5-2019

Discrimination Training to Produce Emergent Relations of Pre-Algebraic Math Skills

Jeanine Rinda Tanz

Discrimination Training to Produce Emergent Relations
of Pre-Algebraic Math Skills

by

Jeanine Rinda Tanz

Bachelor of Arts
Psychology
Florida Institute of Technology
2011

Master of Science
Applied Behavior Analysis and Organizational Behavior Management
Florida Institute of Technology
2014

A dissertation submitted to the School of Behavior Analysis at
Florida Institute of Technology
in partial fulfillment of the requirements
for the degree of

Doctor of Philosophy
in
Behavior Analysis

Melbourne, Florida
May, 2019

The undersigned committee hereby recommends that the attached document be accepted as fulfilling in part the requirements for the degree of Doctor of Philosophy in Behavior Analysis. “Discrimination Training to Produce Emergent Relations of Pre-Algebraic Math Skills,” a dissertation by Jeanine Rinda Tanz.

____

A. Celeste Harvey, Ph.D., BCBA-D
Associate Professor, School of Behavior Analysis
Major Advisor and Committee Chair

____

Catherine Nicholson, Ph.D., BCBA-D
Assistant Professor, School of Behavior Analysis
Committee Member

____

Ivy M. Chong, Ph.D., BCBA-D
Senior Vice President, May Institute
Committee Member

____

Bryon Neff, Ph.D., BCBA-D
Assistant Professor, School of Behavior Analysis
Committee Member

____

Eraldo Ribeiro, Ph.D.
Associate Professor, Computer Engineering and Sciences
Outside Committee Member

____

Lisa Steelman, Ph.D.
Professor and Dean, College of Psychology and Liberal Arts

Abstract

Title: Discrimination Training to Produce Emergent Relations of Pre-Algebraic Math Skills

Author: Jeanine Rinda Tanz

Principle Advisor: A. Celeste Harvey, Ph.D., BCBA-D

Individuals diagnosed with Autism Spectrum Disorder (ASD) often have deficits with respect to generalization of skills. Procedures designed to induce stimulus equivalence have been shown to promote generalization. This study used a modified multiple probe design with an embedded multiple schedule to: (a) compare one-to-many (OTM) and many-to-one (MTO) training structures to determine which structure results in more positive equivalence outcomes when all variables are held constant, and (b) determine the extent to which children with ASD demonstrate stimulus equivalence and stimulus class mergers when using educationally relevant stimuli. Four children with ASD were taught two classes of stimuli (Class 1 and Class 2) comprised of pre-algebraic math skills across two different training structures. Results indicate that the OTM and MTO training structures are equally as effective at producing positive equivalence outcomes for individuals with ASD. Additionally, no participants in the current study demonstrated a class merger. Implications for teaching educationally relevant materials to children with ASD are discussed.

Table of Contents

Table of Contents ……………………………………………………………………iv

List of Keywords …………………………………………………………………...vii

List of Figures ……………..……………………………………………………...viii

List of Tables ………………..……………………………………………...…...ix

Acknowledgement ……………………………………………………………....…x

Dedication ……………………………………………………………………...….xi

Chapter 1: Introduction …………………………………………...………………..1

Conditional Discriminations …………………………………………….....2

Teaching Techniques ……………………………………………………3

Simple-conditional training ………………………………4

Conditional-only training …………………………………5

Errorless Learning …………………………………………………….6

Stimulus Equivalence ……………………………………………………………8

Variables Enhancing Stimulus Equivalence ……………………...12

Class Mergers ……………………………………………………..22

Purpose of Current Study ……………………………………..……..…25

Chapter 2: Method ………………………………………………………………...27

Participants, Settings, Materials ……………………………………..27

Response Definitions and Measurement …………………….................30

Dependent variable ……………………………………………….30

Experimental Design and Procedures ……………………………….……31

Experimental design ………………………………………………31

Preference assessment ……………………………………………32

General procedures ……………………………………….………33

Randomization of trials ……………………………………..…..35

Testing ……………………………………………………….……36

Baseline …………………………………………….………..38

Training ……………………………………………………...39

Session termination criteria ………………………………………41

Social Validity, Interobserver Agreement, and Treatment Integrity……...42

Social validity ……………………………………………………..42

Interobserver agreement …………………………………………..42

Treatment integrity ………………………………………………..43

Chapter 3: Results ……………………………………………………...44

Izak ………………………………………………………………..44

Alex ……...………………………………………………………..48

Greg ……….……………………………………………………..52

Luke ……………………………………………………………….54

Summary of Results ………………...…………………..………58

Chapter 4: Discussion ……………......……………………….…………………60

Limitations and Future Directions ……………………………………………65

Figures……………………………………………………………..………………………….68

Tables…………………………………………………………………………………...….75

References …………………………………………………………………..………..78

Appendix ……………………………………………………………………………92

List of Keywords

conditional discrimination

stimulus equivalence

training structures

one-to-many (OTM)

many-to-one (MTO)

class merger

derived relations

meaningful stimuli

autism

equivalence-based instruction

List of Figures

Figure 1. OTM Training Structure ……………….……….……………………68

Figure 2. MTO Training Structure ………………….……….……………………69

Figure 3. Izak Results …………………………..…….……………………70

Figure 4. Alex Results ……………………………..……………..…………..71

Figure 5. Greg Results ……………………..……………………..…………..72

Figure 6. Luke Results……………………………..……………….……………..73

Figure 7. Combined Results ………………………….…..………………..74

List of Tables

Table 1. Stimulus Sets Used in Testing and Training ……………….……………75

Table 2. Order of Testing and Training ……………………………………………..76

Table 3. Interobserver Agreement (IOA) Results …………………………………77

Table 4. Treatment Integrity (TI) Results …………………………………………77

Acknowledgement

The completion of this dissertation project became a reality with the unwaivering support and help of many individuals. I would like to extend my deepest gratitude to all of them. First, I would like to thank my academic advisors, Dr. A. Celeste Harvey and Dr. Catherine Nicholson. Their endless encouragement and guidance helped me cross the finish line. This project would not have been possible without the support of Dr. Ivy M. Chong. I would like to thank her for believing in me and for providing me with many opportunities to learn and grow both personally and professionally. I would also like to thank Dr. Bryon Neff, Dr. Erlado Ribeiro, Dr. Lauren Beaulieu, Dr. Mark Harvey, and Dr. Michael E. Kelley for graciously serving on my committee and for providing invaluable feedback and advice throughout this project. I send a special thank you to The Scott Center for Autism Treatment and Florida Autism Center for allowing me to recruit participants for this study and for allowing me to use their facilities to run research sessions. Thank you to the many people who assisted with data collection including: Madeleine Keevy, Krystin Hussain, Sydney Major, Victoria Ryan, Marilynn Colato, Victoria Johnson, Ashley Shafter, Rosemary Ravelo, Paige Brooks, and Sierra Walker. Finally, I send my appreciation to my fellow students and colleagues for their ongoing and immeasurable support. Thank you all.

Dedication

This project is dedicated to the many children and families touched by Autism Spectrum Disorder. I have been so fortunate to be able to work with and alongside some of the most amazing humans who bring me inspiration every day and who are a constant reminder about what is good in the world. I would also like to dedicate this project to my family and friends who gave me the strength to bring this project to fruition and who continually provided their love and emotional support.

CHAPTER 1

Individuals diagnosed with Autism Spectrum Disorder (ASD) face many challenges associated with skill deficits. Skill deficits range from mild to severe and can be related to social interactions, verbal and non-verbal communication, and repetitive and restrictive behaviors (American Psychiatric Association, 2013). Behavior analytic researchers have designed and scientifically validated numerous teaching procedures that focus on teaching discriminations between stimuli to ameliorate these skill deficits. A *discrimination* occurs when a target response is emitted and reinforced in the presence of a specific stimulus and not in the presence of other stimuli (Green, 2001). For example, in the presence of a car the spoken word “car” is reinforced, whereas the word “motorcycle” is not. Conversely, in the presence of a motorcycle the spoken word “motorcycle” is reinforced and “car” is not. Most behavior requires some level of discrimination, beginning with the most basic of skills (e.g., labeling, following instructions) to more complex skills (e.g., taking a shower, completing a math problem; Green, 2001). Consequently, individuals with ASD receiving intensive behavioral intervention (IBI) learn numerous skills that require discrimination such as identical matching, non-identical matching, and receptive identification (Green, 2001). Whereas a *simple* *discrimination* (example above) includes three variables, (a) an antecedent stimulus, (b) a response, and (c) a consequence (Green, 2001), a *complex conditional discrimination* includes a fourth variable, a conditional stimulus.

Conditional Discriminations

Conditional discriminations occur when a response is reinforced in the presence of a specific stimulus *if and only if* additional specific stimulus conditions exist (Green, 2001). Therefore, the presence or absence of conditional stimuli signal whether reinforcement is available for a specific response (Green, 2001). For example, following the instruction, “touch car,” touching a car produces reinforcement but touching a motorcycle does not, *and* following the instruction, “touch motorcycle,” touching a motorcycle results in reinforcement but touching a car does not.

Conditional discriminations require prerequisite skills including attending to and differentially responding to sample stimuli, as well as observing and differentially responding to comparison stimuli (McIlvane, Dube, Kledaras, Iennaco, & Stoddard, 1990). Typically developing children acquire these skills (i.e., conditional discriminations) easily through repeated interactions with the environment (Hart & Risley, 1975). However, children with ASD and other related disabilities may have greater difficulty making these discriminations, which may be due in part to skill deficits in vocal (speaker) and non-vocal (listener) responding.

In fact, it has been reported that conditional discriminations can be especially challenging to teach to children with ASD or intellectual disabilities (e.g., Carp, Peterson, Arkel, Petursdottir, & Ingvarsson, 2012; Green, 2001; McIlvane et al., 1990; Perez-Gonzalez & Williams, 2002; Saunders & Spradlin, 1989, 1990, 1993; Romski, Sevcik, & Pate, 1988; Ward & Yu, 2000; Williams, Perez-Gonzalez, & Queiroz, 2005). Though difficult to teach, conditional discriminations open the door for acquisition of new skills, promote independence, and increase the likelihood of generalization of skills to new situations. Therefore, it is important to use effective and efficient procedures to teach discriminations.

Teaching Techniques

Match-to-sample (MTS) is a commonly recommended procedure for teaching conditional discriminations to individuals with intellectual disabilities (e.g., Greer & Ross, 2008; Lovaas, 2003; Maurice, Green, & Luce, 1996; Sundberg & Partington, 1998). During a MTS task the individual is presented with a sample stimulus and prompted to make an observing response. An *observing response* is a response (e.g., touching, pointing) toward the sample stimulus. A sample stimulus serves as the conditional stimulus and the observing response is included to increase the likelihood that the individual is attending to the sample stimulus, thereby increasing the probability that a conditional discrimination will be made. Next, the individual is presented with multiple comparison stimuli in an array and taught to match the correct comparison stimulus (i.e., the target stimulus) to the sample stimulus (Green & Saunders, 1998). The array of comparison stimuli consists of the target stimulus and one or more distractor stimuli. Responding toward any of the distractor stimuli results in no programmed consequences, whereas responding toward the target stimulus produces reinforcement. Therefore, the target stimulus that produces reinforcement is said to be ‘discriminative’ for reinforcement. Two methods commonly used to teach MTS are the simple-conditional and conditional-only training methods.

Simple-conditional training. According to Grow, Carr, Kodak, Jostad, and Kisamore (2011) the simple-conditional training method follows a nine-step teaching process. During the first few steps the individual is required to respond to a single stimulus (i.e., the target) presented in a massed-trial format. However, as the individual advances through the steps, simple discriminations are required and those discriminations become increasingly difficult as distractor stimuli are introduced. During the final step, the individual is required to make conditional discriminations between multiple stimuli (usually three). One problem associated with simple-conditional training is that it may promote faulty stimulus control (Green, 2001) leading to error patterns in responding (Grow et al., 2011). Faulty stimulus control may occur because: (a) the individual may not have the prerequisite skills necessary to make conditional discriminations (i.e., attending and differentially responding to the sample stimulus, and observing and responding to an array of comparison stimuli), and (b) the individual is not required to attend to the relevant stimuli to contact reinforcement (Grow et al., 2011). For example, during Steps 1, 2, and 6, the individual is presented with a sample stimulus and three comparison stimuli. The three comparison stimuli include the target stimulus and two blank distractor stimuli. For these steps, there is no need for the individual to attend to the sample stimulus when engaging in a response because the distractor stimuli are blank. Therefore, it is possible that the individual is simply responding to the stimulus that produced reinforcement during the previous trial and not the relevant features of the target stimulus in relation to the sample stimulus. Even though simple-conditional training may produce faulty stimulus control, a review by Love, Carr, Almason, and Petursdottir (2009) revealed that a greater number of IBI programs teach conditional discriminations using the simple-conditional method versus the conditional-only method, 37% and 31%, respectively.

Conditional-only training. Whereas simple-conditional training includes nine steps, conditional-only training consists of the ninth step alone. While the procedures for these training methods are quite different, the terminal goals are the same, namely, that the individual is taught to make conditional discriminations between multiple stimuli. As previously stated, conditional discriminations require the demonstration of specific prerequisite skills. Using the ninth step exclusively allows the individual to contact reinforcement for only those skills required in the terminal skill and may reduce the likelihood of faulty stimulus control and error patterns in responding. More importantly, using only the ninth step may increase the likelihood that conditional discriminations will be demonstrated. In a study comparing the simple-conditional and conditional-only methods, Grow et al. (2014) found the conditional-only method to be more efficient than the simple-conditional method for acquisition of conditional discriminations in two children with ASD.

Although the literature suggests that the conditional-only method may be more efficient, individuals with ASD or other developmental disabilities may not respond to signals, such as instructions (Green, 2001). In other words, certain instructions (e.g., “Sit down.”) may not be followed by a correct response (e.g., sitting down), which may indicate that the individual is not sensitive to the consequence for correct responding (i.e., availability of reinforcement). However, as suggested by Grow (2014), providing a prompt together with the spoken instruction may strengthen learning.

Errorless Learning

Green (2001) recommends using an errorless teaching strategy to further reduce the number of errors that occur during training. During errorless teaching, the instructor presents prompts that facilitate correct responding simultaneously with or immediately following an instruction. The prompt is then gradually faded, by topography or time, until the individual responds independently. Categories of errorless teaching strategies include within-stimulus prompts and extra-stimulus prompts (Cooper, Heron, & Heward, 2007). A within-stimulus prompt involves changing specific features of the discriminative stimuli (Cooper et al., 2007). For example, the stimulus that produces reinforcement may be larger than the comparison stimuli that produce nothing. Although several within-stimulus prompting strategies have been identified such as intensity fading, criterion-related prompting, and sample stimulus control shaping (see Green, 2001 for a more in-depth explanation of each strategy), these prompts are typically more labor intensive for the instructor. Therefore, extra-stimulus prompts may be more easily incorporated into teaching procedures.

An extra-stimulus prompt is an additional prompt independent from the discriminative stimuli (Cooper et al., 2007). For example, the investigator may gesture toward the stimulus that produces reinforcement or model the action to be completed. These extra-stimulus prompts can easily be faded with respect to intensity and distance or the prompt can be delayed. In a prompt delay the prompt is provided simultaneously with the comparison stimuli (i.e., 0-s prompt delay). On subsequent trials, the prompt is delayed by a predetermined interval, for example 2-s. Following the initial delay, the specified time either remains fixed or is progressively increased (e.g., 0-s, 2-s, 5-s; Walker, 2008). Prompt dependency is an obstacle associated with extra-stimulus prompts. In other words, there is a possibility that the individual may wait for the prompt rather than respond before the prompt is provided. Therefore, as with all prompting procedures the extra-stimulus prompt should be used with caution.

Stimulus Equivalence

Once conditional discriminations are reliably demonstrated between stimuli it is often inferred that equivalence relations are also established (Sidman & Tailby, 1982). In other words, not only is the sample stimulus discriminative for the target stimulus, but the two stimuli are also assumed to be interchangable. Additionally, the presentation of stimuli in a stimulus equivalence arrangement can produce further conditional discriminations that are acquired without direct teaching, making this arrangement extremely efficient.

The specific stimulus arrangement used to teach conditional discriminations results in learners demonstrating stimulus equivalence (Sidman, Kirk, & Willson-Morris 1985). More recently researchers have referred to this arrangement as equivalence-based instruction (EBI; Critchfield, 2014; Fields et al., 2009; Fienup, Covey, & Critchfield, 2010; Fienup & Critchfield, 2010, 2011; Pytte & Fienup, 2012). Using this arrangement for instruction helps facilitate generality of skills across stimuli within a class. In other words, if the individual is taught that the spoken word *"dog"* goes with a picture of a dog and the printed word *dog*, then in the absence of any teaching the individual should be able to associate the picture of the dog with the printed word *dog*. That is, the picture and the printed word are now said to "go together" (Sidman, 1994).

Sidman and Tailby (1982) discuss three different properties of the stimulus equivalence arrangement. The first property is *reflexivity*. Reflexivity is simply identity matching (e.g., A = A). This property has the quality of sameness and is demonstrated in identical matching tasks. The second property associated with stimulus equivalence is *symmetry*. Symmetry has the quality of bi-directionality and therefore states that if A = B, then B = A. In symmetry, only one relation is taught (e.g., A = B) and the other relation (e.g., B = A) emerges due to the previous reinforcement history with the taught relation. The property of symmetry can be demonstrated with non-identical matching tasks. The final property of stimulus equivalence is that of *transitivity*. In transitivity, both relations (e.g., B = C and C = B) are not the result of direct teaching but are relations that emerge because of other taught relations within the stimulus class (e.g., A = B and A = C, therefore B = C). When all taught and emergent relations are demonstrated, a stimulus equivalence class is formed.

For individuals with ASD and other related disabilities, generalization often presents unique challenges. However, researchers have demonstrated that teaching conditional discriminations using EBI with this population promotes generalization

(Green, 2001; McLay, Sutherland, Church, & Tyler-Merrick, 2013; Sprinkle & Miguel, 2012). Generalization occurs when individuals respond correctly to untaught relations (i.e., symmetrical, transitive, and equivalence relations). Additional benefits of EBI are that these procedures can: (a) be used to teach a wide variety of subjects, and (b) increase efficiency in learning for individuals with ASD and other related disabilities, thereby narrowing the gap in skills between them and their typical peers (McLay et al., 2013).

Although most studies examining stimulus equivalence used arbitrary stimuli, more recently, specific stimulus class subjects were taught to individuals with ASD including music (Arntzen, Halstadtro, Bjerke, & Halstadtro, 2010), nouns (Groskreutz, Karsina, Miguel, & Groskreutz, 2010; Sprinkle & Miguel, 2012), money skills (Keintz, Miguel, Kao, & Finn, 2011), geography (LeBlanc, Miguel, Cummings, Goldsmith, & Carr, 2003), and activity schedules (Miguel, Yang, Finn, & Ahearn, 2009). Using educationally relevant stimuli during EBI to teach individuals with ASD increases efficiency by reducing the total number of trials required to teach multiple relations. For example, if a child with ASD were taught six target relations and each relation required (on average) 100 teaching trials to reach mastery, it would take 600 trials to teach those six new target relations. On the other hand, if targets are carefully planned and taught using EBI, that same child could learn six new target relations after teaching just two. In the latter example, only 200 trials would be required to teach the same six new target relations. Therefore, the specific stimulus arrangement used to teach those six new target relations would result in 400 fewer trials. In this same example, the remaining 400 trials could be used to teach an additional 12 relations, resulting in acquisition of 18 relations in the same amount of teaching time. The instructional arrangement therefore allows for a substantial increase in efficiency for teaching children in IBI programs exposed to thousands of teaching trials per day.

Unfortunately, this specific stimulus arrangement (i.e., equivalence-based instruction) for teaching children diagnosed with ASD is not customary in most IBI programs (Green, 2001). This may be due to the effort involved in planning and preparing the specific stimulus classes to be taught or the lack of research on EBI with the ASD population. In a recent review, McLay et al., (2013) identified only nine studies evaluating training of conditional discriminations using EBI. Across those nine studies, McLay and colleagues found that most individuals with ASD demonstrated equivalence. In fact, of the 49 individuals who participated in these studies, 26 responded with equivalence following the first test and an additional 20 demonstrated equivalence with training modifications or after repeated testing. Given the positive outcomes demonstrated within the ASD population, two recommendations by McLay et al. (2013) included that: (a) future research evaluate the effectiveness of these procedures across multiple curriculum domains and (b) future studies examine teaching procedures and variables that may facilitate equivalence.

Although IBI programs have not adopted EBI as a preferred method, researchers continue to study stimulus equivalence, including variables that promote emergent relations. Over the past two decades, researchers identified multiple variables that, when manipulated, increase the likelihood of conditional discrimination and equivalence class formations. These variables include a specific training structure used to teach each relation (i.e., linear series, one-to-many, many-to-one), the use of meaningful stimuli or class-specific reinforcers, and the class size and number of classes being taught.

Variables Enhancing Stimulus Equivalence

Different training structures used to study stimulus equivalence include linear-series training (LS training), one-to-many training (OTM training), and many-to-one training (MTO training) (Arntzen & Holth, 2000). The main difference between the arrangements for EBI is with respect to the training sequence for baseline conditional discriminations. During LS training, baseline conditional discriminations are trained in succession. In other words, the A-B relation is taught until mastery, followed by the B-C relation and so on. In contrast, OTM training occurs when just one of the stimuli within the class serves as the sample during training and the remaining stimuli serve as comparisons. For example, training of the A-B relation occurs first followed by training of the A-C relation (see *Figure 1*). Finally, MTO training is the opposite of OTM training. During MTO training, only one of the stimuli within the class serves as the comparison during training and all other stimuli serve as samples. For example, during MTO training, training of the B-A relation is followed by training of the C-A relation (see *Figure 2*). Fields, Verhave, and Fath (1984) referred to a stimulus that relates to multiple stimuli as a “node.” Thus, in the LS training example above, stimulus “B” is the node, whereas in the OTM and MTO examples, stimulus “A” is the node. Given this, OTM training is also referred to as “sample as node” (SaN) and MTO training is referred to as “comparison as node” (CaN; Fields, Verhave, & Fath, 1984).

Researchers have extensively studied all three training structures (i.e., LS, MTO, and OTM) and most agree that the LS training structure is least effective in producing equivalence (Arntzen, Grondahl, & Eilifsen, 2010; Arntzen & Holth, 1997, 2000; Buffington, Fields, & Adams, 1997; Fields, Landon-Jimenez, Buffington, & Adams, 1995; Fields et al., 1997; Holth & Arntzen, 1998; Saunders & McEntee, 2004). However, results are mixed with respect to whether the MTO training structure or the OTM training structure results in equivalence class formation more consistently. For example, Saunders, Drake and Spradlin (1999) trained five-member stimulus classes with 11 typically developing preschool children. In this study, CaN (MTO) resulted in equivalence class formation for five out of five participants, whereas SaN (OTM) was only effective for two out of six participants. However, Arntzen and Holth (1997) reported results that suggest the opposite. In this study, five out of 10 college students responded positively to equivalence tests following MTO training, whereas 10 out of 10 college students responded positively following OTM training.

Numerous variables may be responsible for differences in outcomes reported as a result of MTO or OTM training. Types of participants have also varied across studies, including: (a) adults or college students (e.g., Arntzen et al., 2010; Arntzen & Hansen, 2011; Arntzen & Holth, 1997, 2000; Fields, Hobbie-Reeve, Adams, & Reeve, 1999; Grisante et al, 2013; Hove, 2003), (b) typical children (e.g., Arntzen & Nikolaisen, 2011; Arntzen & Vaidya, 2008; Saunders, Drake, & Spradlin, 1999; Smeets & Barnes-Holmes, 2005), (c) individuals with intellectual disabilities (e.g., Saunders, Saunders, Williams, & Spradlin, 1993; Saunders et al., 1988; Spradlin & Saunders, 1986), or (d) senior citizens (e.g., Saunders et al., 2005). Additionally, the number of stimulus classes (ranging from two to four) and size of each class trained (ranging from three to seven) varied across studies. Therefore, any conclusions drawn from results across these studies regarding which training structure is most effective would likely be unfounded.

In fact, studies reporting that the MTO and OTM training structures are equally effective (Arntzen et al., 2010; Arntzen & Hansen, 2011; Arntzen & Nikolaisen, 2011; Arntzen & Vaidya, 2008; Grisante et al., 2013) used variables that were consistent across both training structures. However, other studies including consistent variables to examine both training structures have cited more positive outcomes with one training structure over the other (Arntzen & Holth, 1997, 2000; Fields et al., 1999; Hove, 2003; Saunders et al., 2005; Saunders et al., 1999; Saunders et al., 1988; Spradlin & Saunders, 1986). One possible reason for these differences may be that although variables within the studies were held constant (e.g., number of stimulus classes, size of each class), outcomes of training structures were examined between groups (Arntzen & Holth, 1997; Fields et al., 1999; Hove, 2003; Saunders et al., 1999; Saunders et al., 1988). In other words, participants were assigned to one of two groups (MTO or OTM). Therefore, differences in outcomes with these studies may be due to individual differences between the groups and not due to the difference in training structures.

Additionally, studies that examined the two training structures within subjects ran them in succession (Arntzen & Holth, 2000; Saunders et al., 2005; Saunders et al., 1988). In other words, in some instances the MTO training structure was examined first followed by the OTM training structure and in other instances the OTM training structure was examined first followed by the MTO training structure. Therefore, with these studies differences in outcomes across training structures may be due to carry-over effects of the previously introduced training structure. Although in most studies the order of presentation was counterbalanced to control for carry-over effects it is possible that mere exposure to EBI may have affected the overall outcomes. In fact, Fields et al. (2000) found that preliminary training and testing of equivalence relations resulted in enhanced outcomes for subsequent equivalence tests following the training of new conditional discriminations.

Finally, in all studies, the number of trials until mastery was achieved varied across the two training structures. In most cases, participants reached mastery more quickly with the OTM training structure. In these studies, the greater number of training trials needed to reach mastery may have led to more positive equivalence outcomes with the MTO training structure. In other words, better outcomes may be the result of increased exposure to the stimulus relations and not due to the specific training structure used to train those relations.

Only one study examined the effectiveness of the MTO and OTM training structures simultaneously within subject. Arntzen et al., (2010) trained four different stimulus sets of music relations to a 16-year-old boy with ASD using both the MTO and OTM training structures. Initially, the investigator taught the participant 3 three-member stimulus sets, one with major chords (MTO) and one with minor chords (OTM). Next, the number of classes and each class size was expanded so that the participant was eventually taught 4 four-member classes. Under the OTM training structure, the participant demonstrated equivalence for all taught classes after retraining only one set, whereas under the MTO training structure two sets required retraining. Next, the investigator trained three different three-member stimulus sets of major chords and minor chords; however, the training structure used to teach the second set was opposite of what was used to train the first set. As with the first set, both the number of classes and class size were expanded to 4 four-member classes. Results of training the new sets were identical to the initial sets trained in that one set under the OTM training structure required retraining, whereas two sets under the MTO training structure required retraining. Thus, the OTM training structure was more effective than the MTO training structure for teaching music relations to an adolescent with ASD.

Results of the Arntzen et al. (2010) study are limited in three respects. First, only a single participant was used to compare the MTO and OTM training structures. Therefore, additional within-subject studies comparing these training structures simultaneously are warranted. Second, trials to mastery data were not reported for each training structure. Given that the MTO training structure required more retraining than the OTM training structure, it is likely that the participant encountered a greater number of trials during the MTO training. Therefore, it is possible that greater differences in training structure outcomes would be obtained if an equal number of trials were presented across both training structures. In other words, increased exposure to training trials may increase the likelihood of equivalence. Finally, during the test for equivalence both symmetrical and transitive relations were tested. As noted in Smeets, Leader, and Barnes (1997), testing for symmetrical relations exposes participants to the alternate training structure. For example, if a three-member class (A-B-C) were taught using the OTM training structure (A-B and A-C), a test for symmetry would expose the participant to the MTO structure (B-A and C-A). Similarly, training with the MTO structure (B-A and C-A) followed by a test for symmetry (A-C and A-B) exposes the participant to both training structures. Therefore, results of equivalence tests (i.e., transitivity) may be due to the initial training structure (OTM), the test of symmetrical relations (MTO), or a combination of both (Smeets et al., 1997). Consequently, including a test for symmetry prevents an independent analysis of the effectiveness of each training structure. Therefore, a more suitable test for the effectiveness of each training structure in establishing equivalence may be a final test of the transitive relation only.

In addition to the specific training structure used, multiple studies examined the effects of meaningful (i.e., familiar, nameable) stimuli on equivalence class formation. Dickins, Bentall, and Smith (1993) found that participants assigned to the group containing nameable stimuli (i.e., sun, moon, and Saturn) were more likely to form equivalence classes than participants assigned to the group containing abstract stimuli (i.e., unusual shapes and Arabic letters). Additionally, Randell and Remington (1999) found that stimulus classes were formed for 100% of participants when the names of the pictures used rhymed with each other. In contrast, only 30% of participants in the control group formed equivalence. Similarly, Holth and Arntzen (1998) found that 100% of college students formed 3 three-member stimulus equivalence classes when two of the three stimuli per class were familiar pictures, whereas only 30% of college students formed equivalence classes when stimuli consisted of all Greek letters.

Arntzen (2004) extended the research on the use of meaningful stimuli to produce equivalence classes using larger stimulus classes. This study evaluated whether placement of the meaningful picture in an LS training structure influenced stimulus class formation. Equivalence classes were formed for only 30% of participants when all stimuli were arbitrary, for 50% of participants when the last stimulus presented was a familiar picture, and for 100% when the familiar picture was presented first.

Expanding further on whether placement of the meaningful stimulus affects class formation, Arntzen and Lian (2010) analyzed equivalence with the meaningful stimulus as node. This study compared training of three-member classes where members of one stimulus class were all abstract shapes and members of a second stimulus class included a meaningful stimulus as node and two abstract shapes. Following initial training, six out of eight participants formed equivalence classes with a meaningful stimulus as node and only two out of eight participants formed equivalence classes with all stimuli as abstract shapes. Finally, Fields, Arntzen, Nartey, and Eilifsen (2012) examined the effects of meaningful stimuli (i.e., familiar), discriminative stimuli, and meaningless stimuli (i.e., abstract) on equivalence following training of five-member classes. When one stimulus was a meaningful picture, 100% of participants demonstrated equivalence. None of the participants demonstrated equivalence when all the stimuli were abstract shapes; however, when one abstract stimulus first became a discriminative stimulus 50% of participant demonstrated equivalence. Therefore, the discriminative function of stimuli might partially account for enhancement of stimulus class formation when meaningful stimuli are used.

In other studies, researchers examined the use of class-specific reinforcers to establish equivalence classes. Dube, McIlvane, Maguire, Mackay, and Stoddard (1989) demonstrated the formation of stimulus classes as a result of distinct reinforcers delivered for correct responding within two different stimulus classes. In this study, participants with intellectual disabilities were initially taught identity matching for two different sets of stimuli (A1, B1, C1, D1 and A2, B2, C2, D2).

Correct responding during training of Stimulus Set 1 resulted in delivery of food (R1) and correct responding during training of Stimulus Set 2 resulted in delivery of drink (R2). Next, participants were taught arbitrary matching using only the A, B, and C stimuli from each set. Results indicate that all participants formed two separate three-member stimulus classes (A1, B1, C1 and A2, B2, C2) as a result of training. Additionally, each set was expanded to include the corresponding D stimuli, most likely as a result of the stimulus-reinforcer relation established during identity matching. Replications by Schenk (1994) and Goyos (2000) provide further evidence for the role of stimulus-reinforcer relations on stimulus class formation.

Finally, Arntzen and Holth (2000) examined the combined effect of class size and number of classes on stimulus class formation with 50 participants. This study used an LS training structure (i.e., A-B, B-C, C-D, etc.) with a meaningful stimulus in the “B” position. Demonstration of stimulus equivalence was less likely with increases in class size than with number of classes. Whereas five out of five participants trained using 4 three-member stimulus classes demonstrated equivalence class formation, only three out of five participants trained using 3 four-member stimulus classes demonstrated equivalence. In other words, stimulus class formation was more likely when the stimulus class consisted of three members rather than four members. Additionally, the inverse relationship between stimulus class formation and class size was more and more apparent as the class size increased. The same was not true for increased number of classes. In fact, participants continued to demonstrate equivalence class formation with three, four, and five classes when the class size was held constant at three members. Whereas stimulus class formation began to result in negative outcomes when class size increased to four members, it did not result in negative outcomes with increased number of classes until participants reached six classes. Fields et al. (1999) found similar results when class size was increased from five to seven members using a OTM training structure. Although negative outcomes were largely evident during the OTM training structure, this decline was minimal when classes were trained using the MTO structure.

Class Mergers

When individuals learn two or more stimulus classes that share a common member (e.g., A-B-C and A-D-E) the two classes can merge to form one larger stimulus class, referred to as a *class merger* (Sidman et al., 1985). The merger of multiple stimulus classes increases the number of stimulus relations within the stimulus class, many of which are untaught relations. To date, only two studies evaluated the merger of two stimulus classes with one common member using educationally relevant stimuli (i.e., types of alphabets, brain-behavior relations).

In a study by Lane and Critchfield (1998), two adolescents with Down syndrome were taught stimulus relations of letters of the alphabet to spoken words ("vowel" and "consonant") using equivalence-based instruction. Pretests of expected stimulus relations were conducted in a MTS format with reinforcement for on-task behavior only. Correct and incorrect responses produced the next trial. Next, researchers taught participants 2 three-member stimulus classes (one set of vowels and one set of consonants) using 13 different training steps. Initially, a complex sample stimulus (e.g., the spoken word "vowel" and one printed letter) was presented on a computer screen. An observing response removed the sample stimuli and produced two comparison stimuli. Initially, every correct response resulted in reinforcement. The schedule was later thinned so that reinforcement only followed every fourth correct response. Additionally, the number of stimuli in the complex sample was later increased to three (e.g., spoken word "vowel" and two printed letters). Researchers continued training until all baseline relations were mastered. Following training, researchers conducted a posttest of expected stimulus relations (similar to the pretest). Finally, researchers trained and tested two additional three-member stimulus classes (one set of vowels and one set of consonants) using the same procedures. In this study, by training only eight stimulus relations, participants could potentially acquire 32 additional relations without direct teaching. Both participants acquired all emergent relations within each stimulus set and across the two stimulus sets, which demonstrates that a class merger did indeed form for these participants.

In a more recent study by Fienup, et al. (2010), four college students learned facts about brain-behavior relations using MTS procedures that facilitate stimulus equivalence and class mergers. College students learned four sets of two stimulus classes that shared a common member. Researchers hypothesized that, following training of both stimulus classes, a class merger would form and training of only 16 relations would result in a total of 40 acquired relations. Similar to Lane and Critchfield (1998), a pretest, training, post-test design was employed. Prior to any training, all participants scored low on the pretest probes for both stimulus classes. Following training of Class 1, two out of four participants met mastery for the test of emergent relations within that class. Following repeated (i.e., booster) training, the remaining two participants also met mastery. In addition, no improvement in scores was evident for Class 2 or Class Merger relations. Following training for Class 2, all participants demonstrated mastery of emergent relations for that class, retained emergent relations from Class 1 and results of the Class Merger test indicate that the two stimulus classes formed one larger class including members from Class 1 and Class 2, thus demonstrating a class merger.

What remains to be seen is whether these procedures will help facilitate generalization within the ASD population. Individuals diagnosed with ASD and other related disabilities may have deficits with respect to generalization. Thus, teaching skills to individuals diagnosed with ASD using this arrangement may be useful. The efficiency of learning under this paradigm may also help to bridge the gap between individuals diagnosed with ASD and their typical peers. Additionally, discrepancies still exist about which teaching arrangement is more effective, MTO or OTM.

Purpose of Current Study

Teaching conditional discriminations using EBI may result in the emergence of untaught relations. Therefore, the use of this specific stimulus arrangement may facilitate generalization. Teaching two or more classes with one common member may produce a class merger, further improving efficiency and enhancing generalization. However, there are no published studies evaluating the use of EBI to produce class mergers for individuals diagnosed with ASD.

Although the MTO and OTM training structures are widely researched, results across multiple studies vary. The single study that compared these training structures simultaneously found the OTM training structure to be slightly more effective than the MTO training structure (Arntzen et al., 2010). Results of this study are limited, however, because: (a) the MTO and OTM training structures were compared with only a single participant, (b) trials to mastery varied across structures, and (c) tests for symmetry exposed participants to the alternate training structure. Additionally, while the study by Lane and Critchfield (1998) included a complex sample stimulus, all relations were taught using the OTM training structure. Likewise, Fienup et al. (2010) taught all brain-behavior relations using the OTM training structure. Therefore, the effectiveness of the MTO training structure to produce class mergers has yet to be evaluated.

The purpose of the current study was to evaluate the effectiveness of the MTO and OTM training structures in producing equivalence and class mergers using a within-subject simultaneous analysis while holding all variables constant (i.e., number of classes, number of members in each class, and number of trials presented). A second purpose of the current study was to evaluate the extent to which training two stimulus classes with one common member using educationally relevant stimuli would result in derived relations for children diagnosed with ASD. Therefore, this study aimed to: (a) determine which training structure (MTO or OTM) would result in more positive equivalence outcomes, and (b) determine the extent to which children with ASD would demonstrate equivalence and class mergers when using educationally relevant stimuli.

CHAPTER 2

Method

Participants, Settings, Materials

Participants included four children diagnosed with ASD. All participants reliably responded to gestural and echoic prompts. Additionally, each participant demonstrated tacts using flexible three-word phrases and a generalized matching repertoire. Izak, a 7-year, 10-month-old male, attended a local public elementary school in an integrated classroom and had no prior history of behavioral therapy. Alex, a 7-year, 8-month-old male, attended a local public elementary school in a second-grade general education classroom and was receiving approximately 10 hr per week of applied behavior analysis (ABA) therapy. Greg, a 6-year, 2-month old male, attended a local public elementary school in an Exceptional Student Education (ESE) first-grade classroom and was receiving approximately 10 hr per week of ABA therapy. Finally, Luke, a 6-year, 4-month old male, attended a local public elementary school in a Varying Exceptionalities (VE) kindergarten classroom and was receiving approximately 16 hr per week of ABA therapy at the time of the study.

All participants demonstrated weaknesses with respect to communication skills as evidenced by scores on the Vineland Adpative Behavior Scales, Third Edition (Vineland^{TM}-3). The Vineland^{TM}-3 provides a comprehensive, norm-referenced assessment of the adaptive skills of individuals. Scaled scores, having a mean of 15 and a standard deviation of 3, are provided for the adaptive skill areas. The subtests are combined to yield domain scores and an adaptive behavior composite standard score, which have a mean of 100 and standard deviation of 15. Each participant, with the exception of Izak had a Vineland^{TM}-3 conducted by a Board Certified Behavior Analyist within a year of participation in the study.

Alex had v-Scale Scores of 6, 11, and 9 for receptive, expressive, and written subdomains, respectively. His standard score for the communication domain was 69 with an overall adaptive behavior composite score of 71. For Greg, v-Scale Scores were 6, 9, and 7 for receptive, expressive, and written subdomains, respectively. His standard score for the communication domain was 65 with an overall adaptive behavior composite standard score of 69. Finally, Luke's v-Scale Scores were 10, 12, and 9 for receptive, expressive, and written subdomains, respectively. His standard score for the communication domain was 76 with an overall adaptive behavior composite standard score of 77.

During pre-testing and baseline conditions, participants did not respond correctly during greater than 50% of trials across all training targets or investigator defined emergent relations. Children who exhibited severe problem behavior were excluded from this study to minimize risk. The investigator recruited participants through distribution of flyers and with the consent of parents or guardians.

The investigator conducted all sessions in a treatment room at a university-based autism clinic or in the participant’s naturalistic environment (i.e., home, school, therapy center). Izak’s sessions were conducted in a treatment room at a university-based autism clinic. The treatment room was void of distractions and only included a table and chairs, materials required for the study, and an additional observer to collect inter-observer agreement (IOA) and treatment integrity data. Izak’s sessions were conducted three days per week on average. Alex, Greg, and Luke’s sessions were conducted at a local autism center where they received ABA therapy. Research sessions were conducted in the participant’s therapy room that consisted of tables and chairs, toys on a shelf, and wall décor. Additional items present during sessions included materials required for the study and an additional observer to collect IOA and treatment integrity data. Alex and Greg’s sessions were conducted one to two days per week on average and Luke’s session were conducted three days per week on average. Across all participants, an attempt was made to ensure the sessions were administered in a secluded location (absent other adults or children) with other potential distractions removed as well.

Materials present during all sessions included the required instructional materials for training and testing sessions. Also present were preferred edible or tangible items (e.g., iPad) to be used as reinforcers for correct responding during training and to reinforce interspersed mastered tasks and on-task behavior during testing. Materials used during testing and training sessions included visual (i.e., pie charts, fractions, percentages) and auditory (i.e., “What fraction?”, “What percentage?”) stimuli associated with each stimulus class. Individual sample and comparison visual stimuli were printed on laminated cards measuring 7.6 cm by 12.7 cm. Visual stimuli were displayed on a 0.5 cm thick, white, foam graphic art board cut to 30 cm by 51 cm, hereafter referred to as a stimulus board. A second observer collected data for the purpose of calculating inter-observer agreement and treatment integrity measures. Data were collected using paper and pen.

Response Definitions and Measurement

Dependent variable. The dependent variable included the percentage of correct, independent responses across trials during all conditions (i.e., pre-test, baseline, training, interim test, post-test). During the pre-tests, interim tests, and post-tests, percentages were calculated as an average of correct independent responses across the different relations presented during each test. However, for baseline and training conditions, the percentage correct only included the current relation being taught or tested. The target response involved either a selection-based response (e.g., touching, pointing to, removing, or handing to the investigator) of the correct comparison stimulus, with respect to the sample stimulus, when presented in an array with distractors, or a vocal-based response (i.e., tact or intraverbal) when presented with a visual or vocal antecedent. The selection response was individually defined for each participant based on reinforcement history with selection responses. The selection-based response for all participants included pointing to our touching the target stimulus in the comparison array. Responses were scored as correct independent (+), correct prompted (+P), incorrect (-), incorrect prompted (-P), or no response (NR).

Experimental Design and Procedures

Experimental design. The investigator evaluated the data using a modified multiple-probe design across stimulus classes with an embedded multiple schedule (Catania, Horne, & Lowe, 1989; Doughty, Brierley, Eways, & Kastner, 2014). The modified multiple-probe design allowed for analysis of the relation between training and the acquisition of skills by demonstrating that the skills were absent in the participants' repertoire prior to implementation of the training sessions (Cooper et al., 2007). The modified multiple-probe design was similar to that used by Fienup et al. (2010) with two exceptions. First, baseline probes were conducted for each relation prior to training (e.g., an A-B test was conducted prior to A-B training). Second, a pre-test of class merger relations was conducted. The embedded multiple schedule design allowed for analysis of the effectiveness of the MTO and OTM training structures that were implemented simultaneously.

Preference assessment. Prior to each session, the investigator conducted a preference assessment with participants to determine preferred items. Preference assessment choices determined what each participant would earn during research sessions. Preferred items were then presented contingent upon correct responding to mastered tasks interspersed during testing, on-task behavior (i.e., sitting appropriately, attending to the materials, etc.) during testing, and correct responding to experimentally relevant stimuli during training sessions. Each participant had the opportunity to choose a toy or activity they wanted to engage with during breaks.

For Izak, a multiple-stimulus without replacement (MSWO; DeLeon & Iwata, 1996) preference assessment was utilized. During the MSWO, five edibles were presented at equal distances apart in an array on the table in front of him. The order of the presentation was random and items were randomly rotated following each presentation. With each presentation, the investigator instructed Izak to "pick one." Contingent on selecting an edible, Izak was allotted 30 s to consume the edible. Next, the remaining edibles were re-presented. Presentations continued in this manner until all edibles were chosen or Izak failed to respond following two presentations of the same array. During breaks, Izak typically chose to engage with the train set or watch videos on the iPad.

For Alex, Greg, and Luke, the investigator asked what they wanted to work for prior to each session using an open-ended question. Alex typically chose pretzels to be delivered during sessions and the iPad during breaks. Greg typically chose Sour Skittles to be delivered during sessions and play with peers during breaks. Finally, Luke typically chose peppermint candy to be delivered during sessions and play with peers during breaks.

General procedures. Participants learned six sets of stimuli associated with fractions and decimals. The sets of stimuli used for each participant are displayed in Table 1. Stimuli for Set 1 were associated with "two-fifths," Set 2 with "two-sevenths," Set 3 with "five-sevenths," Set 4 with "three-fifths," Set 5 with "three-sixths," and Set 6 with "five-sixths." Each stimulus included five members: (a) a pie chart of the fraction; (b) the written fraction expressed numerically; (c) the vocal fraction expressed numerically; (d) the written fraction expressed as a percentage; and (e) the vocal fraction expressed as a percentage.

Each set of stimuli was divided into two separate stimulus classes with one common member. Class 1 included the A, B, and C stimuli from all stimulus sets and Class 2 included the A, D, and E stimuli from all stimulus sets (see Tables 1 and 2). The sets of stimuli were further divided into two separate groups. Group 1 contained Sets 1 through 3 and Group 2 contained Sets 4 through 6. One group was taught using the MTO training structure and the other group was taught using the OTM training structure. The groups of stimuli taught using each training structure were counterbalanced across participants. In other words, if Participant 1 was taught Group 1 (Sets 1 through 3) using the OTM training structure and Group 2 (Sets 4 through 6) using the MTO training structure, Participant 2 was taught Group 1 (Sets 1 through 3) using the MTO training structure and Group 2 (Sets 4 through 6) using the OTM training structure. Finally, the number of teaching trials was balanced across training structures to keep all variables equal. Izak and Alex were both taught Sets 1 through 3 using the MTO training structure and Sets 4 through 6 using the OTM training structure. Greg and Luke were taught Sets 1 through 3 using the OTM training structure and Sets 4 through 6 using the MTO training structure.

The order in which testing and training sessions were conducted is displayed in Table 2. Testing and training sessions were conducted in the same order across all participants. Figures 1 and 2 depict the stimulus classes, the trained relations for each stimulus class, and the possible emergent relations within and across stimulus classes. For each individual stimulus class, the investigator trained two relations (e.g., A-B and A-C for OTM or B-A and C-A for MTO). It was possible that as a result of the two trained relations, four additional relations would emerge within each class (e.g., B-A, C-A, B-C, and C-B for OTM; A-B, A-C, B-C, and C-B for MTO). Since each class contained a common member (e.g., “A”), it was also possible that eight additional relations across classes would emerge (e.g., B-D, D-B, B-E, E-B, C-D, D-C, C-E, and E-C for both training structures). Therefore, after teaching just four relations across two stimulus classes with one common member, it was possible that 16 additional relations would emerge. Furthermore, because six sets of stimuli (three for each training structure) with four relations each (total of 24 relations) were directly trained, it was possible that 96 additional relations would emerge.

Randomization of trials. The investigator presented sample and comparison stimuli in a pre-determined, randomized order across nine-trial blocks. Sample stimuli were arranged in a pre-determined randomized order so that no one target was repeated more than twice consecutively within a nine-trial block. Additionally, comparison stimuli were presented in a pre-determined randomized order across trials so that each correct comparison appeared in every possible position on an equal number of trials. Visual stimuli were presented with one sample stimulus centered above three comparison stimuli in a linear array and placed on a stimulus board. Prior to the presentation of each trial, participants were required to engage in an observing response (Donahoe & Palmer, 2004). The observing response was required to help increase the likelihood that the participant was attending to the stimuli to be delivered and consisted of the participant touching the sample stimulus (for selection-based responses) or touching a blank stimulus card (for vocal-based responses).

Testing. The investigator conducted tests of possible emergent relations within each stimulus class and across each stimulus class a minimum of three times during the course of the study. Prior to conducting baseline or training sessions, the investigator conducted a pre-test for Class 1, Class 2, and Class Merger (see Table 2 for specific order), and then conducted specific class testing immediately before and immediately following the training for that class. More specifically, the final test before training Class 1 was the test for Class 1 emergent relations and the first test following training for Class 1 was the transitive test for Class 1.

Once a participant reached mastery criteria for both groups within Class 1, the investigator conducted a post-test for the transitive relations within that class. Following the post-test of transitivity for Class 1 the investigator conducted interim testing for Class 2 and the Class Merger. Once a participant reached mastery criteria for both groups within Class 2, the investigator conducted a post-test for the transitive relations within that class. Following the post-test of transitivity for Class 2, the investigator conducted post-tests for Class 1 and the Class Merger. A post-test was defined as a test in which emergent relations were expected due to the prior training sequence. An interim test was defined as a test in which no relations were expected to have emerged.

In the event a participant did not demonstrate transitivity, the investigator conducted a second post-test of symmetry and transitivity. If transitivity was then demonstrated, the investigator moved on to the next step. If transitivity was not demonstrated during the second post-test the investigator reviewed the data to determine which symmetrical relations were demonstrated. If the participant demonstrated symmetry the investigator conducted a third post-test of transitivity. If the participant did not demonstrate symmetrical relations, the investigator trained the symmetrical relations not demonstrated until mastery criteria were met and then conducted a final post-test of transitivity. These procedures were used with Izak, who failed to demonstrate transitive relations during the initial post-test for Class 1, during the second post-test for Class 1, and after direct training of one symmetrical relation (C-A) from the MTO training structure. These procedures were also used with Izak after he failed to demonstrate the transitive relations during the initial and second post-test for Class 2.

For all test conditions (i.e., Pre-Test, Interim Test, Post-Test), the investigator first presented the participant with the stimuli associated with the observing response. For instructions that required a selection response, the investigator delivered an instruction (e.g., “match”) following the observing response and simultaneously presented the relevant stimuli on the stimulus board. For instructions that required a vocal response, once the participant emitted the observing response, the investigator gave an instruction (e.g., “What fraction?”) and presented any relevant stimuli (e.g., written fraction expressed numerically). Participants were given 5 s to respond. There were no programmed consequences for correct, incorrect, and no responses; however, the investigator provided confirmation statements (e.g., “okay”) for responses to all experimentally relevant stimuli. Mastered tasks (i.e., motor imitation, receptive instructions) were interspersed during test sessions. Correct responses to mastered tasks produced reinforcement to facilitate continued responding during the session. Additionally, the investigator provided behavior-specific praise and reinforcement for remaining on-task. Frequent breaks were provided across sessions to prevent participant fatigue.

Baseline. Individual stimulus class and class merger tests included only the possible emergent relations. Therefore, the investigator conducted baseline sessions prior to each training condition to ensure the participant did not already demonstrate the skill being trained. Baseline sessions were conducted for each relation to be taught during Class 1 training (i.e., A-B and A-C for OTM; B-A and C-A for MTO). Similarly, baseline sessions were conducted prior to training for

Class 2. Baseline sessions were conducted in the same manner as test sessions and had no programmed consequences.

Training. The investigator completed training for two relations within each stimulus class. Class 1 trained relations were A-B and A-C for the OTM training structure and B-A and C-A for the MTO training structure. Class 2 trained relations included A-D and A-E for the OTM training structure and D-A and E-A for the MTO training structure. Each stimulus class was trained across six sets of stimuli (e.g., A1-B1, A2-B2, A3-B3, A4-B4, A5-B5, & A6-B6) using two different training structures (i.e., MTO & OTM). The investigator conducted training sessions in the same manner as test sessions; however, responses were prompted using either a gestural prompt (for selection-based responses) or an echoic prompt (for vocal responses). In addition, following each correct response, with respect to the prescribed prompt level or better, the investigator delivered reinforcement using one of the preferred items previously identified during the preference assessment.

The investigator faded gestural and echoic prompts using a progressive prompt delay (Walker, 2008). That is, initially responses were prompted with a 0-s delay (i.e., simultaneously with presentation of the comparison stimuli). Once responding stabilized at the current prompt level, the investigator increased the delay between presentation of the comparison stimuli and the prompt according to the prompt delay schedule. The prompt delay schedule for most participants consisted of five levels; (a) 0-s delay; (b) 2-s delay; (c) 5-s delay; (d) 10-s delay; and (e) independent. Stable responding was defined as two consecutive sessions with no errors at the prescribed prompt level. If at any time the participant erred more than twice within a nine-trial block, the investigator moved back a prompt level (i.e., reduced the prompt delay) during the next nine-trial block for that relation.

Prompt fading procedures varied slightly across participants. For Izak, the prompt delay used during A-B and B-A teaching only included four levels: 0-s delay, 5-s delay, 10-s delay, and independent. For the remainder of teaching trials, a 3-s delay was used instead of a 2-s delay. Additionally, for Izak the investigator did not move back a prompt level following a nine-trial block with two or more errors due to prompt dependency displayed by the participant in the form of waiting for prompts during all prescribed prompt levels. For Alex, all prompting procedures were implemented as outlined and the investigator only moved back a prompt level once, during A-D and D-A teaching. For Greg and Luke, the prompting hierarchy was followed during the initial A-B and B-A teaching. However, due to inattentiveness during research sessions, the investigator returned to a 0-sec delay and added differential reinforcement for independent responses during the prompt fading procedures. Differential reinforcement was used throughout the remainder of the study for Greg and Luke during teaching.

The trained relation was considered mastered when a participant exceeded 88% correct independent responding across three consecutive sessions. If the participant reached mastery criteria for stimuli being taught under one training structure before the other training structure (e.g., OTM before MTO), the investigator continued to present teaching trials under both training structures until each had reached mastery criteria. This was to ensure that the number of teaching trial presentations across the two training structures was equal. Eight potential participants were terminated after either advancing to a Level 5 prompt (i.e., independent) and having to move back more than twice during the training of a single relation or following teaching procedure modifications (e.g., differential reinforcement, conjugate reinforcement, etc.) that were unsuccessful at increasing independent responding.

Session termination criteria. During a session, if a participant engaged in problem behavior lasting longer than two consecutive minutes or posed an immediate danger to himself or herself or the investigator, the session was terminated. If at any time three consecutive or five cumulative sessions were terminated due to problem behavior the participant was terminated from the study. Problem behavior was individually defined for each participant. If termination was required during a session the data gathered during that session was not included in the results. One participant was terminated from the study for problem behavior (i.e., spitting, punching stimulus board, elopement, flopping). Due to termination from the study, that participant's data are not included.

Social Validity, Interobserver Agreement, and Treatment Integrity

Social validity. The investigator provided a survey to caregivers upon completion of the study to collect data on social validity. This measure was designed to assess overall satisfaction and acceptability of the study by caregivers (see Appendix). The survey consisted of five questions on a Likert scale with scores ranging from 1 to 5, with 1 being "strongly disagree" and 5 being "strongly agree".

Interobserver agreement. A second trained investigator observed and scored selection and vocal responses for each participant during a minimum of 35% of sessions. Agreement percentages were calculated using the trial-by-trial agreement method (Kazdin, 2011). This method was calculated by taking the total number of agreements of the target response occurring and dividing it by the total number of agreements plus disagreements and then multiplying this fraction by 100. Table 3 provides detailed information regarding interobserver agreement for each participant. Interobserver agreement was taken during a total of 46.94% of sessions for Izak with an average of 99.85% agreement (range: 88% to 100%). For Alex, interobserver agreement data were collected during 37.25% of sessions with an average of 98.98% agreement (range: 88% to 100%). Interobserver agreement was taken during a total of 55.81% of sessions for Greg with an average of 99.81% agreement (range: 88% to 100%). Finally, for Luke, interobserver agreement data were collected during 33.56% of sessions with an average of 99.55% agreement (range: 88% to 100%).

Treatment integrity. A second trained investigator collected data on treatment integrity and scored each primary investigator behavior as correct (+) or incorrect (–) during a minimum of 35% of sessions. Treatment integrity was calculated as a percentage of primary investigator behaviors implemented correctly during each session. The primary investigator behaviors recorded during testing, baseline, and training sessions included; (a) delivery of the correct instruction, (b) correct prompt provided, and (c) delivery of the correct programmed consequence. Table 4 provides detailed information regarding treatment integrity for each participant. Treatment integrity for Izak was collected during 42.35% of sessions with an average treatment integrity of 99.84% (range: 96% to 100%). For Alex, treatment integrity measures were collected during 40.20% of sessions with an average of 99.33% integrity (range: 88% to 100%). Treatment integrity was collected during 52.81% of sessions for Greg with 100% treatment integrity. Finally, for Luke, treatment integrity was collected during 33.56% of sessions with 100% treatment integrity.

CHAPTER 3

Results

Individual pre-, interim- and post-test results for participants can be viewed in Figure 3, Figure 4, Figure 5, and Figure 6 for Izak, Alex, Greg, and Luke, respectively. Additonally, Figure 7 provides a combined summary of the results across participants. Results are discussed in further detail by participant, followed by a summary of overall results.

Izak

Izak scored at or below 50% across all pre-test probes. More specifically, for the MTO teaching format, he scored 16.67%, 33.33%, and 27.78% for pre-test probes of all possible emergent relations for Class Merger, Class 2, and Class 1, respectively. Additionally, for the OTM teaching format, he scored 8.33%, 36.11%, and 50% for pre-test probes for all possible emergent relations for Class Merger, Class 2, and Class 1, respectively. Based on these scores, Izak did not demonstrate equivalence of the possible emergent relations for any of the classes prior to teaching.

Izak scored low during baseline probes for Class 1. For the MTO teaching format, he scored 33.33% and 44.44% for the B-A and C-A relations, respectively.

For the OTM teaching format, he scored 22.22% and 0% for the A-B and A-C relations, respectively. During B-A (MTO) and A-B (OTM) teaching, Izak met mastery with the A-B (OTM) relation first with 81 teaching trials, followed by the B-A (MTO) relation in 108 teaching trials. However, due to variability in responding to the A-B (OTM) relation following mastery, teaching of both relations (B-A and A-B) continued until they were both high and stable (135 trials each). During C-A (MTO) and A-C (OTM) teaching, Izak met mastery with the C-A (MTO) relation first with 72 teaching trials, followed by the A-C (OTM) relation in 81 teaching trials. However, both relations were presented an equal number of times to ensure equal exposure to each (81 trials). Once mastery of all Class 1 relations was met, interim tests were conducted to determine which possible emergent relations from Class 1 were demonstrated, and to demonstrate experimental control with respect to Class 2 and the Class Merger.

During the interim test for Class 1, Izak did not demonstrate the transitive relations for either the MTO (55.5%) or OTM (55.5%) teaching formats; therefore, testing was conducted to determine if Izak demonstrated any of the symmetrical relations. During the test for symmetry, Izak demonstrated all relations except the symmetrical relation for C-A (A-C = 0%) from the MTO teaching format. Therefore, teaching of the symmetrical relation for C-A (A-C) was conducted until mastery criteria were met (45 trials) and the symmetrical relation for A-C (C-A) was presented for the same number of trials to ensure consistency across teaching formats. Next, a final test of transitivity was conducted. However, Izak continued to respond incorrectly to the B-C relations (0%) across both teaching formats; therefore, we moved on to interim tests for the Class Merger and Class 2 possible emergent relations.

During additional interim tests, Izak scored 19.44% and 0% for the possible emergent relations with respect to the Class Merger for the MTO and OTM teaching formats, respectively. He also scored low (38.88% and 25%) during the Class 2 interim tests for the MTO and OTM teaching formats, respectively. The Class Merger and Class 2 interim test scores remained stable when compared to the pre-test scores demonstrating that the teaching of Class 1 had no effect on the Class Merger or Class 2 outcomes. Therefore, we moved on to baseline probes for Class 2.

Izak scored low during baseline probes for Class 2. For the MTO teaching format, he scored 33.33% and 55.55% for the D-A and E-A relations, respectively. For the OTM teaching format, he scored 0% for both the A-D and A-E relations. During D-A (MTO) and A-D (OTM) teaching, Izak met mastery with the D-A (MTO) relation first (180 trials), followed by the A-D (OTM) relation (198 trials); however, both relations were presented an equal number of times to ensure equal exposure to each (198 trials). During E-A (MTO) and A-E (MTO) teaching, Izak met mastery across the relations at the same time (117 trials). Once mastery of all taught relations for Class 2 was met, post-tests were conducted to determine which possible emergent relations were now demonstrated across both classes and the class merger.

During the post-test for Class 2, Izak did not demonstrate transitive relations for either the MTO (50%) or OTM (44.44%) teaching formats; therefore, testing was conducted to determine if Izak demonstrated any of the symmetrical relations. During the test for symmetry, Izak did not demonstrate any of the symmetrical relations; therefore, teaching on all symmetrical relations was conducted. Izak reached mastery with the MTO symmetrical relations first with 72 trials for the symmetrical relation for D-A (A-D) and 81 trials for the symmetrical relation for E-A (A-E). This was followed by mastery of the OTM symmetrical relations with 126 trials for the symmetrical relation for A-E (E-A) and 135 trials for the symmetrical relation for A-D (D-A). However, the MTO symmetrical relation for D-A (A-D) was presented until the OTM symmetrical relation for A-D (D-A) was mastered and the MTO symmetrical relation for E-A (A-E) was presented until the OTM symmetrical relation for A-E (E-A) was mastered. Following training and mastery of all symmetrical relations, another test for transitivity was conducted. Transitive relations for both MTO (94.44%) and OTM (100%) were demonstrated.

Izak continued to score low on post-tests for Class 1 possible emergent relations with 52.77% and 80.5% for the MTO and OTM teaching formats, respectively. Finally, for this participant, the Class Merger was not demonstrated with 23.61% and 0% for the MTO and OTM teaching formats, respectively.

Two- and four-week maintenance probes were conducted for the Class 1 and Class 2 transitive relations. During the two-week maintenance probes, Izak continued to demonstrate high scores for both Class 1 and Class 2. For the MTO teaching format, Izak scored 100% and 94.44% for the Class 1 and Class 2 transitive relations, respectively. For the OTM teaching format, Izak scored 100% for both Class 1 and Class 2. All transitive relations maintained during the four-week follow-up. For the MTO teaching format, Izak scored 100% and 94.44% for the Class 1 and Class 2 transitive relations, respectively. For the OTM teaching format, Izak scored 94.44% and 100% for the Class 1 and Class 2 transitive relations, respectively. Class merger maintenance probes were not conducted due to no measurable increase occurring from pre- to post-test. Izak's results are displayed in Figure 3.

Alex

Alex scored below 50% on all pre-test probes. More specifically, for the MTO teaching format, he scored 2.77%, 25%, and 25% for pre-tests probes of all possible emergent relations for Class Merger, Class 2, and Class 1, respectively.

Additionally, for the OTM teaching format, he scored 0%, 25%, and 33.33% for pre-tests probes for all possible emergent relations for Class Merger, Class 2, and Class 1, respectively. Based on these scores, Alex did not demonstrate equivalence of the possible emergent relations for any of the classes prior to teaching.

Alex scored low during baseline probes for Class 1. For the MTO teaching format, he scored 0% for both the B-A and C-A relations. Similarly, for the OTM teaching format, he scored 0% for both the A-B and A-C relations. During B-A (MTO) and A-B (OTM) teaching, Alex met mastery with the A-B (OTM) relation first with 54 teaching trials, followed by the B-A (MTO) relation in 81 teaching trials. However, both relations were presented an equal number of times to ensure equal exposure to each (81 trials). During C-A (MTO) and A-C (OTM) teaching, Alex met mastery with the C-A (MTO) relation first with 45 teaching trials, followed by the A-C (OTM) relation in 54 teaching trials. However, both relations were presented an equal number of times to ensure equal exposure to each (54 trials). Once mastery of all Class 1 relations was met, interim tests were conducted to determine which possible emergent relations from Class 1 were demonstrated, and to demonstrate experimental control with respect to Class 2 and the Class Merger.

During the interim test for Class 1, Alex demonstrated the transitive relations for both the MTO and OTM teaching formats at 100% for each.

Therefore, the investigator moved on to interim tests for the Class Merger and Class 2 possible emergent relations. During additional interim tests, Alex scored 0% for the possible emergent relations with respect to the Class Merger for both teaching formats. He also scored low (50% and 44.44%) during the Class 2 interim tests for the MTO and OTM teaching formats, respectively. The Class Merger interim test scores remained stable when compared to the pre-test scores demonstrating that the teaching of Class 1 had no effect on the Class Merger outcomes. The Class 2 interim test scores increased when compared to the pre-test scores; however, they remained at or below the 50% requirement for participation. Therefore, we moved on to baseline probes for Class 2.

Alex scored low during baseline probes for Class 2. For the MTO teaching format, he scored 0% for both the D-A and E-A relations. Similarly, for the OTM teaching format, he scored 0% for both the A-D and A-E relations. During D-A (MTO) and A-D (OTM) teaching, Alex met mastery with the A-D (OTM) relation first (108 trials), followed by the D-A (MTO) relation (117 trials); however, both relations were presented an equal number of times to ensure equal exposure to each (117 trials). During E-A (MTO) and A-E (MTO) teaching, Alex met mastery with the E-A (MTO) relation first (54 trials), followed by the A-E (OTM) relation (72 trials); however, both relations were presented an equal number of times to ensure equal exposure to each (72 trials). Once mastery of all taught relations for Class 2 was met, post-tests were conducted to determine which possible emergent relations were now demonstrated across both classes and the class merger.

During the post-test for Class 2, Alex demonstrated the transitive relations for both the MTO and OTM teaching formats at 100% for each. Additionally, Alex continued to score high on the post-test for Class 1 possible emergent relations with 97.22% and 80.55% for the MTO and OTM teaching formats, respectively. Finally, for this participant, the Class Merger was not demonstrated with 29.16% and 30.55% for the MTO and OTM teaching formats, respectively.

Two- and four-week maintenance probes were conducted for all Class 1 and Class 2 possible emergent relations. During the two-week maintenance probes, Alex continued to demonstrate high scores for Class 1. For the Class 1 MTO teaching format, Alex scored 94.44% and for the Class 1 OTM teaching format, Alex scored 91.67%. Two-week maintenance probes for Class 2 resulted in higher scores for the OTM teaching format at 97.22%; whereas, the MTO teaching format for Class 2 resulted in only 66.67% correct. Four-week maintenance probes were similar; however, there was a decrease in the OTM scores for Class 1. More specifically, Class 1 MTO scores remained high at 97.22%, but Class 1 OTM scores decreased to 55.55%. Four-week maintenance probes for Class 2 remained the same as the two-week maintenance probes. Class merger maintenance probes were not conducted due to no measurable increase occurring from pre- to post-test. Alex's results are displayed in Figure 4.

Greg

Greg scored below 50% on all pre-test probes. More specifically, for the MTO teaching format, he scored 16.66%, 25%, and 25% for pre-tests probes of all possible emergent relations for Class Merger, Class 2, and Class 1, respectively. Additionally, for the OTM teaching format, he scored 18.05%, 33.33%, and 16.66% for pre-tests probes for all possible emergent relations for Class Merger, Class 2, and Class 1, respectively. Based on these scores, Greg did not demonstrate equivalence of the possible emergent relations for any of the classes prior to teaching.

Greg scored low during baseline probes for Class 1. For the MTO teaching format, he scored 22.22% for both the B-A and C-A relations. For the OTM teaching format, he scored 11.11% and 0% for the A-B and A-C relations, respectively. During B-A (MTO) and A-B (OTM) teaching, Greg met mastery with the B-A (MTO) relation first with 153 teaching trials, followed by the A-B (OTM) relation in 243 teaching trials. However, both relations were presented an equal number of times to ensure equal exposure to each (243 trials). During C-A (MTO) and A-C (OTM) teaching, Greg met mastery across the relations at the same time (63 trials). Once mastery of all Class 1 relations was met, interim tests were conducted to determine which possible emergent relations from Class 1 were demonstrated, and to demonstrate experimental control with respect to Class 2 and the Class Merger.

During the interim test for Class 1, Greg demonstrated the transitive relations for both the MTO and OTM teaching formats at 94.44% and 88.88%, respectively. Therefore, the investigator moved on to interim tests for the Class Merger and Class 2 possible emergent relations. During additional interim tests, Greg scored 0.3% and 7.72% for the possible emergent relations with respect to the Class Merger for the MTO and OTM teaching formats, respectively. He also scored low (50% and 44.44%) during the Class 2 interims tests for the MTO and OTM teaching formats, respectively. The Class Merger interim test scores remained stable when compared to the pre-test scores demonstrating that the teaching of Class 1 had no effect on the Class Merger outcomes. The Class 2 interim test scores increased when compared to the pre-test scores; however, they remained at or below the 50% requirement for participation. Therefore, we moved on to baseline probes for Class 2.

Greg scored low during baseline probes for Class 2. For the MTO teaching format, he scored 0% for both the D-A and E-A relations. Similarly, for the OTM teaching format, he scored 0% for both the A-D and A-E relations. During D-A (MTO) and A-D (OTM) teaching, Greg met mastery with the D-A (MTO) relation first (63 trials), followed by the A-D (OTM) relation (108 trials); however, both relations were presented an equal number of times to ensure equal exposure to each (108 trials). During E-A (MTO) and A-E (MTO) teaching, Greg met mastery with the E-A (MTO) relation first (54 trials), followed by the A-E (OTM) relation (108 trials); however, both relations were presented an equal number of times to ensure equal exposure to each (108 trials). Once mastery of all taught relations for Class 2 was met, post-tests were conducted to determine which possible emergent relations were now demonstrated across both classes and the class merger.

During the post-test for Class 2, Greg demonstrated the transitive relations for both the MTO and OTM teaching formats at 100% for each. Additionally, Greg continued to score high on the post-test for Class 1 possible emergent relations with 91.66% and 69.44% for the MTO and OTM teaching formats, respectively. Finally, for this participant, the Class Merger was not demonstrated with 1.38% and 25% for the MTO and OTM teaching formats, respectively.

Two- and four-week maintenance probes were not conducted for this participant due to the family discontinuing services at the autism center. Greg's results are displayed in Figure 5.

Luke

Luke scored below 50% on all pre-test probes. More specifically, for the MTO teaching format, he scored 11.11%, 13.88%, and 22.22% for pre-tests probes of all possible emergent relations for Class Merger, Class 2, and Class 1, respectively. Additionally, for the OTM teaching format, he scored 19.44%, 25%, and 22.22% for pre-tests probes for all possible emergent relations for Class Merger, Class 2, and Class 1, respectively. Based on these scores, Luke did not demonstrate equivalence of the possible emergent relations for any of the classes prior to teaching.

Luke scored low during baseline probes for Class 1. For the MTO teaching format, he scored 33.33% for both the B-A and C-A relations. For the OTM teaching format, he scored 33.33% and 0% for the A-B and A-C relations, respectively. During B-A (MTO) and A-B (OTM) teaching, Luke met mastery with the B-A (MTO) and A-B (OTM) relations at the same time (180 trials). Similarly, during C-A (MTO) and A-C (OTM) teaching, Luke met mastery across the relations at the same time (63 trials). Once mastery of all Class 1 relations was met, interim tests were conducted to determine which possible emergent relations from Class 1 were demonstrated, and to demonstrate experimental control with respect to Class 2 and the Class Merger.

During the interim test for Class 1, Luke demonstrated the transitive relations for both the MTO and OTM teaching formats at 94.44% and 100%, respectively. Therefore, the investigator moved on to interim tests for the Class Merger and Class 2 possible emergent relations. During additional interim tests, Luke scored 12.5% and 18.05% for the possible emergent relations with respect to the Class Merger for the MTO and OTM teaching formats, respectively. He also scored low (19.44% and 25%) during the Class 2 interims tests for the MTO and OTM teaching formats, respectively. The Class Merger and Class 2 interim test scores remained stable when compared to the pre-test scores demonstrating that the teaching of Class 1 had no effect on the Class Merger or Class 2 outcomes. All interim test scores remained at or below the 50% requirement for participation. Therefore, we moved on to baseline probes for Class 2.

Luke scored low during baseline probes for Class 2. For the MTO teaching format, he scored 22.22% and 11.11% the D-A and E-A relations, respectively. For the OTM teaching format, he scored 44.44% and 0% for the A-D and A-E relations, respectively. During D-A (MTO) and A-D (OTM) teaching, Luke met mastery with the A-D (OTM) relation first (90 trials), followed by the D-A (MTO) relation (135 trials); however, both relations were presented an equal number of times to ensure equal exposure to each (135 trials). During E-A (MTO) and A-E (MTO) teaching, Luke met mastery with the E-A (MTO) relation first (81 trials), followed by the A-E (OTM) relation (126 trials); however, both relations were presented an equal number of times to ensure equal exposure to each (126 trials). Once mastery of all taught relations for Class 2 was met, post-tests were conducted to determine which possible emergent relations were now demonstrated across both classes and the class merger.

During the post-test for Class 2, Luke demonstrated the transitive relations for both the MTO and OTM teaching formats at 100% for each. Additionally, Luke scored moderately high on the post-test for Class 1 possible emergent relations with 66.66% for both the MTO and OTM teaching formats. Finally, for this participant, the Class Merger was not demonstrated with 13.88% and 20.83% for the MTO and OTM teaching formats, respectively.

Two- and four-week maintenance probes were conducted for all Class 1 and Class 2 possible emergent relations. During the two-week maintenance probes, Luke continued to demonstrate high scores for Class 1. For the Class 1 MTO teaching format, Luke scored 97.22% and for the Class 1 OTM teaching format, Luke scored 77.77%. Two-week maintenance probes for Class 2 resulted in higher scores for the OTM teaching format at 100%; whereas, the MTO teaching format for Class 2 resulted in only 69.44% correct. Four-week maintenance probes resulted in decreased correct responding across Class 1 and Class 2 for both the OTM and MTO teaching formats. However, more pronounced decreases in responding were seen with MTO training format. More specifically, the Class 1 OTM teaching format decreased by 2.77% and the Class 2 OTM teaching format decreasaed by 41.67%, for an average decrease of 22.22%. Whereas, the Class 1

MTO teaching format decreased by 38.89% and the Class 2 MTO teaching format decreased by 25%, for an average decrease of 31.95%. Class merger maintenance probes were not conducted due to no measurable increase occurring from pre- to post-test. Luke's results are displayed in Figure 6.

Summary of Results

Overall, responding during pre-tests and baseline probes was low across all participants ($\leq$ 50%). With the exception of Greg, the average number of trials to mastery were equal across the different training structures. Greg required more teaching trials during OTM training (average of 261) than during MTO training (average of 166.5). Finally, three out of four participants demonstrated equal outcomes across the two training structures (MTO and OTM). More positive outcomes were demonstrated for the stimulus classes trained using the OTM training structure for Izak. Results for Izak are consistent with the findings of Arntzen et al. (2010) who found that the OTM training structure was slightly more effective than the MTO training structure in producing emergent relations. The primary differences between Arntzen et al. (2010) and the current study are that the number of training trials across training structures was held constant and tests for symmetry were not conducted prior to testing for transitivity. Therefore, results of the current study cannot be attributed to increased exposure to stimulus relations or exposure to the alternative training method (symmetry tests) prior to testing for transitivity.

During tests for transitivity, three out of four participants demonstrated stimulus class formation. Izak required additional teaching of the symmetrical relations prior to demonstrating the transitive relations for Class 1. However, none of the participants demonstrated the Class Merger. These results are not consistent with the finding of Fienup et al. (2010) who found positive outcomes of both equivalence and class merger. However, the current study differs from Fienup et al. (2010) in two respects. Fienup et al. (2010) trained brain-behavior relations to college students using the OTM training structure only; whereas, the current study evaluated equivalence and class mergers in individuals with ASD using both the OTM and MTO training structures. Therefore, failure of participants to demonstrate the Class Merger in this study may be attributed to either the population used (individuals with ASD) or to the fact that two different training structures were evaluated simultaneously. A summary of results are displayed in Figure 7.

CHAPTER 4

Discussion

The current study sought to determine which training structure (MTO or OTM) would result in more positive outcomes and to determine the extent to which children with ASD would demonstrate equivalence and class mergers when using educationally relevant stimuli. Results of this study demonstrate that the different training structures were equally effective across participants. That is, both the MTO and OTM training structure resulted in similar outcomes when compared within and across participants. Additionally, three out of four participants demonstrated equivalence of Class 1 and all participants demonstrated equivalence of Class 2 following initial teaching. Izak required additional teaching to include the symmetrical relations prior to demonstrating equivalence of Class 1. Finally, none of the participants demonstrated a Class Merger.

Results of the MTO and OTM training structure outcomes are consistent with other studies reporting that the MTO and OTM training structures are equally effective when all variables are consistent across training structures (Arntzen et al., 2010; Arntzen & Hansen, 2011; Arntzen & Nikolaisen, 2011; Arntzen & Vaidya, 2008; Grisante et al., 2013). Although results were similar within and across participants, the OTM training structure was slightly more effective at producing derived relations for Izak and Greg. More specifically, Izak only demonstrated 50% of Class 1 emergent relations for the MTO training structure without additional teaching, but demonstrated 75% for the OTM training structure, a 25% difference. Greg demonstrated an average of 66.67% of possible emergent relations for the MTO training structure, but demonstrated 73.61% for the OTM training structure, a 6.94% difference. However, differences in results for Greg should be viewed with caution as these differences could be due to chance responding and were only as a result of responding during the Class Merger post-test. Results of the current study provide additional support for the fact that training structure may not matter when all other variables are held constant, especially for individuals with ASD who could benefit greatly from teaching using EBI.

Results of the Class Merger test were negative across all participants. In other words, none of the participants in this study demonstrated a Class Merger following training and testing of Class 1 and Class 2 relations. These results differ from those reported by Fienup et al. (2010) whereby all participants demonstrated both equivalence within each class and a class merger. Differences in results may be due to the fact that this study evaluated equivalence and a class merger with individuals with ASD, whereas Fienup et al. (2010) used college students only. Another possible reason for differences in results across the two studies is that the current study evaluated the difference between two different training structures (MTO and OTM) trained simultaneously. Fienup et al. (2010) used the OTM training structure alone. Therefore, differences in outcomes could be attributed to the use of multiple training structures or to the number of stimuli being taught. That is, the use of two training structures simultaneously increases the number of stimulus relations being taught at one time.

Similar to Fienup et al. (2010), the current study evaluated stimulus equivalence and class merger outcomes using educationally relevant stimuli. This is especially important given the population used. The positive results for within class emergent relations in the current study provide evidence that EBI can significantly increase the number of relations acquired during teaching. This saves valuable time, especially for clinicians working with individuals with ASD who demonstrate gaps across skill areas when compared to their neuro-typical peers. Since EBI is an effective format for teaching a wide variety of subjects, utilizing this training method in IBI clinics may help bridge that gap. Additionally, determining the specific variables that enhance the formation of stimulus equivalence is important to ensure clinicians are using the most effective and efficient procedures possible.

The current study does not provide much evidence regarding which training structure (MTO or OTM) is more effective at producing emergent relations. Additionally, it does not provide evidence for which variables may be responsible for positive Class Merger outcomes. It may be the case that the training structure (MTO or OTM) does not impact outcomes, but that other variables do. For example, it may be that the number of stimulus classes trained affects outcomes. In the current study, there were only two stimulus classes (Class 1 and Class 2); however, because this study compared two training structures at the same time, there were two Class 1s and Class 2s. That is, there was a OTM Class 1 and a MTO Class 1, as well as a OTM Class 2 and a MTO Class 2. Therefore, teaching four different classes simultaneously could have led to negative outcomes when testing for the Class Merger. Additionally, it may be that individuals with autism have greater difficulty with generalization of skills, which may have led to negative results with respect to the Class Merger.

Similarly, the size of each stimulus class may affect whether a Class Merger is demonstrated. As with the current study, stimulus classes in Fienup et al. (2010) consisted of three members each (e.g., A1, A2, A3; B1, B2, B3; C1, C2, C3). However, unlike the Fienup study, the current study trained two classes containing three members each, one in the MTO training structure and the other in the OTM training structure. More specifically, in the current study, each participant was learning six different relations at the same time. Therefore, it could be that the number of relations being taught at one time affects Class Merger outcomes.

Finally, although errorless teaching methods are strongly recommended for teaching individuals with ASD, errorless teaching alone was not effective for two out of four participants. For these participants, it was necessary to implement differential reinforcement for independent responses in order to increase independent responding. However, it should be noted that a prompt fading procedure was used along with differential reinforcement or prior to differential reinforcement alone, so the effectiveness of differential reinforcement in isolation cannot be determined. Additionally, although the use of errorless teaching is recommended to improve learning (Green, 2001), it is possible that the use of an errorless teaching method resulted in too-tight of stimulus control, thereby suppressing the emergence of relations across the two classes (i.e., Class Merger).

Unlike prior studies, the current study compared the MTO and OTM training structures using a within subjects design with all variables held constant. This study contributes to the literature by providing further evidence to support that the training structure (OTM and MTO) may be irrelevant for which one is more effective at producing emergent relations of symmetry and transitivity during EBI. What has yet to be determined is which variables are relevant for demonstrating a Class Merger when teaching two classes with a common member.

Limitations and Future Directions

Although the current study provides many contributions to the present literature base, some limitations are worth noting. In the current study, the investigator had a difficult time obtaining and maintaining participants. This may be due to the limited scope of the stimuli being taught and to a lack of pre-assessment measures to determine current skill sets already within a potential participant's repertoire. Although prospective participants reliably responded to gestural and echoic prompts and demonstrated tacts and a generalized matching repertoire, no test for attending or scanning was performed. Additionally, no test was performed to identify if stimuli were known. Therefore, it wasn't until pre-tests were introduced that the investigator was able to identify if the stimuli being taught were previously learned. Future studies should design pre-assessment measures that would test for readiness skills (e.g., attending, scanning an array) and known stimuli that would help determine appropriateness for the study.

Another potential limitation to the current study is whether participants contacted the material being taught outside of the study. Given that each participant was in a school setting and the stimuli were educationally relevant, it is possible that participants may have contacted the material in another setting. However, sessions for Izak were run over the summer and Alex, Greg, and Luke were in second grade, first grade, and Kindergarten, respectively. According to the Florida

State Standards, fractions and percentages are not taught until the third grade. Therefore, it is unlikely that any of the participants in the current study contacted the material being taught during participation in the study. However, future studies should control for this variable.

Finally, the Class Merger was not demonstrated with any of the participants. There are multiple reasons why participants may have failed the Class Merger test. First, participation in the study lasted two months for Izak, three months for Alex and Luke, and five months for Greg. Therefore, the amount of time between learning the first class and the second class could have affected Class Merger outcomes. However, each participant was tested for Class 1 prior to the final post-test for the Class Merger and all but one (Alex) demonstrated maintenance of Class 1 emergent relations. A second reason that participants may not have demonstrated a Class Merger is due to the similarity of the stimuli being taught. For example, during the MTO Class Merger test, Alex frequently answered "5/7" when 2/7 was the correct answer and "2/7" when 5/7 was the correct answer. Additionally, during the OTM Class Merger test, Alex frequently confused 50% and 60% and 3/5 and 3/6. A final reason that participants may not have demonstrated a Class Merger is due to the number of classes or the number of stimuli being taught at the same time. This study consisted of four classes with six relations being taught at the same time while other studies that have demonstrated a class merger only consisted of two classes with three relations being taught at once. It is difficult to tease out the exact reason participants in this study did not demonstrate a Class Merger. Therefore, additional studies are warranted to determine the variables that more reliably result in a Class Merger.

Figures

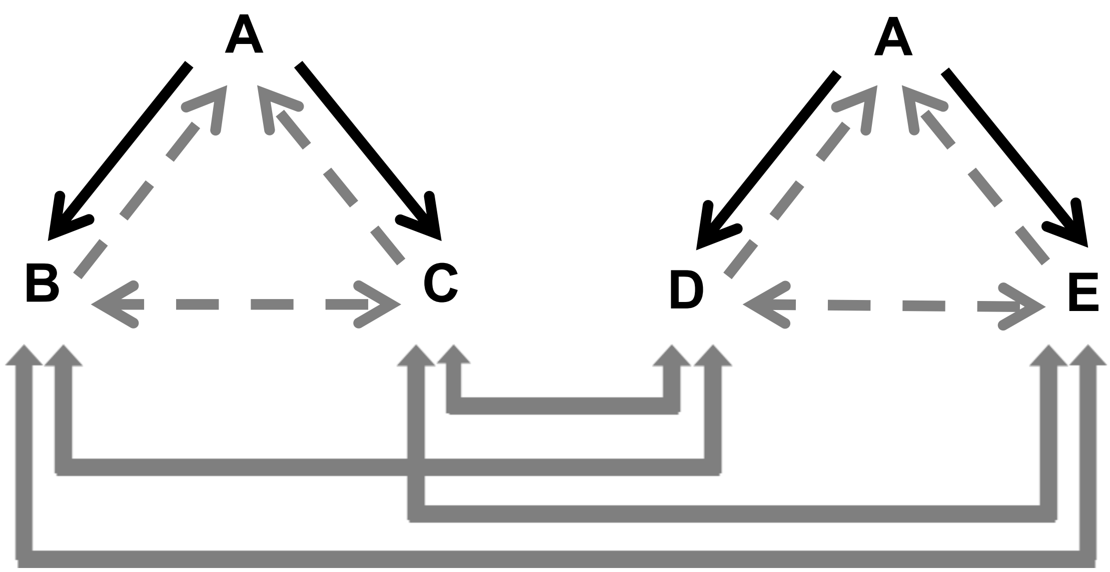


Figure 1. Representation of each stimulus class for OTM training structure. Black arrows indicate trained relations. Grey dashed arrows indicate emergent relations within each stimulus class. Solid grey arrows indicate emergent relations across stimulus classes.

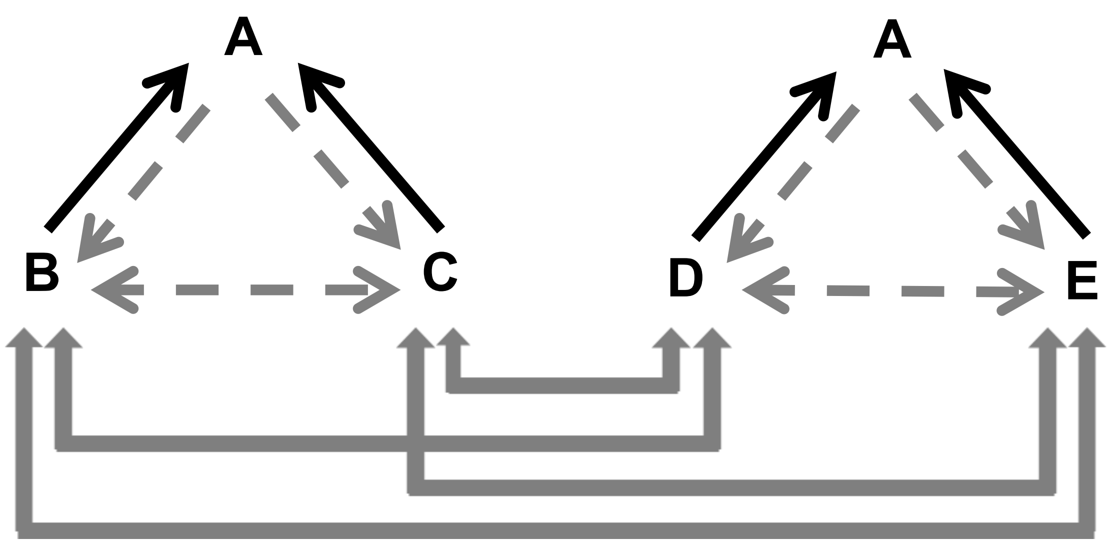


Figure 2. Representation of each stimulus class for MTO training structure. Black arrows indicate trained relations. Grey dashed arrows indicate emergent relations within each stimulus class. Solid grey arrows indicate emergent relations across stimulus classes.

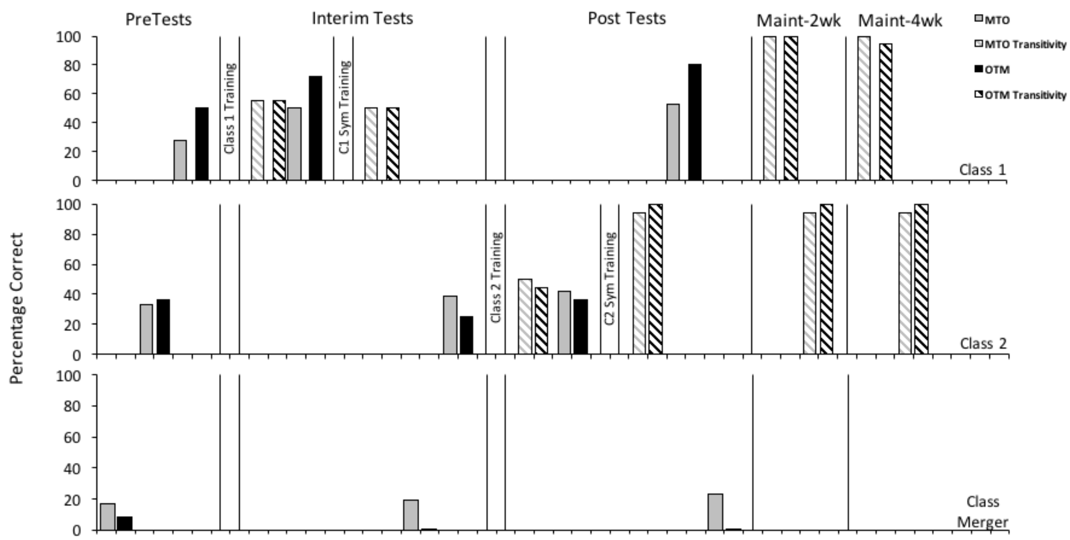


Figure 3. Percentage of correct responding for Izak during pre-, interim-, post-tests and maintenance probes for Class 1 (top panel), Class 2 (middle panel) and Class Merger (bottom panel). Solid gray bars indicate many-to-one testing of all possible emergent relations and solid black bars indicate one-to-many testing of all possible emergent relations. Shaded gray bars indicate many-to-one testing of only transitive relations shaded black bars indicate one-to-many testing of only transitive relations.

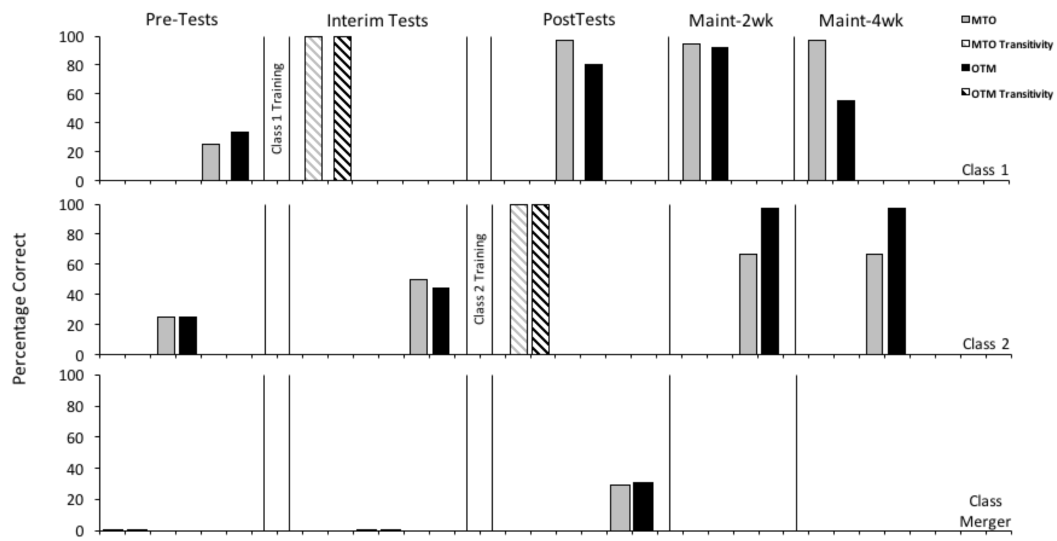


Figure 4. Percentage of correct responding for Alex during pre-, interim-, post-tests and maintenance probes for Class 1 (top panel), Class 2 (middle panel) and Class Merger (bottom panel). Solid gray bars indicate many-to-one testing of all possible emergent relations and solid black bars indicate one-to-many testing of all possible emergent relations. Shaded gray bars indicate many-to-one testing of only transitive relations shaded black bars indicate one-to-many testing of only transitive relations.

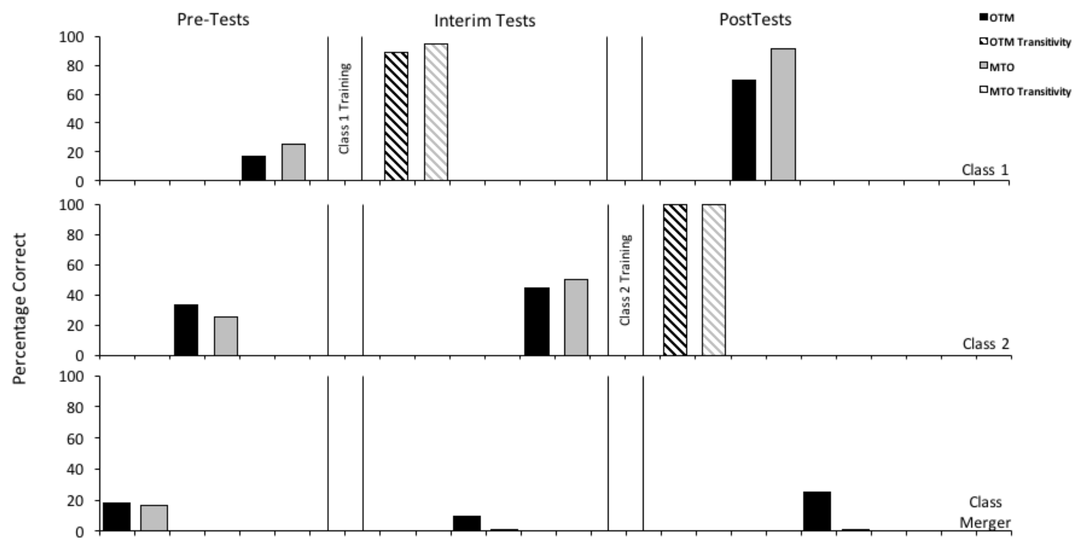


Figure 5. Percentage of correct responding for Greg during pre-, interim-, post-tests and maintenance probes for Class 1 (top panel), Class 2 (middle panel) and Class Merger (bottom panel). Solid gray bars indicate many-to-one testing of all possible emergent relations and solid black bars indicate one-to-many testing of all possible emergent relations. Shaded gray bars indicate many-to-one testing of only transitive relations shaded black bars indicate one-to-many testing of only transitive relations.

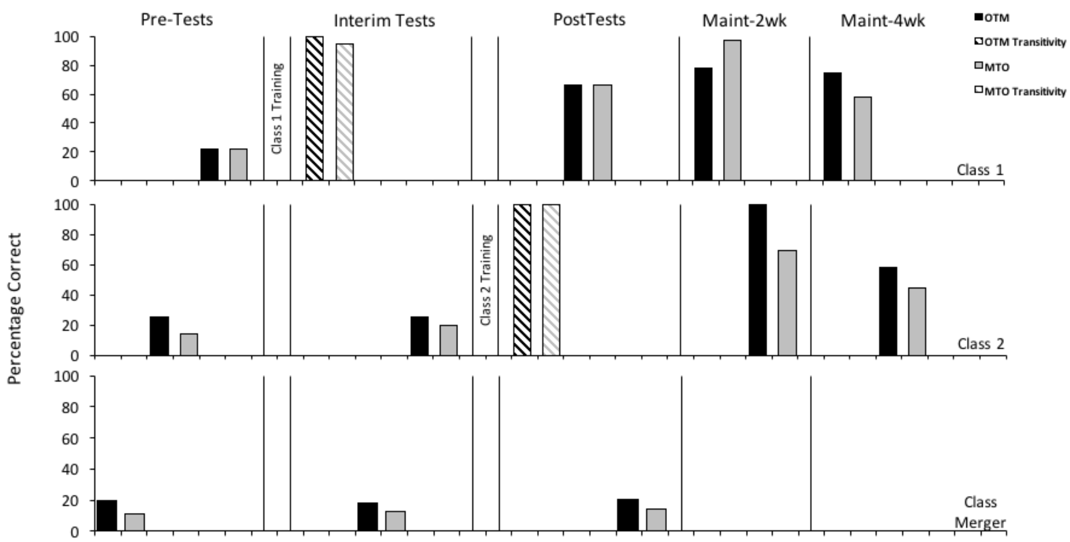


Figure 6. Percentage of correct responding for Luke during pre-, interim-, post-tests and maintenance probes for Class 1 (top panel), Class 2 (middle panel) and Class Merger (bottom panel). Solid gray bars indicate many-to-one testing of all possible emergent relations and solid black bars indicate one-to-many testing of all possible emergent relations. Shaded gray bars indicate many-to-one testing of only transitive relations shaded black bars indicate one-to-many testing of only transitive relations.

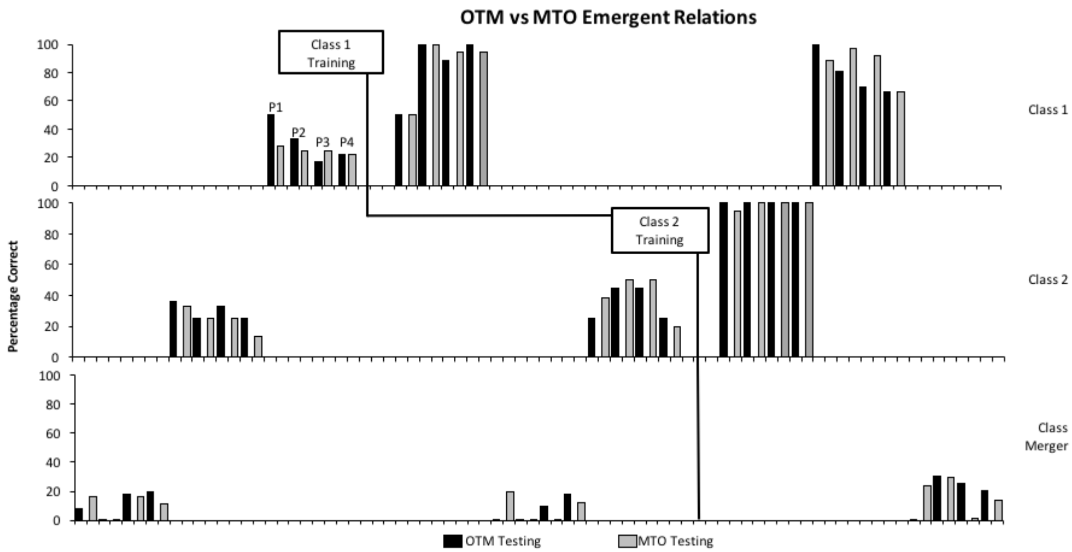


Figure 7. Percentage of correct responding for four participants (Izak-P1, Alex-P2, Greg-P3, & Luke-P4) during pre-, interim- and post-tests of Class 1 (top panel), Class 2 (middle panel), and Class Merger (bottom panel) relations. Solid bars represent tests completed for stimuli trained using the OTM structure and shaded bars represent tests completed for stimuli trained using the MTO structure.

Tables

Table 1

Stimulus Sets used in Testing and Training

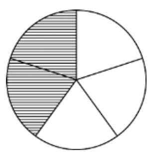
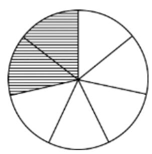
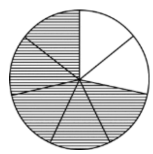
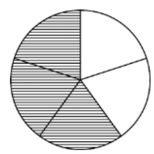
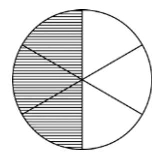
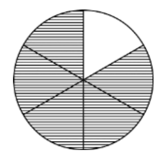
Stimulus	Set 1	Set 2	Set 3	Set 4	Set 5	Set 6
A						
B	2/5	2/7	5/7	3/5	3/6	5/6
C	“two-fifths	“two-seventh	“five-sevenths”	“three-fifths”	“three-sixth	“five-sixths”
D	40%	29%	71%	60%	50%	83%
E	“40 percent”	“29 percent”	“71 percent”	“60 percent”	“50 percent”	“83 percent”

Table 2

Order of Testing and Training

Phase	**Class**	**Requirement**
Pre-Test	Class Merger (OTM & MTO)	Average ≤ 50%
Pre-Test	Class 2 (OTM & MTO)	Average ≤ 50%
Pre-Test	Class 1 (OTM & MTO)	Average ≤ 50%
Baseline	A-B (OTM) and B-A (MTO)	≤ 50%
Baseline	A-C (OTM) and C-A (MTO)	≤ 50%
Training	A-B and B-A	3 Sessions > 88%
Training	A-C and C-A	3 Sessions > 88%
Post-Test	Class 1 Transitivity (OTM & MTO)	Average > 88%
Interim Test	Class Merger (OTM & MTO)	Average ≤ 50%
Interim Test	Class 2 (OTM & MTO)	Average ≤ 50%
Baseline	A-D (OTM) and D-A (MTO)	≤ 50%
Baseline	A-E (OTM) and E-A (MTO)	≤ 50%
Training	A-D and D-A	3 Sessions > 88%
Training	A-E and E-A	3 Sessions > 88%
Post-Test	Class 2 Transitivity (OTM & MTO)	Average > 88%
Post-Test	Class 1 (OTM & MTO)	Average > 88%
Post-Test	Class Merger (OTM & MTO)	Average > 88%
2-Week Maintenance	Class 1, Class 2, Class Merger (OTM & MTO)	
4-Week Maintenance	Class 1, Class 2, Class Merger (OTM & MTO)	

Table 3

Interobserver Agreement (IOA) Results

Participant	IOA		
	Collected on % of Sessions	Average Percentage	Range
Izak	46.94%	99.85%	88%-100%
Alex	37.25%	98.98%	88%-100%
Greg	55.81%	99.81%	88%-100%
Luke	33.56%	99.55%	88%-100%

Table 4

Treatment Integrity (TI) Results

Participant	IOA		
	Collected on % of Sessions	Average Percentage	Range
Izak	42.35%	99.84%	96%-100%
Alex	40.20%	99.33%	88%-100%
Greg	52.81%	100%	N/A
Luke	33.56%	100%	N/A

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.

Arntzen, E. (2004). Probability of equivalence formation: Familiar stimuli and training sequence. *The Psychological Record*, *54*, 275-291. doi: 10.1007/BF03395474

Arntzen, E., Grondahl, T., & Eilifsen, C. (2010). The effects of different training structures in the establishment of conditional discriminations and subsequent performance on tests for stimulus equivalence. *The Psychological Record, 60*, 437-462. doi: 10.1007/BF03395720

Arntzen, E., Halstadtro, L. B., Bjerke, E., & Halstadtro, M. (2010). Training and testing music skills in a boy with autism using a matching-to-sample format. *Behavioral Interventions, 25*, 129-143. doi: 10.1002/bin.301

Arntzen, E., & Hansen, S. (2011). Training structures and the formation of equivalence classes. *European Journal of Behavior Analysis, 12*, 483-503. doi: 10.1007/s40732-018-0290-3

Arntzen, E., & Holth, P. (1997). Probability of stimulus equivalence as a function of training design. *The Psychological Record, 47*, 309-320. doi: 10.1007/BF03395227

Arntzen, E., & Holth, P. (2000). Probability of stimulus equivalence as a function of class size vs. number of classes. *The Psychological Record, 50,* 79-104. doi: 10-1007/BF03395344

Arntzen, E., & Lian, T. (2010). Trained and derived relations with pictures as nodes. *The Psychological Record*, *28*, 181-191. doi: 10.1007/BF03395738

Arntzen, E., & Nikolaisen, S. L. (2011). Establishing equivalence classes in children using familiar and abstract stimuli and many-to-one and one-to-many training structures. *European Journal of Behavior Analysis*, *12*, 105-120. doi: 10.1080/15021149.2011.11434358

Arntzen, E., & Vaidya, M. (2008). The effect of baseline training structure on equivalence class formation in children. *Experimental Analysis of Human Behavior Bulletin*, *29,* 1-8. doi: https://www.researchgate.net/publication/237010587

Buffington, D. M., Fields, L., & Adams, B. J. (1997). Enhancing equivalence class formation by pretraining of other equivalence classes. *Psychological Record*, *47*, 69-96. doi: 10.1007/BF03395213

Carp, C. L., Peterson, S. P., Arkel, A. J., Petursdottir, A. I., & Ingvarsson, E. T. (*2012*). A further evaluation of picture prompts during auditory-visual conditional discrimination training. *Journal of Applied Behavior Analysis, 45*, 737-751. doi: 10.1901/jaba.2012.45-737

Catania, A. C., Horne, P., & Lowe, D. F. (1989). Transfer of function across members of an equivalence class. *The Analysis of Verbal Behavior, 7*, 99-110. doi: 10.1007/BF03392841

Cooper, J. O., Heron, T. E., & Heward, W. L. (2007). *Applied behavior analysis* (2nd ed.). Upper Saddle River, NJ: Pearson.

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. doi: 10.1002/jaba.150

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

Dickins, D. W., Bentall, R. P., & Smith, A. B (1993). The role of individual stimulus names in the emergence of equivalence relations: The effects of interpolated paired-associates training of discordant associations between names. *The Psychological Record, 43*, 713-724. doi: 10.1007/BF03395908

Donahoe, J. W., & Palmer, D. C. (2004). *Learning and complex behavior*. Richmond, MA: Ledgetop Publishing.

Doughty, A. H., Brierley, K. P., Eways, K. R., & Kastner, R. M. (2014). Effects of stimulus discriminability on discrimination acquisition and stimulus-equivalence formation: Assessing the utility of a multiple schedule. *Psychological Record*, *64*, 287-300. doi: 10.1007/s40732-014-0001-7

Dube, W. V., McIlvane, W. J., Maguire, R. W., Mackay, H. A., & Stoddard, L. T. (1989). Stimulus class formation and stimulus-reinforcer relations. *Journal of the Experimental Analysis of Behavior, 51*, 65-76. doi: 10.1901/jeab.1989.51-65

Fields, L., Arntzen, E., Nartey, R. K., & Eilifsen, C. (2012). Effects of a meaningful, a discriminative, and a meaningless stimulus on equivalence class formation. *Journal of the Experimental Analysis of Behavior*, *97*, 163-181. doi: 10.1901/jeab.2012.97-163

Fields, L., Hobbie-Reeve, S. A., Adams, B. J., & Reeve, K. F. (1999). Effects of training directionality and class size on equivalence class formation by adults. *The Psychological Record*, *49*, 703-724. doi: 10.1007/BF03395336

Fields, L., Landon-Jimenez, D. V., Buffington, D. M., Adams, B. J. (1995). Maintained nodal-distance effects in equivalence classes. *Journal of the Experimental Analysis of Behavior, 64*, 129-145. doi: 10.1901/jeab.1995.64-129

Fields, L., Reeve, K. F., Rosen, D., Varelas, A., Adams, B. J., Belanich, J., & Hobbie, S. A. (1997). Using the simultaneous protocol to study equivalence class formation: The facilitating effects of nodal number and size of previously established equivalence classes. *Journal of the Experimental Analysis of Behavior, 67*, 367-389. doi: 10-1901/jeab.1997.67-367

Fields, L., Travis, R., Roy, D., Yadlovker, E., Aguiar-Rocha, L. D., & Sturmey, P. (2009). Equivalence class formation: A method for teaching statistical interactions. *Journal of Applied Behavior Analysis, 42*, 757-593. doi: 10.1901/jaba.2009.42-575

Fields, L., Varelas, A., Reeve, K. F., Belanich, J., Wadhwa, P., Derosse, P., & Rosen, D. (2000). Effects of prior conditional discrimination training, symmetry, transitivity, and equivalence testing on the emergence of new equivalence classes. *The Psychological Record, 50*, 443-466. doi: 10.1007/BF03395365

Fields, L., Verhave, T., & Fath, S. (1984). Stimulus equivalence and transitive associations: A methodological analysis. *Journal of the Experimental Analysis of Behavior, 42*, 143-158. doi: 10.1901/jeab.1984.42-143

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*, 19-33. doi: 10.1901/jaba.2010.43-19

Fienup, D. M., & Critchfield, T. S. (2010). Efficiently establishing concepts of inferential statistics and hypothesis decision making through contextually controlled equivalence classes. *Journal of Applied Behavior Analysis, 43*, 437-462. doi: 10.1901/jaba.2010.43-437

Fienup, D. M., & Critchfield, T. S. (2011). Transportability of equivalence-based programmed instruction: Efficacy and efficiency in a college classroom. *Journal of Applied Behavior Analysis, 44*, 435-450. doi: 10.1901/jaba.2011.44-435

Goyos, C. (2000). Equivalence class formation via common reinforcers among preschool children. *The Psychological Record, 50*, 629-654. doi: 10.1007/BF03395375

Green, G. (2001). Behavior analytic instruction for learners with autism: Advances in stimulus control technology. *Focus on Autism and Other Developmental Disabilities, 16*, 72-85. doi: 10.1177/108835760101600203

Green, G., & Saunders, R. R. (1998). Stimulus equivalence. In K. A. Lattal & M. Perone (Eds.), *Handbook of research methods in human operant behavior* (pp. 229-262). New York: Plenum Press.

Greer, R. D., & Ross, D. E. (2008). *Verbal behavior analysis: Inducing and expanding new verbal capabilities in children with language delays.* Boston, MA: Pearson.

Grisante, P. C., Galesi, F. L., Sabino, N. M., Debert, P., Arntzen, E., & McIlvane, W. J. (2013). Go/no-go procedures with compound stimuli: Effects of training structure on the emergence of equivalence classes. *The Psychological Record, 63*, 63-72. doi: 10-11133/j.tpr.2013.63.1.005

Groskreutz, N. C., Karsina, A., Miguel, C. F., & Groskreutz, M. P. (2010). Using complex auditory-visual samples to produce emergent relations in children with autism. *Journal of Applied Behavior Analysis, 43*, 131-136. doi: 10-1901/jaba.2010.43-131

Grow, L. L., Carr, J. E., Kodak, T. M., Jostad, C. M., & Kisamore, A. N. (2011). A comparison of methods for teaching receptive labeling to children with autism spectrum disorders. *Journal of Applied Behavior Analysis, 44*, 475-498. doi: 10.1901/jaba.2011.44-475

Hart, B. M., & Risley, T. R. (1975). Incidental teaching of language in the preschool. *Journal of Applied Behavior Analysis, 8*, 411-420. doi: 10.1901/jaba.1975.8.411

Holth, P., & Arntzen, E. (1998). Stimulus familiarity and the delayed emergence of stimulus equivalence or consistent nonequivalence. *The Psychological Record, 48*, 81-110. doi: 10.1007/BF03395260

Hove, A. (2003). Differential probability of equivalence class formation following a one-to-many versus a many-to-one training structure. *The Psychological Record*, *53*, 617-634. doi: 10.1007/BF03395456

Kazdin, A. E. (2011). *Single-case research designs: Methods for clinical and applied settings* (2nd ed.). Oxford, NY: Oxford University Press.

Keintz, K. S., Miguel, C. F., Kao, B., & Finn, H. E. (2011). Using conditional discrimination training to produce emergent relations between coins and their values in children with autism. *Journal of Applied Behavior Analysis, 44*, 909-913. doi: 10.1901/jaba.2011.44-909

Lane, S. D. & Critchfield, T. S. (1998). Classification of vowels and consonants by individuals with moderate mental retardation: Development of arbitrary relations via match-to-sample training with compound stimuli. *Journal of Applied Behavior Analysis, 31*, 21-41. doi: 10.1901/jaba.1998.31-21

LeBlanc, L. A., Miguel, C. F., Cummings, A. R., Goldsmith, T. R., & Carr, J. E. (2003). The effects of three stimulus-equivalence testing conditions on emergent US geography relations of children diagnosed with autism. *Behavioral Interventions, 18*, 279-289. doi: 10.1002/bin.144

Lovaas, O. I. (2003). *Teaching individuals with developmental delays: Basic intervention techniques.* Austin, TX: Pro-Ed.

Love, J. R., Carr, J. E., Almason, S. M., & Petursdottir, A. I. (2009). Early and intensive behavioral intervention for autism: A survey of clinical practices. *Research in Autism Spectrum Disorders, 3,* 421-428. doi: 10.1016/j.rasd.2008.08.008

Maurice, C., Green, G., & Luce, S. (Eds.). (1996). *Behavioral intervention for young children with autism: A manual for parents and professionals.* Austin, TX: Pro-Ed.

McIlvane, W. J., Dube, W. V., Kledaras, J. B., Iennaco, F. M., & Stoddard, L. T. (1990). Teaching relational discrimination to individuals with mental retardation: Some problems and possible solutions. *American Journal on Mental Retardation, 95,* 283-296.

McLay, L. K., Sutherland, D., Church, J., & Tyler-Merrick, G. (2013). The formation of equivalence classes in individuals with autism spectrum disorder: A review of the literature. *Research in Autism Spectrum Disorders, 7,* 418-431. doi: 10.1016/j.rasd.2012.11.002

Miguel, C. F., Yang, H., Finn, H. & Ahearn, W. (2009) Establishing derived textual control in activity schedules with children with autism. *Journal of Applied Behavior Analysis, 42,* 703-709. doi: 10.1901/jaba.2009.42-703

Perez-Gonzalez, L. A., & Williams, G. (2002). Multicomponent procedure to teach conditional discriminations to children with autism. *American Journal on Mental Retardation, 107*, 293-301. doi: 10.1352/0895-8017(2002)107<0293:MPTTCD>2.0.CO;2

Pytte, C. L., & Fienup, D. M. (2012). Using equivalence-based instruction to increase efficiency in teaching neuroanatomy. *The Journal of Undergraduate Neuroscience Education, 10*, 125-131. doi: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3598087/

Randell, T., & Remington, B. (1999). Equivalence relations between visual stimuli: The functional role of naming. *Journal of the Experimental Analysis of Behavior, 71*, 395-415. doi: 10.1901/jeab.1999.71-395

Romski M. A., Sevcik, R. A., Pate, J. L. (1988). Establishing of symbolic communication in persons with severe retardation. *Journal of Speech and Hearing Disorders, 53*, 94-107. doi: 10.1044/jshd.5301.94

Saunders, K. J., Saunders, R. R., Williams, D. C., & Spradlin, J. E. (1993). An interaction of instructions and training design on stimulus class formation: Extending the analysis of equivalence. *The Psychological Record, 43*, 725-745. doi: 10.1007/BF03395909

Saunders, K. J., & Spradlin, J. E. (1990). Conditional discrimination in mentally retarded subjects: The development of generalized skills. *Journal of the Experimental Analysis of Behavior, 54,* 239-250. doi: 10.1901/jeab.1990.54-239

Saunders, K. J., & Spradlin, J. E. (1993). Conditional discrimination training in mentally retarded subjects: Programming acquisition and learning set. *Journal of the Experimental Analysis of Behavior, 60,* 571-585. doi: 10.1901/jeab.1993.60-571

Saunders, R. R., Chaney, L., & Marquis, J. G. (2005). Equivalence class establishment with two-, three-, and four-choice matching to sample by senior citizens. *The Psychological Record, 55,* 539-559. doi: 10.1007/BF03395526

Saunders, R. R., Drake, K. M., & Spradlin, J. E. (1999). Equivalence class establishment, expansion, and modification in preschool children. *Journal of the Experimental Analysis of Behavior*, *71*, 195-214. doi: 10.1901/jeab.1999.71-195

Saunders, R. R., & McEntee, J. E. (2004). Increasing the probability of stimulus equivalence with adults with mild mental retardation. *The Psychological Record, 54,* 423-435. doi: 10.1007/BF03395483

Saunders, R. R., & Spradlin, J. E. (1989). Conditional discrimination in mentally retarded adults: The effect of training the component simple discriminations. *Journal of the Experimental Analysis of Behavior, 52*, 1-12. doi: 10.1901/jeab.1989.52-1

Saunders. R. R., Wachter, J., & Spradlin, J. E. (1988). Establishing auditory stimulus control over an eight-member equivalence class via conditional discrimination procedures. *Journal of the Experimental Analysis of Behavior*, *49*, 95-115. doi: 10.1901/jeab.1988.49-95

Schenk, J. J. (1994). Emergent relations of equivalence generated by outcome-specific consequences in conditional discrimination. *The Psychological Record, 44*, 537-558. doi: 10.1007/BF03395143

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

Sidman, M., Kirk, B., & Willson-Morris, M. (1985). Six-member stimulus classes generated by conditional-discrimination procedures. *Journal of the Experimental Analysis of Behavior, 43,* 21-42. doi: 10.1901/jeab.1985.43-21

Sidman, M. & Tailby, W. (1982). Conditional discrimination vs. matching to sample: An expansion of the testing paradigm. *Journal of the Experimental Analysis of Behavior, 37,* 5-22. doi: 10.1901/jeab.1982.37-5

Smeets, P. M., & Barnes-Holmes, D. (2005). Establishing equivalence classes in preschool children with one-to-many and many-to-one training protocols. *Behavioural Processes*, *69*, 281-293. doi: 10.1016/j.beproc.2004.12.009

Smeets, P. M., Leader, G., & Barnes, D. (2000). Establishment of stimulus classes in adults and children using a respondent-type training procedure: A follow-up study. *The Psychological Record, 47*, 285-308.

Spradlin, J. E., & Saunders, R. R. (1986). The development of stimulus classes using match-to-sample procedures: Sample classification versus comparison classification. *Analysis and Intervention in Developmental Disabilities, 6*, 41-58. doi: http://dx.doi.org/10.1016/0270-4684(86)90005-4

Sprinkle, E. C., & Miguel, C. F. (2012). Effects of listener and speaker training on emergency relations in children with autism. *The Analysis of Verbal Behavior*, *28*, 111-117. doi: 10.1007/BF03393111

Sundberg, M. L., & Partington, J. W. (1998). *Teaching language to children with autism or other developmental disabilities.* Pleasant Hill, CA: Behavior Analysts, Inc.

Walker, G. (2008). Constant and progressive time delay procedures for teaching children with autism: A literature review. *Journal of Autism and Developmental Disorders, 38*, 261-275. doi: 10.1007/s10803-007-0390-4

Ward, R., & Yu, C. T. (2000). Bridging the gap between visual and auditory discrimination learning in children with autism and severe developmental disabilities. *Journal on Developmental Disabilities, 7*, 142-155.

Williams, G., Perez-Gonzalez, L. A., & Queiroz, A. B. M. (2005). Using a combined blocking procedure to teach color discrimination to a child with autism. *Journal of Applied Behavior Analysis, 38*, 555-558. doi: 10.1901/jaba.2005.65-04

Appendix

Parent/Caregiver Survey

Dear Parent/Caregiver:

The purpose of this survey is to assess your overall satisfaction with and acceptability of the study titled ***Discrimination Training in the Context of the Stimulus Equivalence Paradigm***.

Please complete the following questions based on your experience with this project.

	Strongly Agree 5	Agree 4	Not Sure 3	Disagree 2	Strongly Disagree 1
1. The researcher respected my child/dependent during the entire length of the study.					
2. I felt this study was a productive use of my child/dependent's time.					
3. I believe this procedure is likely to be effective.					
4. I would be willing to use this procedure to help establish praise as a reinforcer for my child/dependent.					
5. I would recommend others to participate in similar research.					

Thank you for taking the time to participate in this research and for providing your feedback!

Kind Regards,

Jeanine R. Tanz, Primary Investigator