UniN participants emitted the lowest number of correct emergent intraverbal responses, while the Inc-BiN participants emitted the highest number of correct emergent intraverbal responses. Based on these data, it seems that matrix instruction has the best effects on participants' vocabulary skills with participants who demonstrated Inc-BiN. Overall, the greater the degree of Inc-BiN, the greater number of generalized operant responses (emergent intraverbals) emitted following matrix instruction (See Figure 7).

Researchers found a significant negative correlation between participant's percentage of correct listener responses with number of learn units to criterion and a significant correlation between participant's percentage of correct speaker responses with the number of learn units to criterion. Overall, the higher percentage of correct listener/speaker probes, the lower the number of learn units to criterion. An individual with fewer learn units to criterion required less learn units during intervention to meet mastery criterion than another individual with greater learn units to criterion. In other words, the higher the participant's degree of Inc-BiN, the less learn units they require to reach mastery criterion.

Previous matrix instruction research has focused on language skills while measuring matching and pointing responses as the dependent variables (Curiel et al., 2020; Kemmerer, et al., 2021). Curiel et al. (2020) and Kemmerer et al., (2021) reported that participants emitted a mean of 69% of emergent relations following matrix instruction. In the present experiment, participants emitted a mean of 29% of emergent relations following matrix instruction. The mean number of emergent relations in the present experiment is a lot lower than the mean emergent relations in the two systematic reviews. The researcher suggests this difference may be explained by the form of the dependent variable. Most of matrix instruction literature measures emergent responses but not emergent intraverbal responses. Intraverbals are a more complex repertoire than pointing or matching (Skinner, 1957). Additionally, matrix instruction literature has not used consistent language when describing the conceptual properties of target responses. For example, in the seminal experiment Mueller et al. (2000) used the term recombinative generalization responses to describe behaviors targeted for increase during intervention. Other researchers have used terms, receptive and expressive identification, subject-verb-object tacts, scripted and recombined actions and vocalizations (Naoi et al., 2006; Kohler & Malott, 2014; Pauwels et al., 2015; MacManus et al., 2015). Other matrix instruction researchers have typically defined the target skill but neglect to explain the conceptual underpinnings of the target behavior. The lack of explanation of the conceptual underpinnings that the researchers targeted may contribute to the inconsistent findings. For example, researchers who use matrix instruction to teach novel tact responses is a different skill than generalized operant responding and may produce very different results. The complexity of intraverbals as compared to pointing/matching responses may explain when there is such a big difference between the mean number of emergent responses from the systematic reviews and the present experiment.

One possible explanation for these results is that individuals with Inc-BiN learn at a faster rate. Abdool-Ghany & Fienup (in press) and Friedman (2021) found that participants with Inc-BiN learn at a significantly faster rate and emit more emergent relations than those without Inc-BiN. These findings are consistent with the current experiment that the participants with Inc-BiN learn at a faster rate and emit more emergent relations than the Inc-UniN and Pre-UniN participants do. Verbal Behavior Development researchers may explain this relation occurs because individuals with Inc-BiN emit emergent relations across listener and speaker topographies. Individuals with Inc-BiN demonstrated 80% or greater correct responding on listener and speaker probes on Inc-BiN assessments after only hearing the names of the stimuli four times. The speaker responses are a type of emergent relation and so are the emergent intraverbal responses after matrix training. The similarity of these types of responses to different stimuli is a potential explanation for the positive correlation. Like Experiment 1, there are participants who stray from the mean. These data, along with past research, demonstrated general significant correlations between Inc-BiN and emergent relations but call for additional research to account for variability.

The present experiment offers a potential explanation for participant differences after matrix instruction (Curiel et al., 2020; Kemmerer, et al., 2021). There are significant correlations between degree of Inc-BiN and generalized operant responses (emergent intraverbals). The researcher hypothesizes that Inc-BiN is a potential predictor of emergent intraverbals but requires additional research in this area. One limitation is that the researcher only measured Inc-BiN as a potential mediating variable for variable performance in matrix training. Some other variable/s could better explain the differences in Matrix outcomes (i.e. stimuli used, antecedent and response topographies, observational learning, or instructional history). Additionally, something could be explaining Inc-BiN (i.e. observational learning; Greer et al., 2006).

Chapter 4: Incidental Bidirectional Naming and Equivalence-Based Instruction

4.1 Brief Literature Review

Experiment 2 highlighted the differences in emergent intraverbal responding between participants with different degrees of Inc-BiN after relational training (matrix instruction). In Experiment 2, researchers investigated one type of emergent intraverbal response (generalized intraverbal responding). This led researchers to questions about if these same effects would occur after a different type of relational training (Equivalence-Based Instruction). In Experiment 3, researchers investigated if degree of Inc-BiN was positively correlated to other types of emergent intraverbals. Relational Frame Theory researchers classify some emergent relations as mutually entailed and combinatorally entailed. However, previous research has not compared degree of Inc-BiN and emergent intraverbal relations (mutually entailed and combinatorally entailed). Intraverbals are important to study because of their educational and social significance for participants (Jennings et al., 2021).

4.1.1 Stimulus Equivalence

Stimulus equivalence (SE) research also deals with emergent relations. Like matrix training, SE examines how the teaching of particular relations learned result in novel relations with no specific training. However, SE is also different. While matrix training focuses on teaching sufficient exemplars, SE focuses on relations between stimuli and the number of stimulus-stimulus relations that need to be taught for an individual to demonstrate whole classes of responding. *Stimulus equivalence* (SE) research was first developed by Murray Sidman (Sidman, 1994). SE researchers experiment the concept of equivalence and classify stimulus-stimulus relations as either *symmetric*, *transitive*, or *equivalent* (Fryling et al., 2020, Ch. 10).

Symmetric relations are a reciprocal relation between two stimuli in which an individual may be taught that stimulus A is equal to stimulus B and then emits an emergent response that stimulus B is equal to stimulus A. For example, an individual may learn that the image of a cat (A) is equal to the printed word CAT (B) and then emit an emergent response that the printed word (B) is the same as the image (A). Transitive relations occur when there are two stimuli not directly taught together and are related through a mutual stimulus. For example, an individual may learn that the image of a dog (A) is equal to the printed word DOG (B) and that the printed word DOG (B) is equal to a real dog (C). A transitive relation occurs if the individual emits an emergent response that the image of the dog (A) is equal to the real dog (C). Equivalence involves a combined symmetric and transitive relation and a bidirectional transfer and novel association. For example, an individual may learn that the image of a bird (A) is equal to the printed word BIRD (B) and that the printed word BIRD (A) is equal to a real bird (C). An equivalence relation occurs if the individual emits an emergent relation that the real bird (C) is equal to the picture of the bird (A).

4.1.2 Relational Frame Theory

Relational Frame Theory (RFT) looks more broadly at types of behavior principals and emergent relations than SE does (Gibbs & Tullis, 2021). RFT researchers study the development of language and cognition, including the formation of concepts (Blackledge, 2003; Hayes, et al., 2001; Hayes, et al., 2004). Relational Frame theorists regard their theory as a post-Skinnerian approach and use Skinner’s (1957) work as a starting point to study complex relations rather than elementary verbal operants. Additionally, RFT extends equivalence research by studying more types of relations than simply equivalence. For example, RFT utilizes Equivalence Based instruction to specifically target sameness but also uses other types of instruction to target other frames e.g., distinction, comparison. Relational responding is when an organism learns two different stimuli are related in some way. RFT researchers study relational responding and separate relational responding into two categories: *nonarbitrarily applicable relational responding* and *arbitrarily applicable relational responding*. Nonarbitrarily applicable relations are sets of stimuli that are related through their physical properties (Hayes, et al., 2001). For example, if a child learns what a "tree" is, as compared to a bush, they will understand the relation between the varying size of trees and the vertical relational element is physical. Arbitrarily applicable relational responding (AARR) occurs when two or more stimuli are paired together through non-physical properties. For example, if a child hears someone point to a tree and vocally say "tree," they may later see a tree and independently say "tree" as a speaker. This is arbitrarily applicable because the emergent relation of emitting the speaker response after learning the listener is not defined by topography and not physical attributes. AARR is further defined based on properties of *mutual entailment, combinatorial entailment,* and *transformation of stimulus function* (Barnes-Holmes et al., 2004; Blackledge, 2003). Mutually entailed relations are bidirectional and occur when someone is taught that stimulus A is related to stimulus B they emit an emergent response that stimulus B is related to A. For example, if a student hears the word "denominator" (A) is taught to point to a denominator (B), then sees a denominator (B) and vocally tacts it as "denominator" (A), they emitted a mutually entailed relation.

Combinatorially entailed relations occur when an individual is taught two different relations that have one common stimulus. Then, they emit an emergent response as a third relation. For example, if a student is taught that A equals B and A also equals C, the student would emit a mutually entailed relation if they emit the emergent response that B is equal to C and/or that C is equal to B. Transformation of stimulus function is the third property of AARR that occurs when mutually entailed and combinatorally entailed relations become generalized responding repertoires to which the function of a stimulus changes in accordance with the function of stimuli it is related to (Hayes et al., 2021). For example, if math content is established as a reinforcer, transformation of stimulus function occurs when that reinforcing attribute is transferred to other coordinated stimuli, such as reading instruction (Lee-Park, 2005; Ross et al., 2006). Additionally, there are nine relational frame families that further break down types of relations: coordination, opposition, distinction, comparison, hierarchical, temporal, spatial, cordiality/causality, and deictic (Dixon & Stanley, 2020). These frames can be combined into verbal networks (Hayes et al., 2021) to facilitate emergent relations.

***Emergent Relations*.** Verbal behavior researchers have studied emergent relations and have studied them in different ways. Hart and Risley (1995) wrote a book that altered how people think about and study learning and, by extension, emergent learning, because of their reported findings that the number of word exposures as a child is correlated to later academic skills. Hart and Risley's (1995) findings supported the study of verbal behavior and emergent learning because of the common focus of learning new words. These data have served as a spark for various branches of language and cognition research groups; three of which are stimulus equivalence (SE; Fienup et al., 2010; Sidman, 1994), Relational Frame Theory (RFT; Barnes-Holmes, et al., 2014; Hayes, et al., 2001), and Verbal Behavior Development Theory (VBDT; Greer & Keohane, 2006; Greer & Ross, 2008; Pohl, et al, 2018). Researchers across all three of these domains have studied emergent responses (Barnes-Holmes & Harte, 2022; Greer & Ross, 2008; Horne & Lowe, 1997). Emergent responses are responses that are not directly taught. For example, if a student is directly taught that the words "unit fraction" mean "a fraction with one as the numerator," they may derive/emit an emergent response if they write unit fraction when given the definition “a fraction with one as the numerator.” VBDT, RFT, and SE researchers all study the development of language. VBDT and RFT researchers have extended beyond SE researchers because they both study more complex emergent relations than explained with SE.

4.1.3 Equivalence-Based Instruction

Equivalence-based instruction (EBI) is a strategic instructional design technique used by researchers and teachers to teach concepts. For example, young children learn different representations of math families such as addition and subtraction problems, families represented in number bonds, and families represented on number lines. EBI is a form of instruction in which the instructor/researcher lays out the various stimulus relations that a child should ultimately learn, plans for which relations to formally teach, and plans for which relations may emerge via logical deduction, but were never directly taught (For an example, See Table 12).

For example, Verdun, et al. (2022) used EBI to plan for and teach third graders fraction-pictogram-percentage relations. Participants learned fraction to pictogram relations through explicit instruction and were able to relate these stimuli in reverse with no additional teaching. Participants also learned percentage to pictogram relations through explicit instruction and emitted pictogram-percentage emergent relations. More importantly, after researchers taught participants how fractions and percentages were paired with pictograms, participants emitted emergent relations for fractions and percentages with no explicit instruction. Participants emitted emergent relations as a function of instructional planning and generalized operant responding. This extends EBI literature and demonstrates the effectiveness in the general education setting by promoting emergent responding through direct instruction and observational learning.

Equivalence-Based Instruction Outcomes. Lynch and Cuvo (1995) used EBI to teach fraction, ratio, and decimal relations. Researchers utilized a match-to-sample (MTS) procedure and taught seven fifth- and sixth- grade participants to match images of fractions (B) to fraction ratios (A) and match print decimals (C) to images of decimal quantities (B). Participants completed 12 equivalence tests to measure their responses to *symmetrical* and transitive relations. Additionally, participants completed generalization tests that measured responses to converting fractions to decimals and decimals to fractions. Researchers tested CA, AC, BA, and CB relations prior to intervention and tested BA, CB, AC, AC, and CA relations after intervention. Following intervention, all participants accurately matched decimals to fraction ratios (AC relations) after phase one of training. Results indicate that participants acquired *equivalence* relations. Results demonstrated EBI is an effective tool for facilitating equivalence relations for participants who demonstrated difficulty learning fractions and decimals.

Daar et al. (2015) implemented a multi-phase relational training sequence to teach equivalence class membership across three sets of intraverbal response sets (i.e., community helpers, noun-class stimuli, and wh- word relations). Researchers implemented a concurrent multiple baseline across skills across three participants with autism spectrum disorder. Researchers measured intraverbal responses to wh- questions pre- and post-intervention. Researchers assigned each community association response class into three types of wh- questions (i.e. who, where, and what). Researchers began with phase one, Community association training. During this phase researchers established three, three-member equivalence classes. During phase 2, noun association training, researchers established three, four-member equivalence classes for noun-class words. For example, researchers trained a person as doctor, place as hospital, and activity as helps people. Lastly, researchers established equivalent relations between noun-class words and wh- words (i.e. who, where, what). Two out of the three participants emitted accurate responses to wh- intraverbals following the training. Data demonstrated that intraverbals emerges as a function of relational training of specific targets and generalized operant responding (Daar et al., 2015)

Brodsky and Fienup (2018) conducted the first meta-analysis of equivalence-based instruction (EBI) literature within higher education. Researchers summarized EBI research from 31 experiments. They reported a large increase in the number of EBI experiments published at the college level. Most experiments implemented a match-to-sample (MTS) procedure during intervention. In general, Brodsky and Fienup (2018) found that EBI was more effective than no instruction. Gibbs and Tullis (2021) conducted a systematic review of studies with emergent relations with individuals with disabilities. Researchers state there has been an increase in the number of studies and interventions used to study emergent relations (i.e. Equivalence-based instruction). Although the review includes interventions other than Equivalence-Based instruction, McLay et al. (2013) and Gibbs and Tullis (2018) state that researchers should continue to study EBI across different content areas and populations other than college students.

Equivalence-Based Instruction Limitations. Most EBI research has studied the effects of match-to-sample (MTS) procedure on college participants (Brodsky & Fienup, 2018). MTS involves teaching listener responses. There are two major limitations of EBI literature. First, the effects of EBI have not been studied, as extensively, with elementary students in general education classrooms. Second, researchers have predominantly studied EBI with MTS (Fryling et al., 2020, Ch. 10). Most EBI literature implements MTS procedures that focus on listener responses, which is one type of verbal operant. Future research in EBI should study the effects of EBI as it relates to other verbal operants (i.e. intraverbals). To strengthen EBI literature, researchers should extend the existing literature by studying other populations. Specifically, researchers should study the application of EBI to other socially significant populations (i.e.

general education participants). Also, the applications of EBI should be studied with other verbal operants (Skinner, 1957) and include intraverbal responses (Jennings & Miguel, 2017; Sundberg & Sundberg, 2011)

4.1.4 Similarities in Incidental Bidirectional Naming across Theories

Incidental Bidirectional Naming (Inc-BiN) is a verbal developmental cusp that verbal behavior development researchers and relational frame theory researchers have studied (Barnes-Holmes et al., 2014; Horne & Lowe, 1997; Pohl et al., 2019; Skinner, 1957). SE researchers only address Naming, not Incidental Naming and describe equivalent relations were either the listener or speaker relation is taught and the other emerges. RFT defines Inc-BiN as a type of mutually entailed relation that rotates across the listener and speaker topographies specifically (Sivaraman, et al., 2023). RFT researchers refer to symbolic relations and the formation of equivalence classes as functionally synonymous with each other (Sivaraman, 2023). For example, Inc-BiN is considered a symbolic relation between the listener and speaker topography. Relational frame theorists explain Inc-BiN through understanding generalized operant learning, or arbitrarily applicable relational responding (AARR; Hayes et al. 2001). VBDT researchers acknowledge that Inc-BiN is a mutually entailed relation, however, they attribute the source of Inc-BiN to the joining of the listener and speaker, as opposed to RFT researchers who suggest the source of Inc-BiN to occur simultaneously with the acquisition of mutually entailed/symbolic relations. Stimulus control for Inc-BiN develops over time through various experiences and the establishment of other necessary listener and speaker developmental cusps. Inc-BiN allows an individual to emit emergent relations across the listener and speaker topography, therefore, learning in ways they could not before (Greer & Longano, 2010). VBDT researchers also state Inc-BiN is a cusp that allows an individual to emit mutually entailed relations and is acquired through instruction that rotates listener and speaker responses (Fiorile & Greer, 2007; Gilic & Greer, 2011). This differs from the RFT perspective because VBDT researchers affirm that Inc-BiN occurs first and allows for mutually entailed relations. Individuals who have Inc-BiN can emit correct speaker responses after only observing instructional models (Hranchuk, et al., 2018). VBDT researchers have expanded Horne and Lowe's (1996) work and have identified multiple types of Naming including but not limited to Incidental Unidirectional Naming, Naming by Exclusion, and Incidental Bidirectional Naming (Fryling, et al., 2020; Greer & Ross, 2008).

Current attempts to integrate these theories argue which came first (Symbolic relations, Inc-BiN, or Mutual Entailment). While stimulus equivalence research attributes symbolic relations (Inc-BiN) to the formation of equivalence classes, relational frame theorists attribute Inc-BiN to the formation of mutually entailed relations, and verbal behavior development theorist attribute Inc-BiN to the joining of an individual's listener and speaker repertoires. Although stimulus equivalence researchers, relational frame theory researchers, and verbal behavior development researchers all attribute the source of Inc-BiN to different variables, there have been intersections of modern research that encourage the integration of these fields (Fienup, 2018; Sivaraman et al., 2023). All recognize Inc-BiN as an important developmental milestone for language learning that rotates across listener and speaker repertoires. RFT and VBDT researchers have both studied Inc-BiN through the lens of an evolutionary perspective and the role cooperation plays in the formation of emergent relations and classes of responding (Sivaraman et al., 2023). Current research is being conducted on how these variables affect the acquisition of Inc-BiN/Mutual Entailment. Fienup (2018) and Sivaraman et al. (2023) stated that integration is the key to further development of the understanding of language cognition. They propose that a synthesized definition of Inc-BiN may be a step in the right direction. Morgan et al. (2020) has been one of the only experiments to investigate relations between Inc-BiN and RFT relations (e.g., mutual entailment and combinatorial entailment) in an applied study.

In their first Experiment, Morgan et al. (2020) investigated the relation between degree of Inc-BiN and emergent relations. Researchers measured participants' degree of Inc-BiN with unfamiliar and familiar Inc-BiN probes. Researchers categorized participants by degree of Inc-BiN in the same way Baldonado (2021) categorized Inc-BiN. Researchers measured emergent responses with nonarbitrary emergent relations and arbitrary emergent relations assessments. Researchers ran Pearson correlations to investigate the relation between these variables. Overall, data demonstrated a strong correlation between participants' degree of Inc-BiN and emergent responses. These findings suggest future research studying the relation between degree of Inc-BiN and emergent relations for socially significant academic content (i.e. Common Core Math Standards).

4.1.5 Rationale for Experiment 3

The present experiment extends the literature in two major ways: incorporating new verbal operants that were not previously studied in Inc-BiN literature and comparing degree of Inc-BiN with emergent responses after EBI (not MTS). Behavior analysts have traditionally studied Inc-BiN in relation to listener and tact responses (Abdool-Ghany & Fienup, in press; Friedman, 2020). Listener responses include match to sample and multiple-choice response topographies. Tact responses include vocally naming a stimulus. However, much of students' academic performance is measured with intraverbal responses. This research extends the current body of literature by studying Inc-BiN as it relates to emergent intraverbal responses. Previous research has also investigated the relation between Inc-BiN and relational training interventions that implement MTS (Morgan et al., 2020). MTS naturally requires a form of listener response from the individual. The present experiment extends the literature by measuring the relation between Inc-BiN and a different procedure, Equivalence-Based Instruction.

Experiment 3 investigated the relation between participants' degree of Inc-BiN and emergent intraverbal responses (mutually entailed and combinatorally entailed) following Equivalence-Based Instruction. Researchers implemented a relational training intervention where they taught two relations (A to B and A to C) and tested directly taught and emergent relations. Researchers also investigated the relation between participants' degree of Inc-BiN and number of sets to criterion. Researchers implemented a pre/post design with thirteen participants in 4^{th}- and 5^{th}-grade classrooms.

Incidental Bidirectional Naming and Equivalence-Based Instruction

4.2 Methods

4.2.1 Participants and Setting

The researcher selected thirteen, fourth- and fifth-grade participants from two schools outside major metropolitan areas. Researchers selected fourth- and fifth-grade participants who demonstrated instructional readiness. Individuals who were frequently pulled to go to the principal's office were excluded from the experiment. The setting was the same as in Experiment 1 (CABAS® AIL® 4th-grade classroom). Participants ranged from 9-year-olds to 11-year-olds. Seven participants were female and seven were male. Participants' parents/guardians self-reported their race/ethnicity to the participants' school: five identified as white, one identified as Asian, six identified as Hispanic, and one identified as Black. Four participants had an individualized education plan (IEP), one participant had a 504, one participant was classified as an English as second language student, and one participant was classified as intervention and referral services (I & RS). As part of the participants' school-wide assessments, each participant completed an *i-Ready® K-12 Adaptive* Math Diagnostic, a computer-based, adaptive learning and assessment tool (Curriculum Associates, 2014). On the *i-Ready®* math diagnostic participants ranged from a 410 to 532, with a mean of 461. These scores mean that participants ranged from a Grade 1 to a Grade 5, with a mean of grade level 3.3. Participants demonstrated a need for additional math instruction because most participants were performing below grade-level. The researcher measured participants' Inc-BiN repertoires for familiar and unfamiliar stimuli using researcher-created assessments. For familiar Inc-BiN assessments; five participants demonstrated Pre-UniN, five participants demonstrated Inc-UniN, and three participants demonstrated Inc-BiN. For unfamiliar Inc-BiN assessments; 11 participants demonstrated Pre-UniN, 2 demonstrated Inc-UniN, and 0 demonstrated Inc-BiN. See Table 10 for summarized participant information and Table 11 for individual participant information.

Table 10

Summarized participant information; sex, race, learning classification, and Inc-BiN classification for familiar and unfamiliar stimuli.

Variable	*N*	Percent
Sex	M = 6	M = 46%
	F = 7	F = 54%
Race	W = 5	W = 38%
	A = 1	A = 7%
	H = 6	H = 46%
	B = 1	B = 7%
Learning Classification	IEP = 4	IEP = 30%
	504 = 1	504 = 7%
	ESL = 1	ESL = 7%
	I & RS = 1	I & RS = 7%
Inc-BiN Classification: Familiar	Pre-UniN = 5	Pre-UniN = 38%
	Inc-UniN = 5	Inc-UniN = 38%
	Inc-BiN = 3	Inc-BiN = 23%
Inc-BiN Classification: Unfamiliar	Pre-UniN = 11	Pre-UniN = 85%
	Inc-UniN = 2	Inc-UniN = 15%
	Inc-BiN = 0	Inc-BiN = 0%

Note. W = white, M = Mixed, H = Hispanic, B = Black, A = Asian, IEP = Individualized Education Plan, ESL, English as Second Language, I&RS = Intervention and Referral service

Table 11

Individual participant information

					Degree of Naming: Unfamiliar			Degree of Naming: Familiar			*iReady®* Math Diagnostic		*iReady®* Reading Diagnostic	
Participant	Age	Gender	Race	Learning Classification	Classification	Total Listener (40)	Total Speaker (40)	Classification	Total Listener (40)	Total Speaker (40)	Scaled Score	Grade Level	Scaled Score	Grade Level
1	9	M	B	I&RS	Pre-UniN	13	0	Pre-UniN	30	16	424	2	514	3
2	10	F	W	IEP	Pre-UniN	24	0	Pre-UniN	29	5	410	1	460	1
3	9	F	H	IEP, ESL	Pre-UniN	17	2	Pre-UniN	17	11	465	4	542	3
4	10	F	W	IEP	Pre-UniN	9	6	Pre-UniN	20	9	436	3	521	3
5	9	F	H		Pre-UniN	18	6	Inc-UniN	34	9	461	3	559	4
6	9	M	H	IEP	Pre-UniN	23	6	Inc-UniN	38	13	414	2	509	3
7	11	M	W		Inc-UniN	34	6	Pre-UniN	29	8	449	3	488	2
8	9	M	W		Pre-UniN	28	15	Inc-UniN	36	21	479	4	583	4
9	10	F	W		Pre-UniN	24	16	Inc-UniN	32	21	477	4	644	6
10	9	F	H	504	Pre-UniN	23	17	Inc-BiN	38	40	474	4	529	3
11	10	M	H		Pre-UniN	31	19	Inc-UniN	35	19	486	4	587	4
12	9	F	H		Pre-UniN	27	19	Inc-BiN	40	39	492	4	623	4
13	9	M	A		Inc-UniN	36	30	Inc-BiN	39	40	532	5	659	7

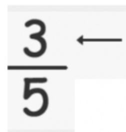
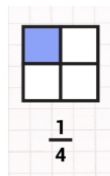
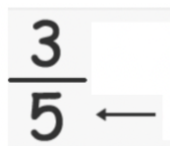
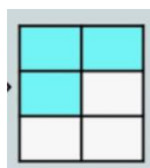
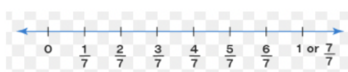
The classroom also used a station rotation model in which participants spent 25-30 min at any given station before rotating to the next station, which focused on a different academic skill. There were three classroom stations: standards-based lessons, intervention, and independent work. The researcher conducted the experiment during the intervention station. The present experiment was embedded within this station rotation model. The researcher split participants into three groups, based on instructional level. Researchers determined instructional groups by the instructional level reported by the *i-Ready*® Diagnostic Assessment data. The researcher selected participants with similar academic performance and placed them in the same group. Participants with high academic performance were in the same group, while participants with moderate academic performance were in the same group, and participants with low academic performance were in the same group.

4.2.2 Materials

Academic Stimuli in Math Relations Tests and Intervention. The Math Relations Tests tested six different relational responses across five operants (numerator, denominator, fraction number line, fraction visual model, unit fraction). The researcher identified these operants as "key words" within the participants' *i-Ready*® curriculum. *i-Ready*® is a math curriculum and diagnostic tool that is used across the participants' school district. The researcher selected the pictures and definitions for each operant from models *i-Ready*® had previously created and pulled images from other sources when necessary (i.e. Google©). See Table 12 for stimuli used.

Table 12

Operant/Relation Combinations

A	B	C
Numerator		number that shows the number of parts of the whole
Unit Fraction		a fraction whose numerator is 1
Denominator		the number that shows how many equal parts the item has been divided into
Fraction Visual Model		a picture that shows the parts of a whole and number of parts in a whole
Fraction Number Line		shows a line with numbers on the lines that are all fractions

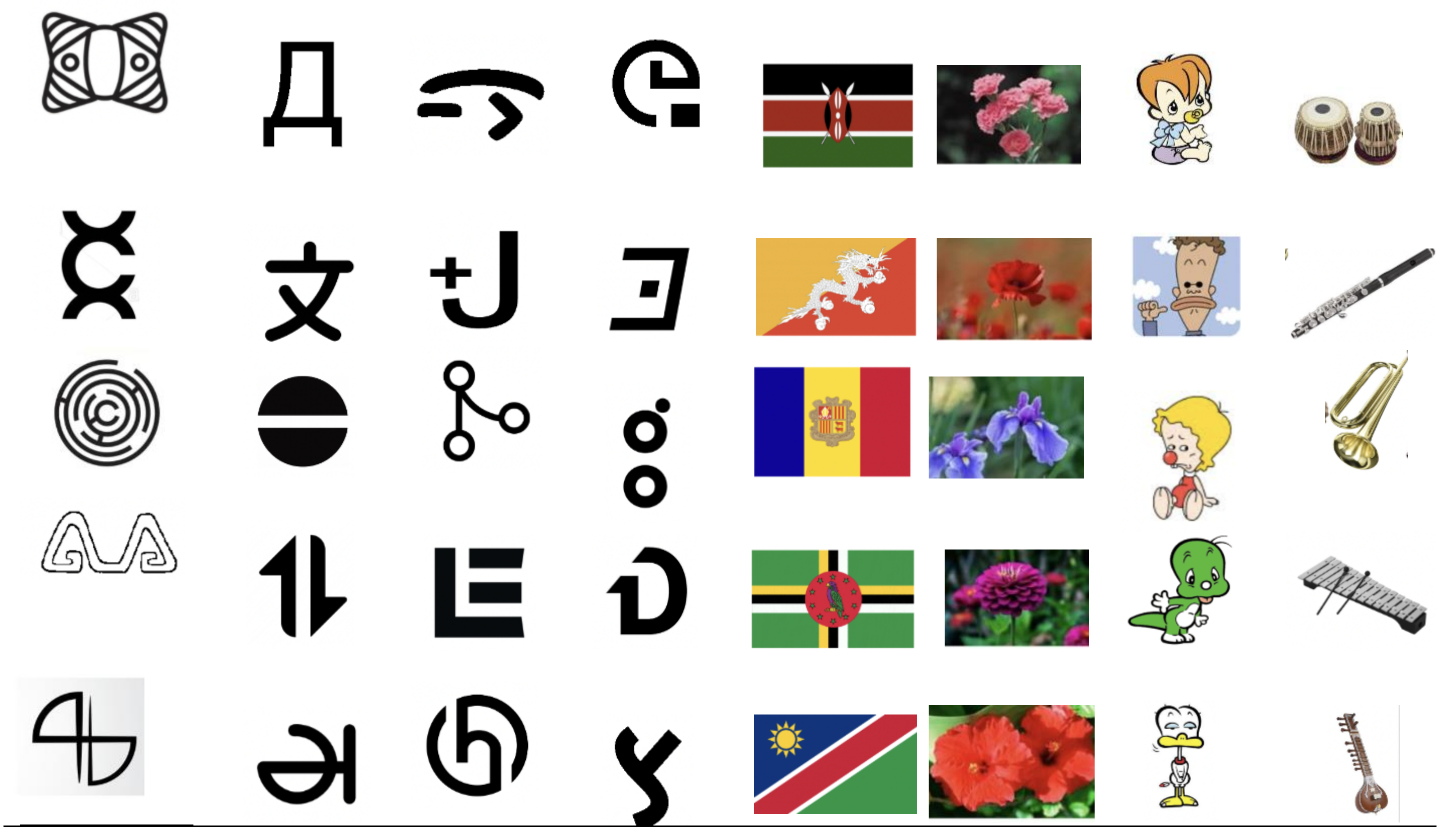
Math Relations Tests. The researcher created the six Math Relations Tests based on the five stimulus classes, each comprised of a printed word (A), picture of a word (B), and definition of a word (C). There were six Math Relations Tests to test for each relation in a class (A to B, A to C, B to A, C to A, B to C, and C to B). A represented the written word, B represented the picture of the word, and C represented the definition of the word (See Figure 11 for types of relations and Table 12 for operant relations combinations). The researcher created Math Relations Tests on GoogleDocs© and printed them for participants to complete using pencil and paper. Each Math Relations Test consisted of each of the five stimuli being presented in either the A (word), B (picture), or C (definition) topography. Additionally, there were printed instructions that stated which topography the participants should write. For example, when participants completed Math Relations Test A to B, each operant was presented in the printed word (A) topography, and the written instructions read "Draw the picture for the word" (B). The five remaining Math Relations Tests were structured in the same format.

Brief Incidental-Bidirectional Naming Assessment. For Brief Inc-BiN Assessments, the researcher used materials identically as in Experiment 1 and 2 except the current experiment also measured Inc-BiN for familiar stimuli. There were two types of Brief Inc-BiN Probes: Familiar and Unfamiliar. Familiar and Unfamiliar Inc-BiN Probes followed identical structures and procedures except for the type of stimuli in those probes. Familiar stimuli were selected because of an assumed history of experiences with types of stimuli. In other words, familiar stimuli were stimuli that participants could potentially encounter in their environment. For example, Familiar Inc-BiN stimuli included types of instruments, cartoon characters, flowers, and flags of countries (See Table 13). Unfamiliar stimuli were selected based on an assumed lack of history of experiences with these types of stimuli. Unfamiliar Inc-BiN stimuli were stimuli that participants were unlikely to contact in their natural environment. For example, unfamiliar stimuli were all black and white, and were obscure symbols. The researcher selected some unfamiliar stimuli from Greek and Chinese letters, as well as archeological symbols.

Table 13

Unfamiliar and Familiar Inc-BiN Stimuli

Unfamiliar				Familiar			
Set 1	Set 2	Set 3	Set 4	Set 1	Set 2	Set 3	Set 4



Intervention Materials. The researcher created intervention materials using GoogleSlides© and PearDeck©. The researcher made one slide for each operant/relation combination (A to B and A to C). For example, the operant "numerator" had one slide for teaching the A to B relation and one slide for teaching the A to C relation. In total, each intervention set included 10 slides. See Figure 9 for an example of an A to C relation slide and Figure 10 for an example of an A to B relation slide.

Figure 9

Sample intervention slide with antecedents for A to C relation.

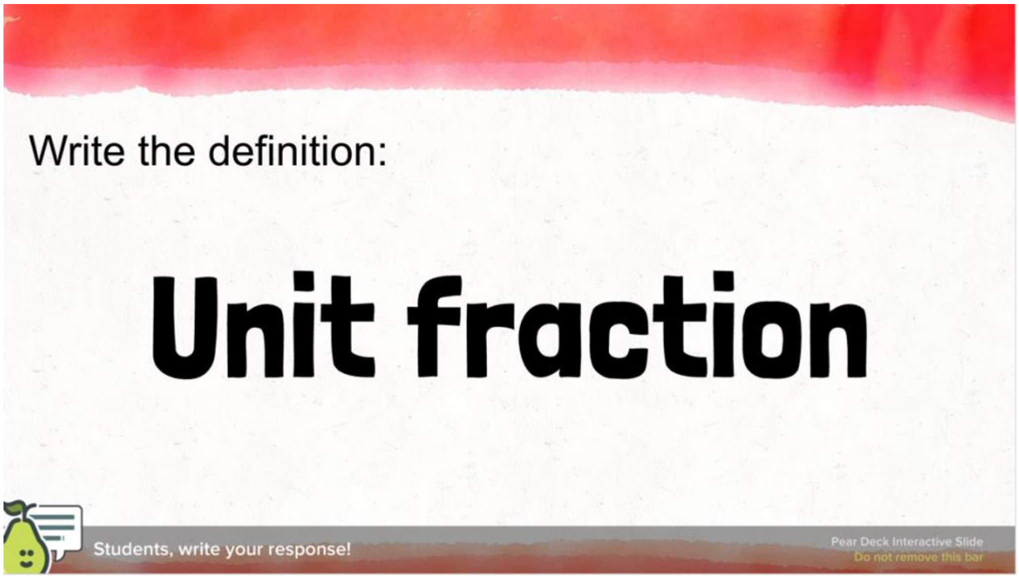
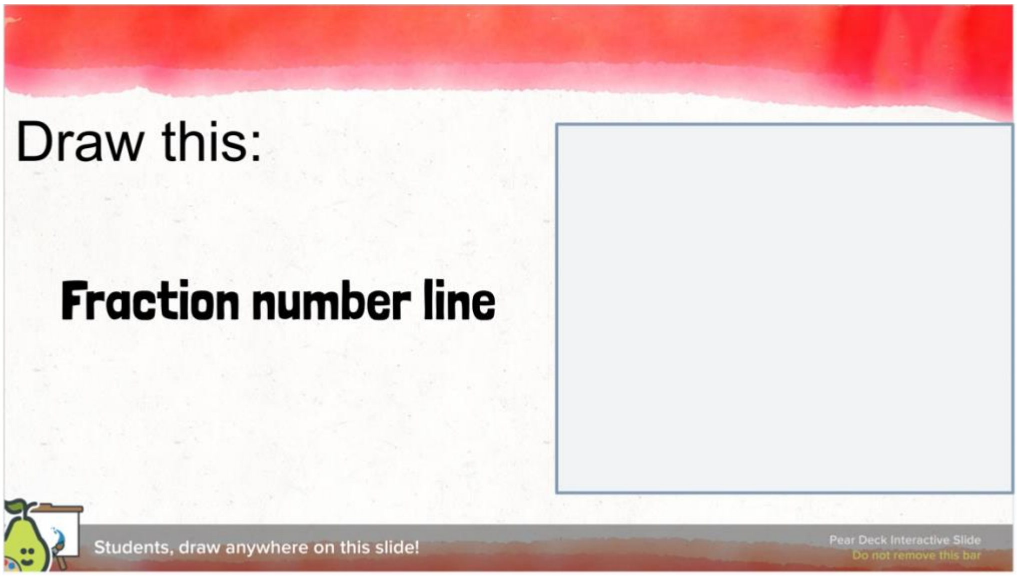


Figure 10

Sample intervention slide with antecedents for A to B relation.



4.2.3 Dependent Variables

Math Relations Tests. The researcher measured the number of correct responses to six individual relations (A to B, A to C, B to A, C to A, B to C, and C to B; See Table 12 for a teacher script). Each Relation Test had a total of five questions. The researcher marked a correct response (plus) if the participant emitted the target response that corresponded to the pre-scripted correct response. The researcher also marked a correct response if the response had all the essential components of the target stimuli. This included responses that were functionally equivalent to the target stimulus. For example, the researcher marked a correct response for the target “denominator” if the participant wrote “danomenator.” The researcher marked an incorrect response (minus) if the participant emitted a partial correct response, incorrect response, or no response. Correct definitions are listed in Table 12. On each Math Relations Test, the researcher summed the number of correct responses out of five and reported that number as their score.

To measure directly taught relations, the researcher summed the number of correct responses for Math Relations Tests A to B and A to C. To measure mutual entailment, the researcher summed the number of correct responses for Math Relations Test B to A and C to A. To measure combinatorial entailment, the researcher summed the number of correct responses for Math Relations Tests B to C and C to B (see Figure 15).

Brief Incidental-Bidirectional Naming Assessment. The researcher measured participants’ degree of Inc-BiN using four unfamiliar Inc-BiN probes and four familiar Inc-BiN probes. The researcher measured and classified participants’ Inc-BiN identically to Experiments 1 and 2 except in the present experiment, the researcher conducted four familiar probes and four unfamiliar probes. The researcher summed the number of correct responses and reported that number for the corresponding Brief Inc-BiN Probe type. Meaning, the researcher summed the total number of familiar listener responses, familiar speaker response, unfamiliar listener responses, and unfamiliar speaker responses.

Sets to Mastery. The researcher measured participants' number of sets to mastery during intervention. Researchers counted one set each time the participant completed the ten target learn units in the set. Researchers calculated sets to mastery by recording the total number of sets delivered to each participant before they met the intervention mastery criterion. Researchers defined mastery criterion as 90% or greater accuracy across two consecutive sets. If participants emitted the same antecedent-response relation incorrectly, for example the A to B relation for numerator, the researcher required the participant to complete another set until the participant either emitted 90% with a different antecedent-response relation that was incorrect or until they emitted 100% accuracy. Researchers counted the total number of sets the participant completed until they emitted mastery criterion. For example, Participant 13 required 5 sets to mastery. Participant 13 completed three sets at 50%, 60%, and 80% correct responses, then on set four and five emitted 100% and 100% correct responses. The participant completed a total of five sets to meet mastery criterion. Researchers used this method to calculate number of sets to mastery for each participant.

4.2.4 Procedure

First, participants, completed pre-intervention probes across all three of the dependent variables (Math Relations Test, Brief Unfamiliar Inc-BiN probe, and Brief Familiar Inc-BiN Probes). Once complete, the researcher started the first group in intervention and delivered learn unit instruction to teach relations A to B and A to C only. Once participants mastered each antecedent-response relation, for a total of ten opportunities, participants completed the post-intervention probes (Math Relations Test, Brief Unfamiliar Inc-BiN probe, and Brief Familiar

Inc-BiN Probes). This procedure continued systematically across all three groups and will be discussed in further detail in the following sections.

Math Relations Test. The researcher gave participants one Relation Test at a time to complete with paper and pencil. Each Relation Test had its own set of brief instructions. For example, Relation Test "A to B" had the printed instructions "Draw the picture of the word," whereas Relation Test "B to C" had the printed instructions "Write the definition for the picture." The researcher provided an additional vocal instruction and said, "Do the best you can. if you don't know something you can write 'I don't know.'" Once participants completed one Relation Test, they handed it to the researcher and the researcher handed them their next Math Relations Test until all six were complete (A to B, A to C, B to A, B to C, C to A, and C to B) so that each participant completed only one at a time. Participants completed Relation Tests at their individual desks and the researcher did not consequate any of the participant responses. Participants completed each Relation Test between five and ten minutes.

Brief Incidental-Bidirectional Naming Assessment. The researcher administered four Brief Inc-BiN probes before intervention and four Brief Inc-BiN probes after intervention, to each participant, for a total of eight Brief Inc-BiN probes. The pre-intervention Inc-BiN probes consisted of two sets of familiar stimuli and two sets of unfamiliar stimuli. The post-Inc-BiN probes also consisted of two sets of familiar stimuli and two sets of unfamiliar stimuli. The researcher implemented the same procedure for Brief Inc-BiN assessments as in Experiment 1 and 2.

Relational Training Intervention. The researcher conducted intervention in small groups with five to seven participants in each group. The researcher instructed the participants to bring their Chromebooks and a pencil to a class U-table. The researcher implemented learn units for all aspects of relational training intervention. Next, the researcher instructed participants to log in to the website for the intervention (joinpd.com). Next researchers asked the participants to input the six-digit class code that the website provided to log in. Once all the participants logged in to PearDeck© the researcher began instruction. The researcher rotated between the five stimuli (numerator, denominator, unit fraction, fraction visual model, and fraction number line) as well as between the two relations (A to B and A to C). The researcher presented the PearDeck© slide, read the directions out loud to the participants, and asked the participants to respond within one to two minutes. The researcher presented either A to B or A to C relations across all five stimuli. Participants were required to independently respond, on their own computers, and tell the researcher when they were finished writing/or drawing. If the participant emitted a correct response, the researcher provided vocal reinforcement and recorded a plus on their data sheet. If the participant emitted an incorrect response, the researcher implemented an error correction procedure by modeling the correct response on a whiteboard and then re-presented the same antecedent. The researcher repeated the correction procedure up to three times or until the participant emitted a correct response. Intervention sessions lasted between 25 and 40 minutes. Participants completed between five and 15 learn units per intervention session. If the participant did not respond after 2 minutes, the researcher implemented the correction procedure. If one participant emitted the correct response and other participants emitted an incorrect response, the researcher told the participant who emitted the correct response to wait quietly until the next question and then completed the correction procedure for the participants who emitted the incorrect responses. Participants continued intervention until they met criterion. Criterion was 90% accuracy across two sessions, contingent on the incorrect responses being different operant/relation combinations between session one and session two. Once participants met criterion, the researcher conducted post-intervention probes for both Brief Inc-BiN Probes and Relations Tests. Once one entire group met criterion for intervention, the next intervention group began intervention.

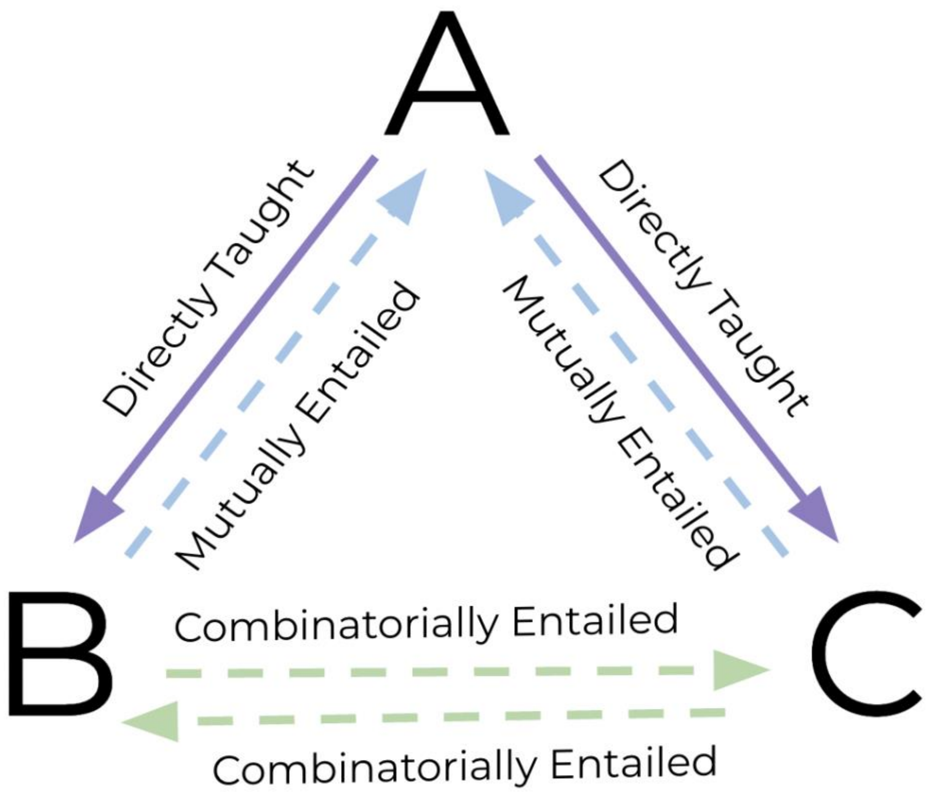
4.2.5 Statistical Analysis

The researcher implemented a pre/post design across 13 participants with independent mastery criterion. All participants completed probes before intervention at the same time. Initial probes included all Math Relations Tests, two Unfamiliar Brief Inc-BiN Probes, and two Familiar Brief Inc-BiN Probes. Once all participants completed probes, the first group began intervention. Participants in Group 1 completed instruction until each participant met the mastery criterion of 90% across two sets, contingent on the incorrect response being different than their previous incorrect response. Once all the participants in Group 1 achieved the intervention criterion, they completed all probes, and the next group began intervention.

The researcher analyzed the data by running spearman correlations that analyzed the relation between participants' degree of Inc-BiN and Math Relations Tests. The researcher measured three different response types within Math Relations Tests (direct relations, emergent relations; mutually entailed, and combinatorally entailed). The researcher conducted nonparametric spearman correlations to test the association between degree of Inc-UniN and directly taught relations, Inc-UniN and emergent relations, Inc-BiN and directly taught relations, and Inc-BiN and emergent relations (See Figure 14 for a diagram of the relations).

Figure 11

Emergent relations, following mastery of A to B and A to C.



4.2.6 Interobserver Agreement and Treatment Fidelity

Two researchers independently collected data on participants' Brief Inc-BiN probe responses. The researcher calculated Interobserver Agreement (IOA) for listener and speaker responses to Inc-BiN probes. The researcher used trial-by-trial agreement and divided the number of agreed-upon trials by the total number of trials. They divided the number of agreed-upon trials by the total number of opportunities to respond and then multiplied the result by one hundred to make a percentage. The researcher conducted IOA for 65% of Inc-BiN assessments, with 99% agreement that ranged from 90% to 100%.

The researcher also collected permanent products from all the unconsequeated Math Relations Tests. Two researchers independently collected data on participants' responses. We divided the number of agreed-upon trials by the total number of opportunities to respond and then multiplied the result by one hundred. ISA was conducted for 97% of Math Relations Tests with 99.9% agreement with a range of 90% to 100%.

Treatment fidelity was collected during intervention sessions to measure the accuracy of intervention implementation. Treatment fidelity was conducted with the use of Teacher Performance Rate and Accuracy forms (TPRA; Ingham & Greer, 1992). The researcher used TPRAs to measure the accuracy of an instructor's delivery of antecedents and consequences. The researcher conducted TPRAs by observing another individual deliver the intervention. TPRAs were conducted to ensure accurate implementation of the intervention. The researcher calculated the percentage of sessions with treatment fidelity. The researcher calculated this by summing the number of trials observed with a TPRA and dividing that number by the total number of trials the instructor delivered. The researcher calculated treatment fidelity agreement by summing the number of correctly delivered trials and dividing that number by the total number of trials observed. Treatment fidelity was conducted for 42.97% of the Intervention with 100% agreement.

4.3 Results

Table 14 displays participants' Inc-BiN classification for unfamiliar and familiar stimuli, sets to mastery, raw scores on math relations test before and after intervention, and the change in their pre-intervention to post-intervention raw score. Participants are listed in order by their change scores from pre-intervention to post-intervention. Change score data demonstrated that all participants emitted novel responses following intervention. Participants' number of sets to mastery gradually decrease as you move down the Table 14.

Table 14

Displays participants' Inc-BiN classification for unfamiliar and familiar stimuli, sets to mastery, raw scores on math relations test before and after intervention, and the change in their pre-intervention to post-intervention raw score.

Participant	Classification (Unfamiliar)	Classification (Familiar)	Sets to Mastery	Pre-Intervention Probe Score (30)	Post-Intervention Probe Score (30)	Change from pre to post Score (0-30)
4	Pre-UniN	Pre-UniN	8	4	11	7
1	Pre-UniN	Pre-UniN	13	0	12	12
2	Pre-UniN	Pre-UniN	15	0	12	12
5	Pre-UniN	Inc-UniN	13	0	13	13
8	Pre-UniN	Inc-UniN	7	6	19	13
6	Pre-UniN	Inc-UniN	19	0	14	14
10	Pre-UniN	Inc-BiN	10	6	21	15
13	Inc-UniN	Inc-BiN	5	13	30	17
7	Inc-UniN	Pre-UniN	12	1	19	18
9	Pre-UniN	Inc-UniN	12	8	26	18
12	Pre-UniN	Inc-BiN	6	8	27	19
3	Pre-UniN	Pre-UniN	15	2	23	21
11	Pre-UniN	Inc-UniN	7	8	30	22

Figure 12 displays the relation between percentage of correct listener/speaker responses for familiar (top row) and unfamiliar stimuli (bottom row) as compared with the total number of emergent intraverbal responses. The top left graph displays the relation between percentage of correct listener responses (Familiar stimuli) as compared with the total number of emergent intraverbal responses. The researcher ran a nonparametric spearman correlation and found a non-significant positive correlation between variables, $r = .390$, $p = .09$. The top right graph displays the relation between percentage of correct speaker responses (Familiar stimuli) as compared with the total number of emergent intraverbal responses. The researcher ran a nonparametric spearman correlation and found a significant positive correlation between variables, $r = .637$, $p = .01$. The bottom left graph displays the relation between percentage of correct listener responses (Unfamiliar stimuli) as compared with the total number of emergent intraverbal responses. The researcher ran a nonparametric spearman correlation and found a significant positive correlation between variables, $r = .605$, $p = .01$. The bottom right graph displays the relation between percentage of correct speaker responses (Unfamiliar stimuli) as compared with the total number of emergent intraverbal responses. The researcher ran a nonparametric spearman correlation and found a significant positive correlation between variables, $r = .814$, $p = .00$.

Figure 12

Participants' percentage of correct listener/speaker responses on unfamiliar and familiar Inc-BiN probes compared with emergent intraverbal responses

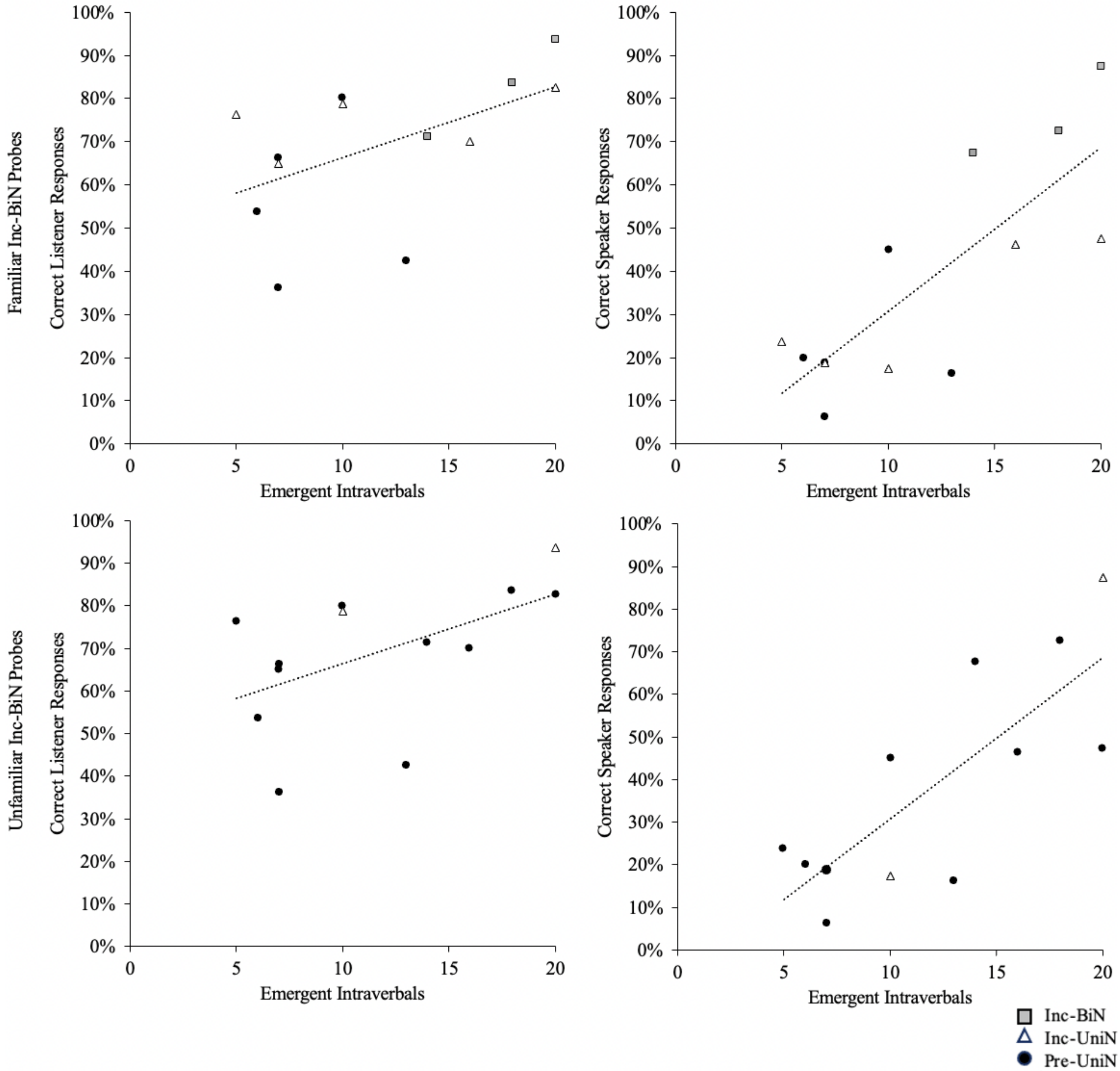
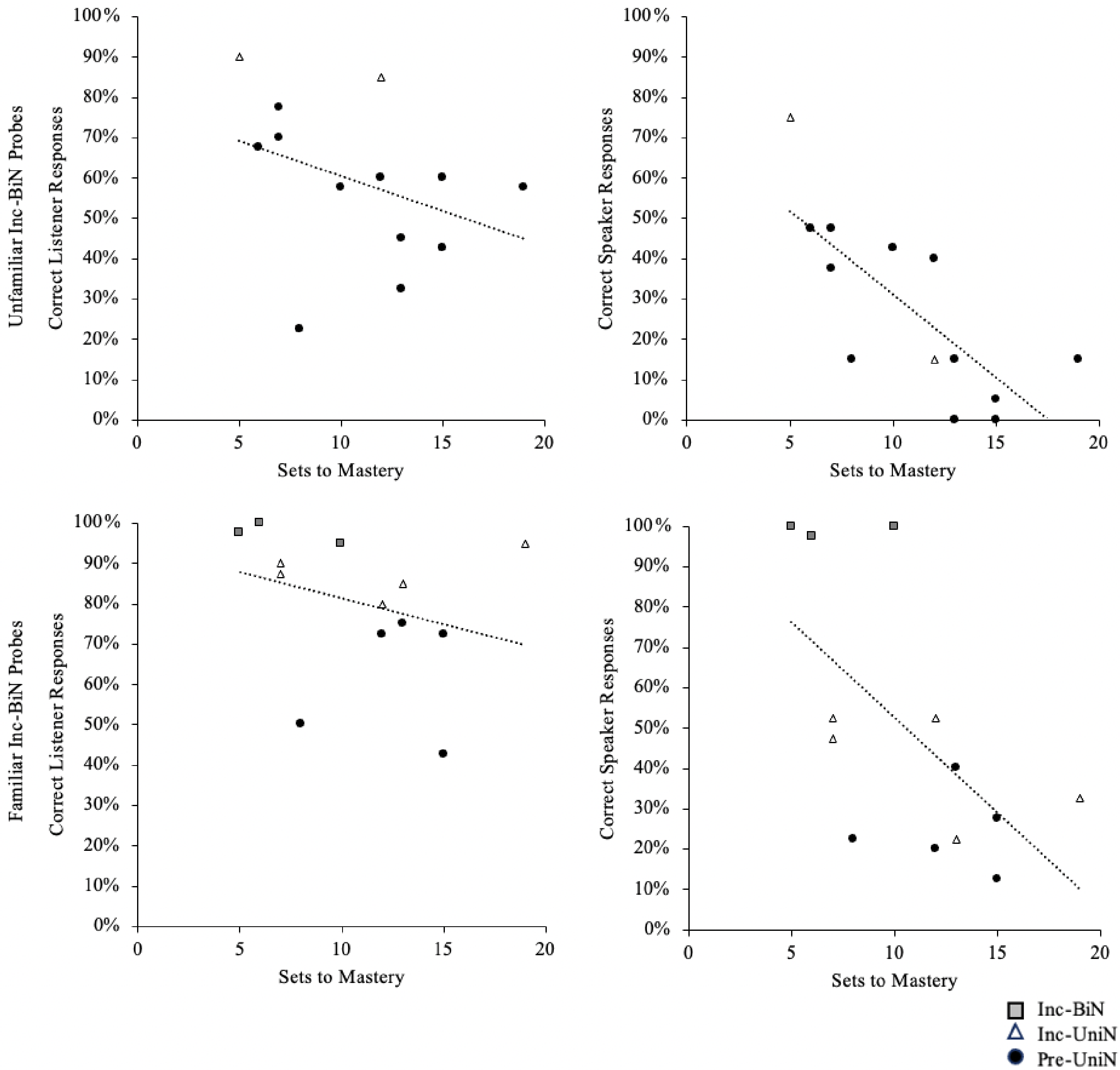


Figure 13 displays participants' number of correct responses on Inc-BiN probes with number of sets to mastery. The left graphs displays the correlation between participants' percentage of correct listener responses with number of sets to mastery. The right graphs displays the correlation between participants' percentage of correct speaker responses with number of sets to mastery. The top graphs display correlations between listener/speaker responses on unfamiliar Inc-BiN probes with number of sets to mastery. The bottom graphs display correlations between listener/speaker responses on familiar Inc-BiN probes with number of sets to mastery. The researcher ran a one-tailed nonparametric spearman correlation and found a significant negative correlation between number of correct listener responses (unfamiliar) and sets to mastery, $r = -.548$, $p = .026$. The researcher ran a one-tailed nonparametric spearman correlation and found a significant negative correlation between number of correct speaker responses (unfamiliar) and sets to mastery, $r = -.824$, $p = .00$. The researcher ran a one-tailed nonparametric spearman correlation and found a significant negative correlation between number of correct listener responses (familiar) and sets to mastery, $r = -.515$, $p = .036$. The researcher ran a one-tailed nonparametric spearman correlation and found a significant negative correlation between number of correct speaker responses (familiar) and sets to mastery, $r = -.642$, $p = .01$.

Figure 13

Displays correlations between participants' percentage of correct listener and speaker responses on unfamiliar and familiar Inc-BiN probes compared with participants' number of sets to mastery during intervention



4.4 Discussion

Results demonstrated a correlation between degree of Inc-BiN and emergent intraverbal responses. There are clear differences in the number of emergent relations between the Pre-UniN, Inc-UniN, and Inc-BiN participants. Pre-UniN participants acquired the fewest number of relations, Inc-UniN participants acquired the next highest number of relations, and the Inc-BiN participants emitted the greatest number of relations. Inc-UniN and Inc-BiN participants emitted greater than 80% of the directly taught relations in the post-intervention probes while the Pre-UniN participants emitted 60% of the directly taught relations in the post-intervention probes. Pre-UniN participants demonstrated the least number of mutually entailed relations (2.25), Inc-UniN participants emitted the second most (6.33), while the Inc-BiN participants demonstrated the greatest number of mutually entailed relations (7.33). Pre-UniN participants demonstrated the fewest number of combinatorally entailed relations (6)), Inc-UniN participants emitted the second least (5), while the Inc-BiN participants demonstrated the greatest number of combinatorally entailed relations (10).

Figure 12 displays the relation between degree of Inc-BiN (listener/speaker responses) and emergent intraverbal responses. These figures demonstrated general trends that the higher degree of Inc-BiN, the more emergent intraverbals are emitted after instruction. Participants' differences in responding, according to their degree of Inc-BiN, were consistent with the original hypothesis that the participants with Inc-BiN would emit the most, and almost all, of the emergent relations. The researcher suggests this relation is like those from Experiment 1 and Experiment 2, in that Inc-BiN may explain these correlations because of the similarity between Inc-BiN probe speaker responses and emergent intraverbal responses after EBI.

Figure 13 displays the relation between degree of Inc-BiN (listener/speaker responses) and sets to mastery. These figures demonstrated general trends that the higher the degree of Inc-BiN, the fewer the number of sets to mastery. Participants' differences in responding, according to their degree of Inc-BiN, were consistent with the original hypothesis that the participants with Inc-BiN would require the fewest sets to mastery. The researcher suggests this has positive implications for measuring Inc-BiN in academic settings because the individuals with Inc-BiN would require less instruction, which could help the instructor be more efficient with their time.

Lynch and Cuvo (1995) implemented a match-to-sample (MTS) procedure that demonstrated a functional relation between the establishment of relations and the demonstration of equivalence classes across their participants. The range of emergent responding, in which the Pre-UniN participants did not consistently emit ME and CE relations, whereas the Inc-UniN participants emitted some, but not all, and the Inc-BiN participants emitted almost all ME and CE relations, is consistent with Verbal Behavior Development Theory (Abdool-Ghany & Fienup, in press; Friedman, 2021; Greer & Ross, 2008; Lynch & Cuvo, 1995). Researchers have identified the verbal development repertoires that influence participant learning, specifically, which emergent relations individuals with Pre-UniN, Inc-UniN, and Inc-BiN can and can't emit (Greer et al., 2017; Morgan et al., 2020). Figure 12 displays differences in emergent intraverbals between individuals with different verbal behavior development repertoires and align with the explanations of Morgan et al. (2020) and verbal behavior development theory (Greer et al., 2017).

Like results from Morgan et al. (2020) there may be multiple explanations for these results. Degree of Inc-BiN could be explaining the number of emergent intraverbals, there could be another variable affecting Inc-BiN and emergent relations (i.e. observational learning, prerequisite skills), or Inc-BiN and emergent intraverbals are measuring the same repertoires. The present researcher supports Morgan et al.'s (2020) preference for the latter. Inc-BiN is an example of a mutually entailed relation which could explain why there is a significant correlation between the variables.

Verdun et al. (2020) strategically implemented EBI to measure emergent responses. The effectiveness of EBI was demonstrated, with varied effects, in the present experiment and is consistent with the findings of Verdun et al. (2020). Verdun et al. (2020) was one of the first to extend EBI and derived relation literature to general education participants. The current experiment extends the literature in important ways by implementing EBI, like Verdun et al. (2020), but incorporating emergent intraverbals and degree of Inc-BiN. However, EBI literature has not studied intraverbals as frequently as match-to-sample responses. The present experiment incorporates intraverbals to fill a gap in the current research. Intraverbals are, however, more complex responses than match-to-sample and warrant more research.

A limitation of the present experiment was that only one Math Relations Tests was conducted prior to and following intervention. For example, participants completed each Math Relations Test once before intervention and each Math Relations Test once after intervention. The researcher made this decision to control for testing effects through the probe. In the future, the researcher should implement a multiple probe design to control to maturation. To continue to control for testing effects, the researcher should conduct multiple delayed probes but only begin participants' intervention if their responding is stable and low.

Many researchers have contributed important insight into EBI and derived relational responding but left us with questions about what made these practices effective (Brodsky & Fienup, 2020; Fienup et al., 2015; Petursdottir & Oliveira, 2020; Shawler et al., 2023). Based on the findings of the current experiment, research should continue to investigate the relation between degree of Inc-BiN and emergent responding and how that should inform instruction. The present data demonstrated that participants with Inc-UniN did not emit all the ME and CE relations. This tells us that the intervention was only partially effective. Future research could use EBI to teach Inc-UniN participants four relations to mastery and test if they emit the other two emergent relations. Additionally, Inc-BiN participants emitted most of ME and CE relations. Future research could investigate if Inc-BiN participants can emit ME, CE, and transformation of stimulus function (TSF) when only presented with a model of one relation.

Results support previous findings that suggest a correlation between degree of Inc-BiN and Derived Relational Responding (Abdool-Ghany & Fienup, in press; Friedman, 2020; Morgan et al., 2021). The present experiment pushes the literature further by incorporating intraverbals. Previous research has only studied matching or selection responses. Future research should build of the present experiment to further analyze the relation between Inc-BiN and emergent intraverbals.

Chapter 5: General Discussion

The purpose of this research was to evaluate the intersection of a specific developmental capability, Inc-BiN, and both general academic performance and performance following specific behavioral instructional design tactics (relational training procedures) that are purported to produce emergent responding. Evaluating this intersection may help inform teachers and researchers of what instructional design strategies are appropriate for which students. This type of research helps fill gaps in the education literature that have persisted over decades (Greer, 1996). First, teachers are told to "use data" to measure performance and inform instruction – often without direct training on how to do this (Datnow & Kennedy-Lewis, 2012; Firestone & Donaldson, 2019). Second, teachers have expressed a need for materials and strategies that are specifically for low-performing students (Moore, 2007). Teachers may employ teaching tactics or curriculum design tactics in an attempt to help their students learn, although many curricula may only benefit high performing students (Brown, 2011; Orlich et al., 2010). Teachers need to teach all students, regardless of level, and they need help differentiating that instruction

I evaluated the relation between a specific verbal development capability and academic performance. The verbal behavior capability was Inc-BiN, which involves learning the names of things from observing models of object-names and not requiring direct reinforcement. Incidental language acquisition is an essential repertoire for any individual to enter the verbal community and succeed in school. It is assumed that learning repertoires typically develop through an individual's experiences across their lifespan (Greer & Longano, 2010; Hart & Risley, 1995). However, not all children develop these repertoires through naturalistic interactions in their environment. Often teachers implement their given curriculum and are left wondering why their lesson was not effective for all their students. Using students' degree of Inc-BiN as a benchmark to predict the potential effectiveness of instruction is one way to help create more effective instruction because instruction would be designed based on principles of behavior of how students learn as opposed to what students need to learn (Moran & Malott, 2006). In three Experiments, I evaluated degrees of Inc-BiN and its relation to: a) general academic performance, b) generalized atomic repertoires, and c) direct, mutually entailed, and combinatorially entailed relations. In Experiment 1, I evaluated the relation between degrees of Inc-BiN for unfamiliar stimuli and both reading and mathematics *iReady*® Diagnostic Scores. Learning listener responses incidentally (Inc-UniN) and learning speaker responses incidentally (Inc-BiN) were significantly correlated with both reading and mathematics scores. That is, as children's degree of incidental language increased, so too did academic performance. Of importance, student's degree of learning speaker responses incidentally was more strongly correlated to academic performance than learning listener responses incidentally. These correlations with broad academic achievement led to questions about specific behavior intervention tactics, in this case, matrix instruction and equivalence-based instruction. Researchers of both matrix instruction and equivalence-based instruction have promoted these instructional design tactics as efficient because they program for learning in addition to what is formally taught. However, mixed data suggest there is a continued need to study for whom these interventions work (Frampton et al., 2016; Morgan et al., 2021).

The purpose of Experiments 2 and 3 was to specifically evaluate the effectiveness of two instructional strategies based on students degree of Inc-BiN. In Experiments 2 and 3, I evaluated the effects of relational training procedures (matrix instruction and equivalence-based instruction) on emergent intraverbals based on participants' degree of Inc-BiN. In Experiment 2, I investigated the effects of teaching atomic units (matrix instruction) on generalized emergent intraverbals and correlated those outcomes with each participant's degree of Inc-BiN. In Experiment 2, I directly taught word combinations (made up of two atomic units) and measured if participants generalized, or recombined those units in ways that were not directly taught. The data demonstrated significant correlations between degree of Inc-BiN and emergent intraverbals as well as degree of Inc-BiN and the amount of instruction required to meet criterion.

In Experiment 3, I investigated the effects of a different relational training, equivalence-based instruction and correlated outcomes with each participant's degree of Inc-BiN. Equivalence-based instruction involved teaching two intraverbal relations (equivalence-based instruction) on the emergence of additional intraverbals. I displayed target words and directly taught the participants write the definition of the word and draw a picture for the target word. Following teaching, I tested if they could write the write the name given a picture or definition (mutual entailment) and draw a picture given a definition or write a definition given the picture (combinatorial entailment). Like Experiment 2, there were a mixture of outcomes, with participants emitting a range of correct and incorrect emergent intraverbals. Overall, the greater a participant's degree of Inc-BiN, the greater the number of emergent responses they emitted. Like Experiment 1, student's degree of learning speaker responses incidentally was more strongly correlated to emergent intraverbals than student's degree of learning listener responses incidentally. To further investigate the relation between verbal development and these outcomes, I also tested the relation between degree of Inc-BiN and learn units/sets to criterion in Experiments 2 and 3. The data demonstrated significant negative correlations between Inc-BiN and learn units/sets to mastery. That is, the participants who incidentally acquired more speaker and listener responses required the fewest number of learning opportunities (learn units/sets to mastery) to meet the teaching criterion. These data suggest that if an individuals who learn more speaker responses and listener responses incidentally emitter greater number of emergent intraverbals and require fewer teaching opportunities to achieve the learning criterion.

To summarize these findings, learning more listener and speaker responses incidentally was positively correlated with overall academic performance (Experiment 1), emergent intraverbal responses (Experiment 2 and 3), and negatively correlated with the amount of instruction required to learn the given skill (Experiment 2 and 3).

5.2 Major Findings

B. F. Skinner's conceptual analysis of verbal behavior laid the groundwork for decades of research on the study of verbal operants and the study of human language and communication (Greer et al., 2017; Greer & Ross, 2008; Pohl et al., 2020; Skinner, 1957). Researchers have extended his work by studying the six verbal operants, as well as higher order operants (Barnes-Holmes & Barnes-Holmes, 2000; Catania, 1995; Gilmore, et al., 2024; Greer & Keohane, 2006). Researchers have studied Inc-BiN, in part, because of the positive relation the learning repertoire has with learning without direct instruction (Gilmore et al., 2024; Sivaraman et al., 2023). Individuals with Inc-BiN learn at a faster rate than those without after only observing a model emit an object-name. Many researchers have studied Inc-BiN in relation to other variables which include rate of learning, emergent relations, and the induction of Inc-BiN. For example, researchers have studied how multiple-exemplar instruction can induce the stimulus control for Inc-BiN. With MEI, teachers rotate opportunities to respond as a listener and speaker to a specific stimulus (e.g., touching a picture of a mouse when told "find mouse!" and naming the mouse when shown the picture). By rotating listener and speaker behavior, children come to develop the stimulus control for emitting both listener and speaker responses from observation of object-name models alone. Most research evaluating correlates of Inc-BiN have measured match-to-sample responses, listener (field of three), or tact responses (Abdool-Ghany & Fienup, in press; Friedman, 2020; Morgan et al., 2021). However, this is the first experiment that specifically looks at the degree of Inc-BiN with intraverbal responses. Intraverbal responses are important in educational settings due to the high rate at which these types of responses are taught in general education. This is further supported by the high number of intraverbals found in academic assessments. The educational significance of intraverbals is highly relevant and warrants further investigation (Carp & Petursdottir, 2015).

It is crucial to identify types and degree of Inc-BiN because it is a direct related to how that individual acquires new language (Hawkins et al., 2018). Identifying individuals' degree of Inc-BiN is especially important considering the results of Experiment 1, which demonstrated a significant positive correlation between Inc-BiN and academic performance (reading and math). This means that children who can learn incidentally acquire a higher number of reading and mathematics responses than children who cannot learn incidentally. Some previous research has demonstrated the correlation between Inc-BiN and other academic variables (Baldonado, 2021; Morgan et al., 2021). Academic performance is important to measure because it assesses an individual's larger skill set in a given area (i.e. mathematics, reading). General education and special education students begin taking these assessments at an early age for a variety of reasons. These reasons may include evaluating teacher performance, assessing need for additional intervention services, or determining appropriate teaching material. However, Experiment 1 is one of the first times researchers evaluated correspondence between degree of Inc-BiN and academic performance. All three experiments demonstrated positive correlations between Inc-BiN and academic responses (*iReady*® Diagnostic test in Experiment 1 and emergent intraverbals in Experiment 2 and 3). These data demonstrate that the greater number of incidentally acquired listener and speaker responses, the better they perform on 4th grade math and reading tests. Conversely, the participants who acquired the fewest number of listener and speaker responses performed poorly on the 4th grade math and reading tests. These data suggests that learning listener and speaker responses incidentally may be an important variable in measuring academic performance. These findings are an essential step forward for bettering instructional design strategies to teach all students.

Based on these findings, the researcher sought to investigate the relationship between incidental language learning (degree of Inc-BiN) and performance of specific instructional design strategies. In Experiments 2 and 3, the researcher measured the effects of two instructional strategies on emergent intraverbal responses while measuring their degree of Inc-BiN. Correlations between Inc-BiN and academic variables were replicated after both types of instructional design strategies (teaching atomic units and teaching two direct relations). The outcomes demonstrated that the greater the number of incidentally acquired listener and speaker responses on the Inc-BiN assessments, the greater the number of emergent intraverbal responses after matrix instruction and equivalence-based instruction. Individuals with Inc-BiN tended to acquire more untaught relations, which means it was, likely, an appropriate and efficient form of instruction, as opposed to teaching all the relations directly for these children. Conversely, individuals with Inc-UniN and Pre-UniN tended to acquire fewer untaught relations. This suggests Pre-UniN and Inc-UniN students may need alternative interventions, or an alternative amount of instruction should be implemented with Inc-UniN and Pre-UniN participants because matrix instruction and EBI was not effective for them. Additionally, For the students that did not acquire all the untaught relations, they may be other variables at play (i.e. joint attention, socioeconomic status), that may have a functional relation with emergent intraverbals. These are discussed in limitations. These outcomes also help to clarify researchers claims that matrix instruction and equivalence-based instruction are efficient tacts – they are efficient for children with high degree of Inc-BiN.

Throughout all three experiments, the percentage of correct speaker responses on Inc-BiN probes were more strongly correlated with academic performance (*iReady*® Scores and emergent intraverbals) than listener responses. The present research extends research in multiple ways. This study extends Morgan et al.'s (2021) and Abdool-Ghany and Fienup (in press) research by measuring emergent intraverbals. Morgan et al. (2021) taught one relation and tested selection responses and tact responses for mutually entailed and combinatorially entailed relations. Their data found strong significant correlations between degrees of BiN and derived listener and match-to-sample relations (mutually entailed and combinatorially entailed). Abdool-Ghany and Fienup (in press) conducted a similar study examining emergent relations following tact or listener training. While Abdool-Ghany and Fienup (in press) taught tacts or listener responses only and measured the opposite. Despite looking at different emergent responses (i.e. selection, tacts, intraverbals) we found similar outcomes and overall relations between Inc-BiN and other emergent responses. Incorporating intraverbals and Inc-BiN is unique to the present paper. The participants who emitted the greatest number of speaker responses after only hearing the names of pictures emitted the greatest number of emergent intraverbals on their academic tests. Researchers hypothesize the speaker responses on Inc-BiN probes are more strongly correlated with emergent intraverbal responses because of the similarities in response topographies. In Inc-BiN probes, correct speaker responses are tact responses, which are verbal operants controlled by a visual stimulus and generalized reinforcement conditions. Intraverbals are another speaker verbal operant controlled by a given stimulus (i.e. vocal, written, visual) and generalized reinforcement conditions in which the responses does not share point-to-point correspondence with the antecedent. (Skinner, 1957). Although the form of responses was different (i.e. vocal and written) they were both production responses, as opposed to listener responses. The listener responses on Inc-BiN probes were selection responses in which individuals were required to point to the target stimulus in a field of three. Almost all participants emitted a greater number of correct listener responses than speaker responses. Results from Experiment 2 and 3 and Morgan et al. (2021) suggest that hearing the names of pictures and acquiring the corresponding speaker responses incidentally is more closely related to emergent relations in selection responses (Morgan et al., 2021) and intraverbal responses (Experiment 2 and 3). This suggests that speaker responses may require more complex stimulus control than listener responses.

The current data highlight the relation between Inc-BiN (listener and speaker) and academic performance. Specifically, that individuals with higher degrees of incidental language acquisition (Inc-BiN) perform better on 4th grade math and reading tests, acquire more intraverbal responses that were not directly taught, and learn at a faster rate. Inc-BiN, likely, plays an important role in academic performance, emergent intraverbal responses, and the effectiveness of a given lesson. Then, the effectiveness of a given lesson, and then multiple lessons, plays out on academic performance on achievement tests. However, these data cannot attribute the function of these variables to Inc-BiN. There may be other variables at play that were not directly measured in these experiments. In the present paper, researchers only used correlational data and did not use causal methods.

In terms of learning directly taught relations, Inc-BiN scores (listener and speaker) were significantly correlated with the number of learn units/sets to criterion. Researchers ran these correlations in Experiment 2 and 3 and found, again, that speaker responses were more strongly correlated to learn units/sets to criterion than listener responses were. Listener and speaker scores were negatively correlated with learn units/sets to criterion. Both Experiments 2 and 3 demonstrated the significant negative correlations between degree of Inc-BiN and learn units/sets to criterion. This suggests that the higher the degree of Inc-BiN, the fewer number of opportunities the individual needs to meet mastery criterion. Specifically, when specific behavior is directly reinforced (Experiment 2 and 3) those with Inc-BiN tended to perform better which is likely related to higher academic performance (Experiment 1). The data demonstrated that the greater the number of incidental listener and speaker responses on Inc-BiN probes, the fewer number of learn units they need to learn a directly taught skill. Conversely, the fewer the number of incidental listener and speaker responses on Inc-BiN probes the greater the number of learn units they need to learn a directly taught skill. We suggest speaker responses were more strongly correlated to learn units/sets to criterion than listener responses for the same reason as the correlations between Inc-BiN and emergent intraverbals. Individuals with Inc-BiN acquire the greatest number of speaker responses and, when asked to respond intraverbally, required the fewest amount of instruction because they learned the target responses faster than those without Inc-BiN. On the other hand, individuals with Pre-UniN acquired the fewest number of speaker responses, and when asked to respond intraverbally, required the most amount of instruction as comparted to Inc-BiN and Inc-UniN participants. It seems that higher speaker responses is related to fewer number of learn units to criterion. This indicates that students who learn tacts after only being exposed to the stimuli can also learn intraverbals faster, too. This seems logical because tacts and intraverbals are different only in terms of the antecedent conditions. These data support previous research from verbal behavior analysts that individuals with Inc-BiN learn at a faster rate and emit more emergent relations than those without Inc-BiN (Greer & Ross, 2008; Hotchkiss & Fienup, 2020; Hranchuk et al., 2019).

In Experiments 2 and 3, the researchers noticed that not all participants emitted 100% correct responses on the post-intervention probes for the directly taught relations. Individuals who learned listener and speaker responses maintained the greatest number of directly taught relations while individuals who learned fewer listener and speaker responses incidentally responses maintained the fewest number of directly taught relations. This may seem counter-intuitive because, theoretically, if an individual meets mastery criterion then that skill is considered in the individuals’ repertoire. However, researchers hypothesize this difference in responding can be explained by two main reasons: (1) the change in reinforcement contingencies and (2) the difficulty of the response. First, during intervention, researchers providing immediate consequences to the participants’ responses. If they emitted a correct response the researcher delivered reinforcement and if they emitted an incorrect response the researcher delivered a correction. During the pre-intervention and post-intervention probes (intraverbals and Inc-BiN probes), researchers did not provide any reinforcement. The absence of a consequence in these probes may explain why all participants did not maintain the directly taught relations in post-intervention probes. Furthermore, maintenance responses tended to be fewer with participants who learned fewer listener and speaker responses incidentally. This may be explained by the difference in strength of stimulus control in the training condition for students. Student’s with Inc-BiN have higher stimulus control for acquiring incidental listener and speaker relations. Behavior of students without Inc-BiN may have been reinforced by teacher delivered reinforcement which is a different stimulus control.

Previously, I discussed results of Experiments 2 and 3 in regard to the degree of Inc-BiN where here I discuss results in terms of Inc-BiN classifications. Previous research has categorized participants in this manner, and it is important to consider results in terms of degree as well as classification to better understand these data. Researchers implemented two relational training instructional design strategies: matrix instruction and equivalence-based instruction. On average, participants in both experiments acquired more untaught intraverbals when they had higher degrees of incidental language acquisition and fewer untaught intraverbals when they had lower degrees of incidental language acquisition. Researchers calculated the mean number of correct intraverbals after matrix training and after equivalence-based instruction and found that EBI was more effective than matrix instruction. Researchers used participants' Inc-BiN classifications for unfamiliar probes and found that, on average, Pre-UniN participants emitted 12% correct responses after matrix instruction and 50% correct after equivalence-based instruction. Inc-UniN participants emitted 40% correct responses after matrix instruction and 58% correct responses after equivalence-based instruction. For Inc-BiN participants, I could not compare outcomes between matrix instruction and equivalence-based instruction because there were no participants with a high degree of incidental language acquisition Experiment 3.

5.3 Implications

5.3.1 Implications for Verbal Behavior Research

This research has implications for verbal behavior research in multiple ways. Inc-BiN seems to be related to more than researchers previously studied, specifically, the relation to intraverbals. A goal of these experiments was to measure elements of verbal behavior development theory and relational frame theory, simultaneously. These data have implications for both perspectives and should be considered together to promote more inclusive and cross-disciplinary research.

Relational training procedures have been used to promote efficient instruction by selecting specific relations to teach and measuring emergent responses (Curiel et al., 2020' Frampton et al., 2016; Gibbs & Tullis, 2021). However, limited research has measured the effectiveness of these interventions as they relate to degree of Inc-BiN (Abdool-Ghany & Fienup, in press; Morgan et al., 2021). Many researchers have studied matrix instruction and Equivalence-Based Instruction. However, none have evaluated the effectiveness of these interventions as they relate to *how students learn*. The appropriateness of these interventions based on students' degree of Inc-BiN may be an important variable in explaining the mixed outcomes of these procedures (Curiel et al., 2020; Kemmerer et al., 2021). Researchers use matrix instruction to arrange atomic units and measure generalized operant responding. If their participants have generalized operant responding and acquire the emergent relations, the instruction is deemed more efficient than teaching all operants. However, several matrix instruction studies have reported mixed outcomes and can only hypothesize why other participants don't acquire those emergent relations. The researcher hypothesized that participants' incidental learning repertoires may provide an explanation for these mixed outcomes. Specifically, individuals who acquire listener and speaker behavior after only hearing the names of pictures would perform better on matrix instruction than individuals who only acquire listener responses or neither listener nor speaker responses. Specifically, relational frame theorists may benefit by including degree of Inc-BiN as a prerequisite measure in their research. While verbal behavior development theorists may benefit from recognizing and including multiple dimensions of relations as opposed to primarily looking at response topography (listener and speaker). For example, verbal behavior development theorists may choose to study the temporal aspects of Inc-BiN probes to further understand these measures.

Similarly, researchers have used Equivalence-Based Instruction to teach given relations and measure emergent relations (mutually entailed and combinatorially entailed). With EBI, researchers select target relations to directly teach and measure emergent relations. Most of EBI research measures emergent listener responses (i.e. match-to-sample, selection; Fryling et al., 2020, Ch. 10). Researchers have reported mixed outcomes, which may be explained by measuring individuals' degree of Inc-BiN. For example, researchers hypothesize that individuals with Inc-BiN also emit more emergent relations than individuals with Inc-UniN and Pre-UniN (Gilmore et al., 2024; Greer & Ross, 2008; Hawkins et al., 2018). If students learn incidentally, EBI is likely to be deemed more efficient than teaching all the relations directly. However, not all students learn incidentally, which means they need some other form or amount of instruction. From the relational frame perspective, verbal behavior development researchers may benefit from more explicitly defining relations. For example, considering which relations are directly taught and the corresponding mutually entailed and combinatorially entailed relations is important when arranging instruction.

Specifically, both matrix instruction and equivalence-based instruction were most successful with participants classified with Inc-BiN. Participants with Pre-UniN and Inc-UniN consistently emitted low to moderate levels of correct responses after relational training. This suggests that these participants need either more instruction or a different type of instruction to emit emergent intraverbals at criterion levels. Previous instruction in matrix instruction and equivalence-based instruction have demonstrated variable results (Curiel et al., 2020; Kemmerer et al., 2021). Experiments 2 and 3 offer a potential explanation for this variability; degree of Inc-

BiN. Researchers hypothesize that Inc-BiN may contribute to academic intervention performance based on the significant correlations in Experiments 1, 2, and 3.

These findings also have implications for Inc-BiN assessment procedures. Inc-BiN assessment procedures consisted of a naming experience, listener probe, and speaker probe. These assessments test individuals' incidental language acquisition after only hearing the names of the symbols. All three experiments demonstrated that speaker responses are more highly correlated with the variables measured than listener responses are. These data may spark future research that investigate variations of Inc-BiN assessments. Specifically, researchers measured emergent written intraverbals in Experiments 2 and 3 and required participants to write definitions and drawings of target words. These responses are more complex than how verbal behavior development researchers have studied intraverbals in the past. This leaves room for further questions regarding how Inc-BiN assessments can be modified. For example, future research may create a modified Inc-BiN assessment that specifically measures written intraverbal responses that require the individual to hear the definitions of novel words, and then later select and produce appropriate definitions after only hearing the word-definition relations (see intraverbal-bidirectional naming, Miguel, 2016).

5.3.2 Implications for Educators

These data have additional implications for general and special education settings. The researcher conducted Experiments 2 and 3 in small groups. In the field of behavior analysis, most research is conducted in clinical settings using single-case designs with 1:1 or 1:2 teacher-to-student ratios. These data were all collected in a public school in a general education CABAS® AIL® 4th-grade classroom. Implementing interventions in a group setting is more similar to typical conditions that occur in general education and special education classrooms and offers more practical interventions for teachers than interventions conducted in a 1:1 setting.

Inc-BiN assessments may be a helpful tool for educators to select appropriate instruction for their students. Data from all three experiments demonstrate a relation between Inc-BiN and multiple academic variables; academic performance, emergent intraverbals, and learn units/sets to criterion. All three variables are highly relevant in educational settings, however, only CABAS® teachers and CABAS® AIL® classrooms have measured Inc-BiN as it relates to selecting and modifying their instruction (Hranchuk et al., 2019). Data from these experiments demonstrated that the relational training interventions used in Experiment 2 (matrix instruction) and Experiment 3 (equivalence-based instruction) do not work for all students. Specifically, matrix instruction and equivalence-based instruction were effective for participants with higher degrees of Inc-BiN. However, matrix instruction and equivalence-based instruction was not effective for participants with lower degrees of Inc-BiN (Inc-UniN and Pre-UniN) in my studies. Individuals with Inc-UniN and Pre-UniN, typically, did not acquire the all the necessary emergent relations for those interventions to be effective and in Experiment 2 these participants also did not maintain directly taught intraverbals. Most did not learn the responses the researchers implementing the interventions are trying to teach. Inc-BiN assessments are relatively brief, especially when compared to other academic assessments such as the *iReady*® Diagnostic assessment.

Teachers may benefit from measuring their students' degree of Inc-BiN and then using that data to select/modify instruction that will be most beneficial for their students. Once degree of Inc-BiN is assessed, teachers may use relational training interventions for their students with Inc-BiN and require other instructional tactics for their students without Inc-BiN. For students without Inc-BiN, this may include teaching more relations directly. For example, in Experiment 2, we taught five out of 25 relations directly as opposed to 10 or all 25 relations. Teachers could choose to teach all 25 relations to their students with Pre-UniN. Additionally, in Experiment 3, we taught 2 out of 6 relations directly. Teachers may teach students with Pre-UniN/Inc-UniN four or all six relations.

Additionally, these data suggest researchers should focus on improving students' repertoires for how they learn. This means researchers should focus on improving verbal behavior development. Verbal behavior development research has extensively studied verbal behavior development cusps and shown their effects in establishing or increasing the relevant stimulus control (Greer et al., 2017; Greer & Ross, 2008; Pohl et al., 2020). One example of how to do this is systematically arranging for multiple exemplar instruction across listener and speaker responses (Greer & Ross, 2008; LaFrance & Tarbox, 2019). To do this, teachers should first, identify the objective they want to teach. Next, they should identify all the topographies those targets could be presented. For example, researchers identified numerator as a word, picture, and definition in Experiment 3. Next, teachers could rotate the way the question is asked between listener and speaker topographies. For example, this may look like rotating between multiple choice and production questions where the first question is to select an example of a numerator in a field of three/four and the following question asks the students to produce their own example of a numerator. It is important to keep in mind that there are many ways to arrange for multiple exemplar instruction and it should be individualized to individual students. Researchers should continue to induce verbal behavior development repertoires so students can benefit from more types of instruction and learn at a faster rate.

5.4 Limitations and Future Research

These experiments had various limitations that are discussed below as well as recommendations for future research. First, a major limitation is that these data were only correlational. There is a need for research, like this, to be conducted that implement experimental control. Researchers examined correlations between Inc-BiN and other variables but did not specifically conduct controlled experiments. Future research should attempt to induce Inc-BiN and measure the effects of those variables before and after Inc-BiN was induced. Once these experiments are conducted, researchers can begin to determine casualty or lack thereof. These correlations are important because it may serve as a necessary step in inspiring experimentally controlled procedures that look at these same variables. Future research should be conducted that induces Inc-BiN and measures academic performance, emergent intraverbals, and learn units/sets to criterion simultaneously.

Another relevant limitation is the participants' history of instruction and whether the outcomes reported here would generalize to individuals with different instructional histories. All participants were in a CABAS® AIL® classroom at the time the research was conducted. These classrooms implemented research-based behavior analytic systems and tactics that promoted individualized instruction. Although the CABAS® AIL® model has many benefits for the students, this may limit the generalizability of these data. Most participants had been in a CABAS® AIL® model for three or more years. This known history of instruction may limit the generalizability of these data to other populations. Future researchers may seek to replicate these experiments in more typical general education classrooms.

There are multiple other variables that may also be at play other than Inc-BiN. For example, observational learning is a critical verbal behavior development cusp that allows individuals to learning by observing the consequences of others (Fryling et al., 2011; Greer et al., 2006). The presence or absence of observational learning changes the relevant stimulus control during instruction and probe sessions for each participant. For example, individuals with observational learning may have acquired a different number of emergent relations than those without and we do not have these data. The presence or absence of observational learning could be the causal variable in these correlations and was not measured in the present experiments. . To determine if observational learning or Inc-BiN is causally related to the variables measured in this experiment, experiments would have to identify participants with similar verbal behavior and academic repertoires and systematically induce observational learning or Inc-BiN across multiple participants and measure emergent intraverbals and rate of learning before and after the given verbal behavior cusp has been induced. Then, researchers may be able to determine if observational learning or Inc-BiN causes increases in these other variables.

Other researchers may explain these data in terms of additional variables that we did not measure such as joint attention and socioeconomic status. Moore et al., (2014) defines joint attention as a "meeting of minds" where individuals attend to a shared context. Researchers may explain the data from these experiments first noting we did not measure joint attention and second explaining that individuals interest in attending to the same stimuli at the same time as their peers and teacher may be variable. Many developmental researchers may stress that joint attention may be the key variable in these relations that may be more directly related to emergent intraverbals and academic performance. Researchers may also stress that socioeconomic status may be the key variable at play here as well (Hart & Risley, 1980). Socioeconomic status has been correlated with many life outcome variables (i.e. health, education; Adler & Ostrove, 1999, Broer et al., 2019). Further research that implements controlled experiments is needed to compare these variables. Nevertheless, because of the significant correlations, measuring students' degree of Inc-BiN helps teachers determine if matrix instruction and equivalence-based instruction is appropriate for their students.

Based on data from Experiments 2 and 3, researchers call for more systematic research that measures listener responses, tacts, and intraverbals simultaneously following matrix instruction and equivalence-based instruction while also measuring Inc-BiN. When these experiments are conducted, this will help researchers examine the full range of outcomes involving verbal operants. Researchers also only tested the effects of two behavior analytic relational training interventions. Future research should replicate measuring these variables across a variety of contexts, such as investigating the relation between Inc-BiN and emergent intraverbals across multiple interventions. For example, researchers should measure multiple response topographies (i.e. hand-written, typed, vocal). Researchers could implement a relational training procedure in a specific topography and test if participants with varying degrees of Inc-BiN readily acquire other topographies of the directly taught response. For example, researchers could directly teach a response that is hand-written, and then test the accuracy of that response when typed or vocally emitted. If participants can or can't do this, it may have serious implications for how teachers implement instruction. Previous research has investigated the effects of mastery criterion and found that, in general, the higher the mastery criterion, the better the outcomes after intervention. However, based on the findings from the present experiment, individuals would likely benefit from a mastery criterion informed by their degree of Inc-BiN. For example, individuals with Inc-BiN may require a less stringent mastery criterion (e.g., few consecutive correct) than participants who demonstrate Pre-UniN. These data may help maximize instructional time and research by determining what each student needs to acquire a skill.

Conclusion

In three experiments, I investigated the relation between Incidental Bidirectional Naming (Inc-BiN) and academic performance. In Experiment 1, I investigated the relation between participants' degree of Inc-BiN and iReady Diagnostic assessment scores. Researchers found significant positive correlations between Inc-BiN (listener and speaker responses) and academic performance on iReady Diagnostic assessments for math and reading. In Experiments 2 and 3, I investigated the effects of relational training interventions (matrix instruction and equivalence-based instruction) on emergent intraverbals and learn units-to-criterion with participants with varying degree of Inc-BiN. Researchers found significant correlations between Inc-BiN and emergent intraverbal and learn units-to-criterion. Since a higher degree of Inc-BiN is significantly correlated with greater academic performance and fewer learn units-to-criterion, and a lower degree of Inc-BiN is significantly correlated with lower academic performance and a greater number of learn units-to-criterion, it seems logical for practitioners to measure student's degree of Inc-BiN to help modify and select appropriate instruction. Future research should expand on these findings by implementing more controlled experiments and studying Inc-BiN with multiple emergent verbal operants simultaneously.

References

Abdool-Ghany, F. A., & Fienup, D. F., (in press). A preliminary analysis of incidental bidirectional naming and derived listener and speaker relations for literacy responses. *Behavior and Social Issues.*

Abdool-Ghany, F. A. (2020). *Degrees of Bidirectional Naming Are Related to Derived Listener and Speaker Responses*. ProQuest Dissertations Publishing.

Adler, N. E., & Ostrove, J. M. (1999). Socioeconomic status and health: what we know and what we don't. *Annals of the New York academy of Sciences*, *896*(1), 3-15.

Albers, A. E., & Greer, R. D. (1991). Is the three-term contingency trial a predictor of effective instruction? *Journal of Behavioral Education*, 1, 337-354. https://doi.org/10.1007/BF00947188

Allen, M. (2019). Positive reinforcement 101. *Teachers Matter*, 44, 46–47.

Arafeh, S. (2016). Curriculum mapping in higher education: a case study and proposed content scope and sequence mapping tool. *Journal of Further and Higher Education*, *40*(5), 585–611. https://doi.org/10.1080/0309877X.2014.1000278

Ayllon, T., & Azrin, N. H. (1968). The token economy: A motivational system for therapy and rehabilitation. New York: Appleton-Century-Crofts.

Bahadourian, A. J., Tam, K. Y. (Brian), Greer, R. D., & Rousseau, M. K. (2006). The Effects of Learn Units on Student Performance in Two College Courses. *International Journal of Behavioral and Consultation Therapy*, *2*(2), 246–264. https://doi.org/10.1037/h0100780

Baldonado, L. A. (2021). *Increased Complexity in Bidirectional Naming Stimulus Control Enhances Reading in First Graders*. ProQuest Dissertations Publishing.

Barnes-Holmes, Y., Barnes-Holmes, D., & McHugh, L. (2004). Teaching derived relational responding to young children. *Journal of Early and Intensive Behavior Intervention*, 1, 3–12. https://doi.org/10.1037/h0100275

Barnes-Holmes, D., & Harte, C. (2022). Relational frame theory 20 years on: The Odysseus voyage and beyond. *Journal of the Experimental Analysis of Behavior*, *117*(2), 240-266.

Belmonte-Mulhall, C. P., & Harrison, J. R. (2023). Multi-tiered systems of support: a pilot study of teacher interpretation and application of graphed behavioral data. *Journal of Applied School Psychology*, 39(2), 151-178. https://doi.org/10.1080/15377903.2022.2113945

Bishop, D. V. M., Snowling, M. J., Thompson, P. A., & Greenhalgh, T. (2016). Catalise: A multinational and multidisciplinary delphi consensus study. Identifying language impairments in children. *PloS One*, *11*(7), e0158753–e0158753. https://doi.org/10.1371/journal.pone.0158753

Blackledge, J. T. (2003). An introduction to relational frame theory: basics and applications. *The Behavior Analyst Today*, *3*(4), 421–433. https://doi.org/10.1037/h0099997

Brodsky, J., & Fienup, D. M. (2018). Sidman goes to college: A meta-analysis of equivalence-based instruction. *Perspectives on Behavior Science*, 41, 95–119. doi:10.1007/s40614-018-0150-0

Broer, M., Bai, Y., Fonseca, F., Broer, M., Bai, Y., & Fonseca, F. (2019). A review of the literature on socioeconomic status and educational achievement. *Socioeconomic inequality and educational outcomes: Evidence from twenty years of TIMSS*, 7-17

Cahill, C. S., & Greer, R. D. (2014). Actions vs. words: How we can learn both. *Acta Investigacion Psichologia*, *4*, 1716–1745.

Carp, C.L. and Petursdottir, A.I. (2015), Intraverbal naming and equivalence class formation in children. Journal of the Experimental Analysis of Behavior, 104: 223-240. https://doi.org/10.1002/jeab.183

Cassidy, S., Roche, B., Colbert, D., Stewart, I., & Grey, I. M. (2016). A relational frame skills training intervention to increase general intelligence and scholastic aptitude. *Learning and Individual Differences*, 47, 222-235. https://doi.org/10.1016/j.lindif.2016.03.001

Catania, A. C. (2013). A natural science of behavior. *Review of General Psychology*, 17(2), 133-139. https://doi.org/10.1037/a0033026

Cengher, M., Kim, J. Y., & Fienup, D. M. (2020). A Review of Prompt-Fading Procedures, Part II: Parameters and Components. *Journal of Developmental and Physical Disabilities*, *32*(4), 553–574. https://doi.org/10.1007/s10882-019-09711-2

Choi, J., Greer, R. D., & Keohane, D.-D. (2015). The effects of an auditory match-to-sample procedure on listener literacy and echoic responses. *Behavioral Development Bulletin (Philadelphia, Pa.)*, *20*(2), 186–206. https://doi.org/10.1037/h0101313

Chomsky, N. (1959). A review of B. F. Skinner's verbal behavior. language, 35(1), 26–58. https://doi.org/10.2307/411334

Conine, D. E., Frampton, S. E., Buote, K. A., & Keller, C. E. (2023). A scoping review of empirical research on emergent intraverbal behavior. *Behavioral Interventions*, e1986. https://doi.org/10.1002/bin.1986

Cooper, J. O., Heron, T. E., & Heward, W. L. (2007). Applied behavior analysis.

Cooper, J. O., Heron, T. E., & Heward, W. L. (2019). Applied behavior analysis (Third edition.). Pearson.Greer, R. D., & Ross, D. E. (2008). Verbal behavior analysis: inducing and expanding new verbal capabilities in children with language delays. *Pearson/Allyn & Bacon*.

Corwin, A. (2011). *A Functional Analysis of the Effects of the Induction of Naming and Observing Teacher-Modeling on Accelerated Learning of Academic Skills for Children with Autism*. ProQuest Dissertations Publishing.

Crawford, E. R., LePine, J. A., & Rich, B. L. (2010). Linking job demands and resources to employee engagement and burnout: a theoretical extension and meta-analytic test. Journal of applied psychology, 95(5), 834. https://doi.org/10.1037/a0019364

Critchfield, T. S. (2014). Online equivalence-based instruction about statistical inference using written explanation instead of match-to-sample training. *Journal of Applied Behavior Analysis*, 47, 606–611. https://doi.org/10.1002/jaba.150

Critchfield, T. S., & Twyman, J. S. (2014). Prospective instructional design: establishing conditions for emergent learning. *Journal of Cognitive Education and Psychology*, *13*(2), 201-217.

Csillag, J. (2016). Differentiated reading instruction : strategies and technology tools to help all students improve. Routledge. https://doi.org/10.4324/9781315630649

Curiel, E. S. L., Axe, J. B., Sainato, D. M., & Goldstein, H. (2020). Systematic Review of Matrix Training for Individuals With Autism Spectrum Disorder. *Focus on Autism and Other Developmental Disabilities*, *35*(1), 55–64. https://doi.org/10.1177/1088357619881216

Daar, J. H., Negrelli, S., & Dixon, M. R. (2015). Derived emergence of wh-question–answers in children with autism. *Research in Autism Spectrum Disorders*, 19, 59–71. doi:10.1016/j.rasd.2015.06.004

Decasper, A., & Spence, M. (1987). Prenatal experience with low-frequency maternal-voice sounds influence neonatal perception of maternal voice samples. *Infant Behavior and Development 10, 2*. 133-142.

DeLuca, C., & Bellara, A. (2013). The current state of assessment education: aligning policy, standards, and teacher education curriculum. *Journal of Teacher Education*, *64*(4), 356–372. https://doi.org/10.1177/0022487113488144

Dinsmoor, J. A. (1983). Observing and conditioned reinforcement. *Behavioral and Brain Sciences*, *6*, 693–728. https://doi.org/10.1017/S0140525X00017969

Dixon, M. R., Hayes, S. C., Stanley, C., Law, S., & al-Nasser, T. (2020). Is Acceptance and Commitment Training or Therapy (ACT) a Method that Applied Behavior Analysts Can and Should Use? *The Psychological Record*, *70*(4), 559–579. https://doi.org/10.1007/s40732-020-00436-9

Dixon, M. R., Belisle, J., & Stanley, C. R. (2018). Derived relational responding and intelligence: assessing the relationship between the PEAK-E pre-assessment and IQ with individuals with autism and related disabilities. *The Psychological Record*, *68*(4), 419–430. https://doi.org/10.1007/s40732-018-0284-1

Donley, C.R., & Greer, R. D. (1993). Setting events controlling social verbal exchanges between students with developmental delays. *Journal of Behavioral Education, 3*, 387-401.

Du, L., Broto, J., & Greer, R.. (2015). The effects of establishment of conditioned reinforcement for observing responses for 3D stimuli on generalized visual match-to-sample in children with autism spectrum disorders. *European Journal of Behavior Analysis*. 16. 82-98.

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving Students' Learning With Effective Learning Techniques: Promising

Directions From Cognitive and Educational Psychology. *Psychological Science in the Public Interest*, *14*(1), 4–58. https://doi.org/10.1177/1529100612453266

Eby, C. M., & Greer, R. D. (2017). Effects of social reinforcement on the emission of tacts by preschoolers. *Behavioral Development Bulletin*, 22(1), 23–43. https://doi.org/10.1037/bdb0000043

Engelmann, S., & Carnine, D. (1991). Theory of instruction: principles and applications. Eugene, Oregon: ADI Press.

Fienup, D. M., Covey, D. P., & Critchfield, T. S. (2010). Teaching brain-behavior relations economically with stimulus equivalence technology. *Journal of Applied Behavior Analysis*, *43*(1), 19-. https://doi.org/10.1901/jaba.2010.43-19

Fienup, D.M. The future of verbal behavior: integration. *Analysis Verbal Behavior* 34, 18–23 (2018). https://doi.org/10.1007/s40616-018-0108-z

Fiorile, C.A., Greer, R.D. The induction of naming in children with no prior tact responses as a function of multiple exemplar histories of instruction. *Analysis Verbal Behavior* 23, 71–87 (2007). https://doi.org/10.1007/BF03393048

Foody, M., Barnes-Holmes, Y., Barnes-Holmes, D., Törneke, N., Luciano, C., Stewart, I., & McEnteggart, C. (2014). RFT for clinical use: The example of metaphor. J*ournal of Contextual Behavioral Science,* 3(4), 305-313. https://doi.org/10.1016/j.jcbs.2014.08.001

Frampton, S.E., Wymer, S.C., Hansen, B. and Shillingsburg, M.A. (2016), The use of matrix training to promote generative language with children with autism. *Journal of Applied Behavior Analysis*, 49: 869-883. https://doi.org/10.1002/jaba.340

Friedman, L. F. (2020). *Establishment of Increased Stimulus Control for Bidirectional Naming Increased Stimulus Control for Other Derived Relations in 20- to 40-Month-Old Toddlers*. ProQuest Dissertations Publishing.

Fryling, M. J., Rehfeldt, R. A., Tarbox, J., & Hayes, L. J. (2020). An introduction to applied behavior analysis of language and cognition. *Applied Behavior Analysis of Language and Cognition*. Context Press.

Gibbs, A. R., & Tullis, C. A. (2021). The emergence of untrained relations in individuals with autism and other intellectual and developmental disabilities: A systematic review of the recent literature. Review *Journal of Autism and Developmental Disorders*, 8(2), 213-238. https://doi.org/10.1007/s40489-020-00211-0

Gibbs, A. R., Tullis, C. A., Thomas, R., & Elkins, B. (2018). The effects of noncontingent music and response interruption and redirection on vocal stereotypy. *Journal of Applied Behavior Analysis*, 51(4), 899-914. https://doi.org/10.1002/jaba.485

Gilmore, A., Barnes-Holmes, D. & Sivaraman, M. A (2024). Modern Collaborative Behavior Analytic Approach to Incidental Naming. *Perspect Behavioral Science*. https://doi.org/10.1007/s40614-024-00399-0

Gilic, L., & Greer, R. D. (2009). Establishing Naming in typically developing two-year-old children as a function of multiple exemplar speaker and listener experiences. *The Analysis of Verbal Behavior*, 27(1), 157–177. https://doi.org/10.1007/BF03393099

Greenberg, J.H., Tsang, W. and Yip, T. (2014) The effects of intensive tact instruction with young children having speech delays on pure tacts and mands in non-instructional settings: A partial replication. *Behavioral Development Bulletin*, 19, 35-39. http://dx.doi.org/10.1037/h0100572

Greer, R. D. (1991). Teaching practices to save America's schools: The legacy of B. F. Skinner. *Journal of Behavioral Education*, 1, 159-164. https://doi.org/10.1007/BF00957002

Greer, R. D. (1994b). The measure of a teacher. In R. Gardner HI, D. M. Sainato, J. O. Cooper, T. E. Heron, W. L. Heward, J. W. Eshleman, & T. A. Grossi (Eds.), Behavior analysis in education: Focus on measurably superior instruction (pp. 161-171). *Pacific Groves, CA: Brooks/Cole*.

Greer, R. D. (1997a). The crisis in education: Contributing factors and solutions. In B. Thyer & R. Mattaini, (Eds.). Problems in society and solutions from the science of behavior. Washington, D. C.: American Psychological Association.

Greer, R. D. (1997b). Acting to save our schools (1984-1994). In J. Cautela & W. Ishaq (Eds.). The science of behavior and the human condition. *New York: Praeger*

Greer, R. D. (2002). Designing teaching strategies: An applied behavior analysis systems approach. *New York: Academic*.

Greer, R. D. (2008). The ontogenetic selection of verbal capabilities: Contributions of Skinner's verbal behavior theory to a more comprehensive understanding of language. *International Journal of Psychology and Psychological Therapy*, *8*(3), 363-386.

Greer, R. D., Chavez-Brown, M., Nirgudkar, A. S., Stolfi, L., & Rivera-Valdes, C. (2005). Acquisition of fluent listener responses and the educational advancement of young children with autism and severe language delays. *European Journal of Behavior Analysis*, *6*(2), 125.

Greer, R. D., Corwin, A., & Buttigieg (2011). The effects of the verbal developmental capability of Naming on how children can be taught. *Acta de Investigacion Psicologia*, 1(1), 23–54. Retrieved from http://scielo.unam.mx/pdf/aip/v1n1/v1n1a5.pdf

Greer, R. D., & Du, L. (2010). Generic instruction versus intensive tact instruction and the emission of spontaneous speech. *The Journal of Speech-Language Pathology and Applied Behavior Analysis, 5(*1), 1–19.

Greer, R. D., & Du, L. (2015a). Experience and the onset of the capability to learn names incidentally by exclusion. *The Psychological Record, 65,* 355–373

Greer, R. D., & Du, L. (2015b). Identification and establishment of reinforcers that make the development of complex social language possible. *International Journal of Behavior Analysis and Autism Spectrum Disorders*, *1*(1), 13–34.

Greer, R., Dudek-Singer, J., & Gautreaux, G. (2006). Observational learning. *International Journal of Psychology, 41*(6), 486–499. https://doi.org/10.1080/00207590500492435

Greer, R. D., & Keohane, D.-D. (2006). The Evolution of Verbal Behavior in Children. *The Journal of Speech and Language Pathology, Applied Behavior Analysis*, *1*(2), 111–140. https://doi.org/10.1037/h0100194

Greer, R. D., & Longano, J. (2010). A rose by naming: How we may learn how to do it. *The Analysis of Verbal Behavior*, 26, 73-106. https://doi.org/10.1007/BF03393085

Greer, R.D., Pohl, P., Du, L. and Moschella, J.L. (2017) The separate development of children's listener and speaker behavior and the intercept as behavioral metamorphosis. *Journal of Behavioral and Brain Science,* 7, 674- 704. https://doi.org/10.4236/jbbs.2017.713045

Greer, R. D., & Ross, D. E. (2008). Verbal behavior analysis: Inducing and expanding complex communication in children with severe language delays. Boston, MA: *Allyn & Bacon*.

Greer, R. D., & Speckman, J. (2009). The integration of speaker and listener responses: A theory of verbal development. *The Psychological Record*, *59*(3), 449.

Greer, R. D., Stolfi, L., Chavez-Brown, M., & Rivera-Valdez, C. (2005). The emergence of the listener to speaker component of Naming in children as a function of multiple exemplar instruction. *The Analysis of Verbal Behavior*, 21, 123–134. https://doi.org/10.1007/BF03393014

Gunersel, A. B., Mason, B. A., Wills, H. P., Caldarella, P., Williams, L., & Henley, V. M. (2023). Effective Classroom Management in Middle Level Schools: A Qualitative Study of Teacher Perceptions. *RMLE Online : Research in Middle Level Education*, *46*(8), 1–13. https://doi.org/10.1080/19404476.2023.2252714

Hart B. M., & Risely, T. R. (1995). Meaningful differences in the everyday life of Americas children. Baltimore: Paul Brookes.

Hawkins, E., Gautreaux, G., & Chiesa, M. (2018). Deconstructing common bidirectional naming: A proposed classification framework. *Analysis of Verbal Behavior*, 34, 44–61.

Hayes, S. C., Barnes-Holmes, D., & Roche, B. (Eds.). (2001). Relational frame theory: A post-Skinnerian account of human language and cognition. *New York, NY: Plenum Press*.

Hayes, S. C., Law, S., Assemi, K., Falletta-Cowden, N., Shamblin, M., Burleigh, K., ... & Smith, P. (2021). Relating is an operant: A fly over of 35 years of RFT research. *Perspectivas em Análise do Comportamento*, 12(1), 005-032. https://doi.org/10.18761/PAC.2021.v12.RFT.02

Hayes, S. C., Strosahl K. D., Bunting K., Twohig M. P., & Wilson K. G. (2004). What is acceptance and commitment therapy? In Hayes S. C., Strosahl K. D. (Eds.), A practical guide to acceptance and commitment therapy (pp. 1–30). New York, NY: Guilford Press.

Ho, H., Perry, A., & Koudys, J. (2021). A systematic review of behaviour analytic interventions for young children with intellectual disabilities. *Journal of Intellectual Disability Research*, *65*(1), 11–31. https://doi.org/10.1111/jir.12780

Holth, P. (2017). Multiple exemplar training: some strengths and limitations. *Journal of Behavior Analysis.* 40, 225–241. https://doi.org/10.1007/s40614-017-0083-z

Horne, P. J., & Lowe, C. F. (1996). On the origins of naming and other symbolic behavior. *Journal of the Experimental Analysis of Behavior*, 65, 185–241. https://doi.org/10.1901/jeab.1996.65-185

Horne, P.J. and Lowe, C.F. (1997). Toward a theory of verbal behavior. *Journal of the Experimental Analysis of Behavior*, 68: 271-296. https://doi.org/10.1901/jeab.1997.68-271

Hotchkiss, R. M., & Fienup, D. M. (2020). A parametric analysis of a protocol to induce bidirectional naming: effects of protocol intensity. *The Psychological Record*, *70*(3), 481–497. https://doi.org/10.1007/s40732-020-00383-5

Hranchuk, K., Douglas Greer, R. & Longano, J. (2019). Instructional demonstrations are more efficient than consequences alone for children with Naming. *Analysis Verbal Behavior*, 35, 1–20. https://doi.org/10.1007/s40616-018-0095-0

Ingham, P., & Greer, R. D. (1992). Changes in student and teacher responses in observed and generalized settings as function of supervisor observations of teachers. *Journal of Applied Behavior Analysis*, 25, 153-164. https://doi.org/10.1901/jaba.1992.25-153

Irie, T., Yokomitsu, K., & Sakano, Y. (2019). Relationship between cognitive behavioral variables and mental health status among university students: A meta-analysis. *PloS One*, *14*(9), e0223310–e0223310. https://doi.org/10.1371/journal.pone.0223310

Jenkins, H. M., & Harrison, R. H. (1960). Effect of discrimination training on auditory generalization. *Journal of Experimental Psychology*, 59(4), 246–253. https://doi.org/10.1037/h0041661

Jennings, A.M. and Miguel, C.F. (2017), Training intraverbal bidirectional naming to establish generalized equivalence class performances. *Journal of Experimental Analysis Behavior*, 108: 269-289. https://doi.org/10.1002/jeab.277

Kemmerer, A.R., Vladescu, J.C., Carrow, J.N., Sidener, T.M. and Deshais, M.A. (2021), A systematic review of the matrix training literature. *Behavioral Interventions*, 36: 473-495. https://doi.org/10.1002/bin.1780

Kim, J. Y., Fienup, D. M., Oh, A. E., & Wang, Y. (2022). Systematic review and meta-analysis of token economy practices in K-5 educational settings, 2000 to 2019. *Behavior Modification*, *46*(6), 1460–1487. https://doi.org/10.1177/01454455211058077

Kohler, K. T., & Malott, R. W. (2014). Matrix training and verbal generativity in children with autism. *The Analysis of Verbal Behavior*, *30*, 170–177.

LaFrance, D.L. and Tarbox, J. (2020), The importance of multiple exemplar instruction in the establishment of novel verbal behavior. *Journal of Applied Behavioral Analysis*, 53: 10-24. https://doi.org/10.1002/jaba.611

Lamm, N., & Greer, R. D. (1991). A systematic replication of CABAS® in Italy. *Journal of Behavioral Education*, 1, 427-444. https://doi.org/10.1007/BF00946776

Lee Park, H.-S. (2005). *Multiple exemplar instruction and transformation of stimulus function from auditory-visual matching to visual-visual matching*. ProQuest Dissertations Publishing.

Lodhi, S. & Greer, R. D. (1989). The speaker as listener. *Journal of the Experimental Analysis of Behavior, 51*, 353-359

Longano, J. M., & Greer, R. D. (2006). The effects of a stimulus-stimulus pairing procedure on the acquisition of conditioned reinforcement on observing and manipulating stimuli by young children with autism. *Journal of Early and Intensive Behavior Intervention, 3*(1), 62-80. http://dx.doi.org/10.1037/h0100323

Lowe, C.F., Horne, P.J., Harris, F.D.A. and Randle, V.R.L. (2002). Naming and categorization in young children: vocal tact training. *Journal of the Experimental Analysis of Behavior*, 78: 527-549. https://doi.org/10.1901/jeab.2002.78-527

Lynch, D. C., & Cuvo, A. J. (1995). Stimulus equivalence instruction of fraction-decimal relations. *Journal of Applied Behavior Analysis*, 28, 115–126. https://doi.org/10.1901/jaba.1995.28-115

Maffei, J., Singer-Dudek, S. and Keohane, D. (2014) The effects of the establishment of adult faces and/or voices as conditioned reinforcers for children with ASD and related disorders. *Acta de Investigacion Psicologia*, 4, 1621-1641. https://doi.org/10.1016/S2007-4719(14)70970-6

Marzullo-Kerth, D., Reeve, S. A., Reeve, K. F., & Townsend, D. B. (2011). Using multiple-exemplar training to teach a generalized repertoire of sharing to children with autism. *Journal of Applied Behavior Analysis, 44*(2), 279-. https://doi.org/10.1901/jaba.2011.44-279

Maurilus, E. N. (2018). *The Effect of the Establishment of Reinforcement Value for Math on Rate of Learning for Pre-Kindergarten Students*. ProQuest Dissertations Publishing.

McClay, C.-A., Waters, L., McHale, C., Schmidt, U., & Williams, C. (2013). Online cognitive behavioral therapy for bulimic type disorders, delivered in the community by a nonclinician: qualitative study. *Journal of Medical Internet Research*, *15*(3), e46–e46. https://doi.org/10.2196/jmir.2083

Menzies, H. M., Lane, K. L., Oakes, W. P., & Ennis, R. P. (2017). Increasing students' opportunities to respond: a strategy for supporting engagement. *Intervention in School and Clinic, 52*(4), 204–209. https://doi.org/10.1177/1053451216659467

Michael, J. (1982). Distinguishing between discriminative and motivational functions of stimuli. *Journal of Experimental Analysis of Behavior*, 37, 149–155. https://doi.org/10.1901/jeab.1982.37-149

Moore, C., Dunham, P. J., & Dunham, P. (2014). *Joint attention: Its origins and role in development*. Psychology Press.

Morgan, G. A., Greer, R. D., & Fienup, D. M. (2021). Descriptive analyses of relations among bidirectional naming, arbitrary, and nonarbitrary relations. *The Psychological Record*, *71*(3), 367–387. https://doi.org/10.1007/s40732-020-00408-z

Mueller, M. M., Olmi, D. J., & Saunders, K. J. (2000). Recombinative generalization of within-syllable units in prereading children. *Journal of Applied Behavior Analysis*, *33*(4), 515–531. https://doi.org/10.1901/jaba.2000.33-515

Naoi, N., Yokoyama, K., & Yamamoto, J. (2006). Matrix training for expressive and receptive two-word utterances in children with autism. *Japanese Journal of Special Education*, *43*, 505–518.

Nuzzolo-Gomez, R. & Greer, R. D. (2004). Emergence of untaught mands and tacts with novel adjective-object pairs as a function of instructional history. *The Analysis of Verbal Behavior, 24,* 30-47.

Oliver, B., Ferns, S., Whelan, B., & Lilly, L. (2010). Mapping the curriculum for quality enhancement: Refining a tool and processes for the purpose of curriculum renewal. Proceedings of AuQF 2010. Quality in Uncertain Times., 80-88.

Page, S. V., Zimmerman, D. M., & Pinkelman, S. E. (2023). A systematic review of dependent group contingencies (1970–2019). *Journal of Positive Behavior Interventions*, *25*(3), 198–209. https://doi.org/10.1177/10983007211054519

Palmer, D. C. (2016). On intraverbal control and the definition of the intraverbal. *The Analysis of Verbal Behavior*, 32(2), 96–106. https://doi.org/10.1007/s40616-016-0061-7

Pauwels, A. A., Ahearn, W. H., & Cohen, S. J. (2015). Recombinative generalization of tacts through matrix training with individuals with autism spectrum disorder. *The Analysis of Verbal Behavior, 31*, 200–214.

Peiris, D. L. I. H. K., Duan, Y., Vandelanotte, C., Liang, W., Yang, M., & Baker, J. S. (2022). Effects of in-classroom physical activity breaks on children’s academic performance, cognition, health behaviours and health outcomes: a systematic review and meta-analysis of randomized controlled trials. *International Journal of Environmental Research and Public Health, 19*(15), 9479-. https://doi.org/10.3390/ijerph19159479

Perlin, M. S. (2011). Curriculum mapping for program evaluation and CAHME accreditation. *Journal of Health Administration Education*, 28(1), 33-53.

Pohl, P., Greer, R. D., Du, L., & Moschella, J. L. (2018). Verbal development, behavioral metamorphosis, and the evolution of language. *Perspectives on Behavior Science*, 1–18. https://doi.org/10.1007/s40614-018-00180-0

Richling, S. M., Williams, W. L., & Carr, J. E. (2019). The effects of different mastery criteria on the skill maintenance of children with developmental disabilities. *Journal of Applied Behavior Analysis*, 52(3), 701-717. https://doi.org/10.1002/jaba.580

Roediger, H. L., Finn, B., & Weinstein, Y. (2012). Applications of cognitive science to education. Neuroscience in education: The good, the bad, and the ugly, 128-151.

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. https://doi.org/10.1901/jaba.1997.30-533

Rosenberg, D., & Anderson, T. (2021). Teacher turnover before, during, & after COVID. In *Education Resource Strategies*. Education Resource Strategies.

Ross, D. & Greer, R. D., (2003). Generalized imitation and the mand: inducing first instances of speech in young children with autism. *Research in Developmental Disabilities*,*24* (58-74).

Ross, D. E., Nuzzolo, R., Stolfi, L., & Natarelli, S. (2006). Effects of Speaker Immersion on Independent Speaker Behavior of Preschool Children with Verbal Delays. *Journal of Early and Intensive Behavior Intervention*, *3*(1), 135–150. https://doi.org/10.1037/h0100327

Schmelzkopf, J., Greer, R. D., Singer-Dudek, J., & Du, L. (2017). Experiences that establish preschoolers' interest in speaking and listening to others. *Behavioral Development Bulletin*, *22*, 44.

Selinski, J., Greer, R. D., & Lodhi, S. (1991). A functional analysis of the comprehensive application of behavior analysis to schooling. *Journal of Applied Behavior Analysis*, 24, 108 - 118 . https://doi.org/10.1901/jaba.1991.24-107

Sidman, M. (1994). Equivalence relations and behavior: A research story. Boston, MA: Authors Cooperative.

Sidman, M. (2000). Equivalence relations and the reinforcement contingency. *Journal of the Experimental Analysis of Behavior*, 74: 127-146. https://doi.org/10.1901/jeab.2000.74-127

Singer-Dudek, J., Oblak, M., & Greer, R. D. (2011). Establishing books as conditioned reinforcers for preschool children as a function of an observational intervention. *Journal of Applied Behavior Analysis*, *44*(3), 421-. https://doi.org/10.1901/jaba.2011.44-421

Singer-Dudek, J., Speckman, J., & Nuzzolo, R. (2010). A comparative analysis of the CABAS® model of education at the Fred S. Keller School: A twenty-year review. *The Behavior Analyst Today*, *11*(4), 253–265. https://doi.org/10.1037/h0100705

Sivaraman, M., Barnes-Holmes, D., Greer, R. D., Fienup, D. M., & Roeyers, H. (2023). Verbal behavior development theory and relational frame theory: Reflecting on similarities and differences. *Journal of the Experimental Analysis of Behavior*, 119(3), 539–553. https://doi.org/10.1002/jeab.836

Sivaraman, M, Bhabu, P. (2018). Establishment of exclusion responding in children with autism spectrum disorder. *Behavioral Interventions*. 414– 426.

Skinner, B. F. (1957). Verbal behavior. New York, NY: Appleton-Century-Crofts.

Speckman-Collins, J., Lee Park, H. and Greer, R.D. (2007) Generalized selection-based auditory matching and the emergence of the listener component of naming. *Journal of Early and Intensive Behavioral Intervention*, 4, 412-429. https://doi.org/10.1037/h0100382

Stokes, T. F., & Baer, D. M. (1977). An implicit technology of generalization. *Journal of Applied Behavior Analysis*, 10, 349–367. https://doi.org/10.1901/jaba.1977.10-349

Sun, J., Przybylski, R., & Johnson, B. J. (2016). A review of research on teachers' use of student data: from the perspective of school leadership. *Educational Assessment, Evaluation and Accountability*, *28*(1), 5–33. https://doi.org/10.1007/s11092-016-9238-9

Sundberg, M. L., & Sundberg, C. A. (2011). Intraverbal behavior and verbal conditional discriminations in typically developing children and children with autism. *The Analysis of Verbal Behavior*, 27(1), 23–44. https://doi.org/10.1007/BF03393090

Tsai, H., & Greer, R. D. (2006). Conditioned observation of books and accelerated acquisition of textual responding by preschool children. *Journal of Early and Intensive Behavior Intervention, 3*(1), 35-61. http://dx.doi.org/10.1037/h0100322

Tsiouri, I., Greer, R.D. (2003) Inducing vocal verbal behavior in children with severe language delays through rapid motor imitation responding. *Journal of Behavioral Education* 12, 185–206. https://doi.org/10.1023/A:1025508311022

Twyman, J.S. (1996). The functional independence of impure mands and tacts of abstract stimulus properties. *Analysis Verbal Behav* 13, 1–19.

Unholz-Bowden, E., McComas, J. J., McMaster, K. L., Girtler, S. N., Kolb, R. L., & Shipchandler, A. (2020). Caregiver training via telehealth on behavioral procedures: a systematic review. *Journal of Behavioral Education, 29*(2), 246–281. https://doi.org/10.1007/s10864-020-09381-7

Van den Bergh, L., Ros, A., & Beijaard, D. (2013). Teacher feedback during active learning: Current practices in primary schools. *British Journal of Educational Psychology*, *83*(2), 341–362. https://doi.org/10.1111/j.2044-8279.2012.02073.x

Vargas, J. S. (2013). Behavior analysis for effective teaching. Routledge.

Verdun, V.R., Fienup, D.M., Chiasson, B.A. and Greer, R.D. (2022), Arranging peer-tutoring instruction to promote inference-making. *Journal of Applied Behavior Analysis*, 55: 369-394. https://doi.org/10.1002/jaba.895

Wasik, B. A., & Hindman, A. H. (2011). Improving vocabulary and pre-literacy skills of at-risk preschoolers through teacher professional development. *Journal of Educational Psychology*, 103(2), 455–469. https://doi.org/10.1037/a0023067

Wong, K. K., Fienup, D. M., Richling, S. M., Keen, A., & Mackay, K. (2022). Systematic review of acquisition mastery criteria and statistical analysis of associations with response maintenance and generalization. *Behavioral Interventions*, *37*(4), 993–1012. https://doi.org/10.1002/bin.1885

Yoon, S. Y., & Bennett, G. M. (2000). Effects of a stimulus-pairing procedure on conditioning vocal sounds as reinforcers. *The Analysis of Verbal Behavior*, 17(1), 75–88. https://doi.org/10.1007/BF03392957

Zodik, I., & Zaslavsky, O. (2008). Characteristics of teachers' choice of examples in and for the mathematics classroom. *Educational Studies in Mathematics*, *69*(2), 165–182. https://doi.org/10.1007/s10649-008-914