# Generality as an Interaction between Recombinative Generalization and Derived Relational Responding

By

Jessica Lee Paranczak

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Special Education

May 10, 2024

Nashville, Tennessee

Approved:

Joseph M. Lambert, Ph.D., BCBA-D

Elizabeth E. Biggs, Ph.D.

Justin D. Lane, Ph.D., BCBA-D

Jennifer R. Ledford, Ph.D., BCBA-D

Johanna L. Staubitz, Ph.D., BCBA-D

*To my mother and father, whose unwavering belief in me,*
*taught me to believe in myself*

# ACKNOWLEDGMENTS

I am grateful for the financial support of Vanderbilt University's Peabody College Professional Development Funds and Grant #H325D190005, *Preparing Certified Behavior Analysts to be Leaders in Special Education* of the Office of Special Education Programs, U.S. Department of Education.

Further, I am grateful for the support and feedback from the members of my committee – Drs. Joe Lambert, Elizabeth Biggs, Justin Lane, Jennifer Ledford, and Johanna Staubitz. I am especially grateful for the guidance and encouragement from my advisor, Joe Lambert, throughout this entire project. Further, I am extremely grateful for the many contributions of the research assistants, who served as therapists during this study.

I would also like to thank my friends, near and far. To Haylee, Melissa and Alice for reminding me that I am a person first and a behavior analyst and researcher second. To Paige, Erin, Sarah and Sienna, this journey would have been much less fun without you by my side. To Janey and Bailey, I could not possibly thank you enough. You truly held me together over the past 5 years. Anything and everything I accomplish, is a testimate to you.

Most importantly I would like to thank my family: my mom, dad, Krista and Eric, for your love and support during this process. To my mom, who frequently said "if your teacher has a problem with you, tell them to call me" and my dad who paid me $20 for every yellow card I received in soccer, your actions taught me that when I was fighting for something, I was never alone. I could not have done this without that deeply rooted belief. To Krista and Eric, thank you for the endless encouragement and unconditional love. Whenever I needed a pep-talk or a distraction, you were there. I love you all so much.

# TABLE OF CONTENTS

# LIST OF TABLES

# LIST OF FIGURES

# CHAPTER 1

## Introduction

From a behavioral perspective, generality refers to a behavior change that (a) persists over time, (b) occurs in environments distinct from the one in which the initial behavior change occurred, and/or (c) spreads to a variety of related behaviors which were not directly addressed through the intervention (Baer, Wolf, and Risley, 1968). These outcomes, when they occur, can be the product of multiple distinct but interacting natural learning mechanisms (Stokes & Baer, 1977).

### Generalization

For example, response generalization is a term which describes circumstances in which novel responses, with physical similarities to previously learned behavior, occur in the presence of eliciting or discriminative stimuli (Kazdin, 1994; Mayer et al., 2011; Cooper et al., 2020 Skinner, 1953). For example, a child may be taught to cut paper by opening and closing their scissors, but when wrapping a birthday gift, they discover they can also slide their scissors to cut the paper. Similarly, in the presence of a flash of bright light in your eyes, one might tend to turn their head away from the light; one may also instinctively look down, close their eyes, or cover their eyes.

Relatedly, stimulus generalization describes situations in which novel stimuli elicit or evoke learned behavior because these stimuli share common properties with conditioned or discriminative stimuli which currently control behavior (Cooper et al., 2020). For example, novel smells might elicit favorable emotional responses if they resemble smells produced by a favorite meal. Similarly, a child might pick up and ride a random bike left on the street if it looks like the bike she rides at home.

In more complex cases of generalization, referred to as stimulus adduction (Johnson & Layng, 1992) or recombinative generalization (Goldstein & Mousetis, 1989; Suchowierska, 2006), independently acquired component responses come together to form a novel compound under a novel circumstance that possesses properties common to the conditions in which component responses were acquired and/or maintained (Johnston, 2014). For example, a child who has been taught to expressively label specific adjectives (e.g., jumping, rolling, sparkling) in one context and specific common nouns (e.g., water, sky, bean) in another, will sometimes be able to generate a novel expressive label (e.g., “jumping bean”) appropriate to a novel experience (e.g., seeing a Mexican jumping bean) that contains properties relevant to both a previously learned adjective (e.g., bouncing up and down) and a previously learned noun (e.g., a small brown roundish object).

In both simple and complex cases, generalization is a byproduct of Pavlovian (stimulus-stimulus) and/or operant (stimulus-response-stimulus) conditioning procedures. In complex cases, multiple exposures to conditioning are necessary. Specifically, recombinative generalization is only possible after an individual has already learned to discriminate component concepts. This is typically achieved following multiple reinforced exposures to distinct exemplars which share features critical to the relevant concept and which differ across many non-critical features.

## Teaching Nonarbitrary Relations via Multiple Exemplar Training

Because concepts are abstractions and represent only properties or elements of a given stimulus experience (Skinner, 1957), it is impossible to teach natural concepts by presenting only one example (Engelman & Carnine, 2016). That is, any given stimulus experience (e.g., seeing a blue toy train) which represents an example of a concept (e.g., "blue") possesses both critical (e.g., "blueness") and non-critical (e.g., "toy-trainness") features. Thus, to teach a concept, multiple exemplars are needed. For example, a parent can feel confident that a child understands the concept of *blue* when they can identify the property across any number of untrained items (e.g., blue train, blue desk, blue chair, blue box).

## Teaching Arbitrary Relations via Match to Sample

Unlike the relation between component members of a given concept (which share similarities across critical and discriminable features), the relation between component members and the symbols used to represent them in everyday discourse (e.g., vocal productions, text) is arbitrary. That is, these relations are dictated by social convention (e.g., English is spoken in the United States, Spanish is spoken in Mexico) and must be directly taught. One instructional procedure commonly employed to accomplish this for children with disabilities is sometimes referred to as "match to sample" (MTS; Pilgrim, 2020).

Broadly, MTS entails presenting a label (e.g., a vocal production of the word "blue" or the printed characters B-L-U-E) as a "sample" stimulus and then presenting an array of "comparison stimuli" or exemplars that include one example (e.g., a blue toy train) along with a number of near-non-examples which are the same across non-critical features but which differ across critical features (e.g., a red toy train, a yellow toy train), and then differentially reinforcing a correct selection response (e.g., pointing to the blue toy train). Instead of producing the label as the sample stimulus in subsequent trials, an instructor might produce an example of the target concept instead (e.g., a blue toy train) and then differentially reinforce a correct production response (e.g., a vocal production of the word "blue," the printed characters B-L-U-E). Because MTS is most effective when the procedure requires learners to make conditional discriminations (e.g., only point to a blue object when you hear "blue" [not when you hear "yellow" or "red"), concept labels (e.g., red, yellow, blue) are often taught in clusters. Stimuli that serve as non-examples (e.g., yellow trains and red trains) for one sample (e.g., "blue"), serve as examples for the others (e.g., yellow, red; Sidman & Tailby, 1982).

## Stimulus Equivalence

Importantly, not all novel performances and discriminations can be explained as byproducts of conditioning. For example, Sidman (1971) sought to teach a participant with a developmental delay how to read. Prior to instruction, this participant could select specific pictures (A stimuli) in the presence of corresponding spoken words (B stimuli). The participant could also produce correct spoken words (B) when presented with corresponding pictures (A). During instruction, Sidman employed a conditional-discrimination teaching protocol (i.e., MTS) to teach this participant to select appropriate textual stimuli (C stimuli) in the presence of corresponding spoken words (B). Following this instruction, the participant displayed several derived (untaught) performances, including a reversal of directly taught relations—referred to as *symmetrical responding* or *symmetry.* Specifically, following C-B instruction, the participant independently produced the appropriate spoken words (B) in the presence of corresponding textual stimuli (C).

Additionally, the participant could combine learning from instruction on more than one previously taught relation to respond appropriately under novel conditions—referred to as *transitive responding* or *transitivity*. Specifically, following A-B and C-B instruction, the participant was independently able to choose appropriate textual stimuli (C) in the presence of pictures (A), and vice versa (i.e., C-A and A-C responding). Because the derived relations between these stimuli led participants to respond to each similarly, Sidman referred to the overarching phenomenon (i.e., a manifestation of both symmetry and transitivity across a set of arbitrarily related stimuli) as *stimulus equivalence.* Relevant to generality, Sidman (1971) demonstrated that direct instruction on 20 conditional discriminations not only led to mastery of those specific targets, but also to the emergence of an additional 40 derived relations; thus demonstrating the generative potential of instructional paradigms which exploit equivalence relations (see *equivalence-based instruction*, below).

## Relational Framing

Since the groundbreaking work of Sidman (1971; 1994), it has become clear that humans are also capable of deriving relations other than equivalence (e.g., opposition, hierarchy) and that these relations can be as generative as the equivalence relation (Gibbs et al., 2023). It has also become clear that derived relational responding (i.e., a generalized pattern of behavior performed with respect to stimuli in one's environment that involves responding to at least one stimulus in terms of at least one other stimulus [Stewart et al. 2013]) can come under contextual control (Steele & Hayes, 1991). The generalized ability to respond to contextual stimuli is referred to as arbitrarily applicable relational responding (AARRing; Hayes et al., 2001).

To account for these realities, the concept of stimulus equivalence has been supplanted by a more general concept referred to as *relational framing* (Hayes et al., 2001). Relational frames are defined by the properties of mutual entailment, combinatorial entailment, and transformation of stimulus function. *Mutual entailment* serves as a more generic term for what would be called symmetry in stimulus equivalence, describing the bidirectionality of relational responding even when that bidirectionality is not symmetrical. For example, If A is larger than B, then B is smaller than A. *Combinatorial entailment* serves as a more generic term for transitivity and refers to derived responding, which emerges as a product of the combination of two or more directly trained relations. For example, if we learn that a nickel is more than a penny, and that dime is more than a nickel, these discriminations combine to yield a response of "yes" when asked, "is a penny is less than a dime?"

A *transformation of stimulus function* underscores the conditional nature of the manifestation of relational frames and is unique to the definition of relational framing in that there is not a correlate-defining feature in stimulus equivalence. For example, hearing "dinner" can serve as a discriminative stimulus ($S^D$) for coming to the table under some circumstances, and can serve as an $S^D$ for meal prep under others. Contextualizing variables determine which reaction is appropriate (e.g., wonderful smells wafting from the kitchen would select a "come-to-the-table" function. By contrast, an expectant and stern look might select a "start-preparing-dinner" function). When a given stimulus (e.g., "dinner") serves different functions under different circumstances (e.g., coming to the dinner table, meal prep), any one of these functions can be transferred to other members of a relational frame (e.g., a bell chime, a hand clap) such that the original contextual variables (e.g., wonderful smells, expectant looks) also select the functions of the arbitrarily related stimuli (bell chime, hand clap).

In contrast with generalization, which is a *byproduct* of conditioning, AARRing is purportedly a direct product of conditioning. That is, AARRing is a learnable skill and the mechanisms responsible for its occurrence are considerably different than those which promote generalization (Stewart et al., 2013). Notwithstanding, these mechanisms can interact to promote instances of novel responding not accounted for by either in isolation (Adams et al., 1993). For this reason, a technology of generality with aspirations of consistently promoting effective action under novel circumstances requires us to understand both the isolated, and combined, impacts of these mechanisms (Lambert et al., 2024).

## Generative Instruction

Broadly, generative instruction involves systematically arranging and teaching one set of skills while observing the emergence of other skills without direct teaching or a history of reinforcement (Alessi, 1987; Johnson & Layng, 1992). Through this process, previously acquired skills facilitate or accelerate the acquisition of other skills (e.g., Horne & Lowe, 1996). Generative instruction is instrumental when many skills must be taught, such as for individuals with language deficits (Axe & Sainato, 2010). Given the extent of skills that must be taught, generative instruction could preclude the necessity of directly teaching every skill individually, thus improving efficiency by reducing instructional time (Kemmerer et al., 2021). Recent research has outlined several instructional approaches for producing generative language, including matrix training and equivalence-based instruction (Axe, 2015).

### *Matrix Training*

Matrix training has been used for decades to establish untrained skills with individuals with autism and developmental disabilities (See Kemmerer et al., 2021, for a recent review). Matrix training is not a teaching procedure but a method of organizing learning targets. Multi-component skills, such as those required to make grammatically correct statements (e.g., noun and verb combinations, adjective and noun combinations, preposition and noun combinations), are arranged in a table (or matrix) with components of each skill isolated on each axis (e.g., nouns on one axis and verbs on the other). Within the cells of the matrix are the instructional targets, consisting of combinations of components outlined on each axis (e.g., each cell contains a noun-verb combination). Cells within the matrix are strategically selected for instruction, and the other combinations are not directly taught but instead assessed for recombinative generalization following mastery of the trained responses (Curiel et al., 2020; Kemmerer et al., 2021). Typical strategies to produce recombinative generalization call for training responses with overlapping components so that the participants can learn to emit novel responses made up of those components arranged in a new order (Hanna et al., 2004). Some studies suggest that the optimal targets for intervention are found along the diagonal of the matrix because these targets share no overlapping components but contain each component present in the matrix (Axe & Sainato, 2010; Kohler & Malott, 2014; Pauwels et al., 2015).

For example, Frampton and colleagues (2016) evaluated the use of diagonal matrix training on generative language in children with autism. Known nouns (e.g., duck, rabbit, pig) and known verbs (e.g., reading, painting, sitting) were arranged in a 3x3 matrix. Participants received instruction on noun-verb combinations for targets arranged along the diagonal (e.g., duck reading, rabbit painting, pig sitting). Following mastery, recombinative generalization was assessed for non-diagonal targets (e.g., duck painting, duck sitting, rabbit reading, rabbit sitting, pig reading, pig painting). Matrix training established recombinative generalization of known nouns and verbs for all participants. Further, Frampton et al. (2016) evaluated the extent to which the effects of recombinative generalization would extend to nouns and verbs that did not receive instruction (e.g., dog, bear, alligator, drinking, jumping, eating). Results indicate that all participants could emit noun-verb combinations for targets beyond the training matrix. Four participants could engage in this generalized response following one set of training targets. One participant required multiple exemplars (instruction in three matrices) before engaging in recombinative generalization with targets without matrix training.

## Equivalence-Based Instruction

Equivalence Based Instruction (EBI) is another approach to generative instruction. Like matrix training, EBI is not instruction per se but is a strategy for organizing instruction in ways that promote generative outcomes based on Sidman's theory of stimulus equivalence (1971). Specifically, the strategy entails first identifying how social convention equates arbitrary stimuli (e.g.., the sound "dog," the letters D-O-G, and four-legged furry animals) and then relating these stimuli (via differential reinforcement, often through MTS paradigms) in ways which establish bidirectional relations between them (e.g., by using the stimulus D-O-G as both the sample stimulus and as the comparison stimulus). When this is done, it is often the case that untrained relations between stimuli are derived in the absence of direct instruction. It is also often the case that the previous stimulus functions of each stimulus (e.g., emotional reactions to four-legged furry animals) are transferred to other stimuli related via MTS (e.g., the sound "dog") when the relations between them are derived.

EBI has been used to teach individuals with disabilities skills across a variety of domains, including language, academics (e.g., math, geography), and other functional skills (e.g., visual schedules; Pilgrim, 2020; McLay et al., 2013). For example, Elias and colleagues (2010) evaluated the effects of MTS instruction on the emergence of manual signs. Researchers taught seven individuals with disabilities relations between videos of manual signs (A), pictures (B), and printed words (C). Following MTS instruction for A-B and B-C relations, A-C relations emerged for 6 out of 7 participants. Researchers then evaluated additional relations (B-D, C-D) between pictures and (B) printed words (C) with participant-generated manual signs (D). Five participants were able to generate a corresponding manual sign (D) with no additional instruction. While promising, the effects of EBI on more complex language tasks, including grammar and pragmatics, are relatively unknown.

## Purpose

There is compelling (e.g., generalized equivalence classes; Adams et al., 1993) but limited evidence to suggest that relational framing can interact with complex generalization processes (e.g., recombinative generalization) to promote novel and contextually appropriate responding under novel circumstances. However, there are few examples of integrating established instructional methodologies (i.e., matrix training and EBI) to promote this type of response in children with disabilities (e.g., Neves et al., 2018). Because commercially available curriculums marketed to serve this population have scopes and sequences that assume the potential for this interaction (Dixon, 2014a, b, 2015, 2016), we sought to assess whether it occurs. Specifically, we were interested in demonstrating, that responding that has historically been attributed to generalization processes (recombinative generalization) can interact with responding attributed to derived relational responding (AARRing). We explored this interaction through a series of sequential research questions across four forms of generativity: (a) recombinative generalization, (b) AARRing for component stimuli, (c) AARRing for compound stimuli, and (d) AARRing in a socially valid context.

### ***Recombinative Generalization***

1. In a decontextualized discrete trial teaching (DTT) MTS paradigm and for participants who have recently learned ASL signs for component nouns and verbs, to what extent can they produce correct compound ASL responses (i.e., noun-verb signs; B stimuli) to novel compound-stimulus toy exemplars (e.g., a toy hippo sliding; A stimuli)?
2. In a decontextualized DTT MTS paradigm and for participants who do not automatically produce correct compound ASL responses (i.e., noun-verb signs; B stimuli) to novel compound stimulus toy exemplars (e.g., a toy hippo sliding; A stimuli) to what extent will matrix training lead to correct responding in response to novel compounds?

### ***Component AARRing***

3. In a decontextualized DTT MTS paradigm and for participants who can produce correct compound ASL responses (i.e., B stimuli) to novel compound stimulus toy exemplars (A stimuli), to what extent will additional instruction relating component stimulus situations (B stimuli) to arbitrary shapes (for nouns) or colors (for verbs) (C stimuli) lead to AARRing as evidenced by correct responding when tested on component C-A relations?

### ***Compound AARRing***

4. In a decontextualized DTT MTS paradigm and for participants who have derived component C-A relations, to what extent will they also derive *compound* C-A relations?

### ***Gameplay***

5. In a socially valid gameplay context (i.e., Twister), and following the emergence of derived compound C-A relations in a decontextualized format, to what extent will participants emit novel selection responses (i.e., correct gross motor actions) to novel compound shape-color compound exemplars (C stimuli), when novel compound novel compound-stimulus toy exemplars are presented as samples?

# CHAPTER 2

## Method

### Participants

The study was conducted with three siblings referred to a university-based behavior analysis clinic for treatment of severe problem behavior (i.e., Lambert, Copeland, et al., 2022). All study activities occurred within the constraints of each child's individualized behavior plans.

#### *Inclusion Criteria*

Although they were a population of convenience, had they not met the following inclusion criteria, additional participants would have been recruited. To be eligible to participate in the study, children must (a) have a documented intellectual or developmental disability (IDD), (b) be school-aged at the time they are enrolled in the study, (c) tact (i.e., expressively label) at least 100 objects and 20 actions in their native language, and (d) have parental consent to participate.

#### *Jerome*

Jerome was a 10-year-old Black male diagnosed with autism spectrum disorder (ASD) and profound Deafness. Jerome wore cochlear implants, and his primary mode of communication was American Sign Language (ASL). The Verbal Behavior-Milestones Assessment and Placement Program (VB-MAPP; Sundberg, 2008) was administered to Jerome prior to the study onset. Jerome's overall score was 83. His tact domain score was 8, and the listener responding domain score was 9. He did not show proficiency in the listener responding or tact domains for verb-noun combinations.

Jerome received behavior intervention for aggression and property destruction. Following a functional analysis, multiple functions of problem behavior were identified (Iwata et al., 1982/1994). Escape from demands and access to adult attention were confirmed, and access to tangibles was suggested (Lambert, Copeland, et al., 2022). At study onset, Jerome completed at least 15 demands over three activities in the absence of challenging behavior before accessing a break.

### ***Zeke***

Zeke was a 9-year-old Black male diagnosed with attention deficit disorder, a visual impairment (i.e., Cortical Visual Impairment and Nystagmus), a traumatic brain injury, and ASD. Zeke communicated vocally in complete sentences. Zeke received behavior intervention for tantrums. Following a functional analysis, multiple functions of problem behavior were confirmed, including escape from demands and access to attention and tangibles. At study onset, Zeke completed at least 11 demands in the absence of problem behavior before accessing reinforcement.

### ***Serenity***

Serenity was an 11-year-old biracial (i.e., White and Indigenous) child diagnosed with ASD and mood dysregulation. Serenity reported her gender as nonbinary, and she prefers she/her pronouns. Serenity communicated vocally in complete sentences. Serenity received behavior intervention for physical aggression, tantrums and verbal threats. Following a functional analysis, multiple functions of problem behavior were confirmed, including escape from demands and access to attention and tangibles. At study onset, Serenity completed at least 15 demands without problem behavior before accessing reinforcement.

Although both hearing siblings used vocal speech as their primary mode of communication, they each knew some basic signs to communicate with their brother. They were, however, not fluent ASL users. Specifically, neither participant used two-word combinations, including nouns and verbs (e.g., dog swimming).

### Setting

All appointments occurred in the family's home. Appointments were scheduled 3-6 days per week, lasting between 1.5 and 2 hours. All sessions were conducted with individual participants in a room away from the other participants (i.e., the participant's bedroom) to protect against diffusion of intervention effects.

### Implementers

Members of the research team served as therapists. The research team consisted of two White female doctoral students in special education and applied behavior analysis who were board-certified behavior analysts (BCBA) and graduate students seeking certification in behavior analysis and special education (e.g., Lambert, Paranczak et al., 2022). All study activities occurred under the direction of a White Latino doctoral-level BCBA (BCBA-D). Prior to the onset of the study, all members of the research team were trained in study-specific procedures, data collection, and crisis management (i.e., Safety Care®).

While members of the research team served as therapists, an ASL interpreter was present during two weekly appointments. The ASL interpreter was informed of study procedures and was instructed to provide no teaching or reinforcement during study sessions. We also consulted the ASL interpreter regarding all signs used throughout the study.

## Materials

General session materials included materials consistent with each child’s behavior intervention plan (i.e., visual schedule, delta signal) and tangible reinforcers (e.g., computer, LEGO). Researchers collected data on paper and pencil data sheets. Examples of data sheets are available in Appendix A.

Instructional materials included toy animal figurines and accessories (A-stimuli), which corresponded to targeted nouns (e.g., hippo, donkey) and verbs (e.g., slide for sliding), and pictures of different colored shapes (e.g., green star; C-stimuli).

Materials used to assess the interaction between recombinative generalization and relational framing through gameplay included A-stimuli (toys) and the Twister Shapes game (comprising a game board and spinner). The Twister Shapes game board featured a four-by-four grid with colored spots, each column representing a distinct color (red, green, yellow, blue). Each spot on the grid displayed a shape (star, triangle, square, circle), and every combination of color and shape was presented only once on the game board. Appendix B provides a picture of the game board for visual reference.

## Stimulus Sets

To facilitate matrix training, instructional targets were organized in a 3x3 matrix, with nouns listed across rows and verbs listed across columns. The training matrix is depicted in Figure 1.

**Figure 1:** *Training Matrix*

| | Sliding | Spinning | Seesawing |
|---|---|---|---|
| Hippo | **Hippo Sliding** | Hippo Spinning | Hippo Seesawing |
| Donkey | Donkey Sliding | **Donkey Spinning** | Donkey Seesawing |
| Hedgehog | Hedgehog Sliding | Hedgehog Spinning | **Hedgehog Seesawing** |

Note: Diagonal targets are presented in bold typeface. Diagonal targets were the only A-B compounds to receive instruction.

To facilitate equivalence-based instruction, we organized stimulus sets (rows) in terms of planned equivalence relations (columns) we aimed to establish between arbitrarily paired stimuli across sets over the course of this study. Compound stimuli relations are denoted by capital letters (e.g., A-B), component stimuli are represented with a subscript. A-stimuli were toy animals (noun, $A_N$) engaging with accessories (verb, $A_V$). B-stimuli were ASL signs corresponding to nouns ($B_N$) and verbs ($B_V$). C-stimuli were colored shapes. Noun components were depicted by shape ($C_N$), and verbs by color ($C_V$). Component stimulus sets are presented in Table 1 and Compound stimulus sets in Table 2.

**Table 1:** *Component Stimuli*

| Stimulus Set | Animal Toys ($A_N$) | Toy Accessories ($A_V$) | ASL ($B_N$, $B_V$) | Shapes ($C_N$) | Colors ($C_V$) |
|---|---|---|---|---|---|
| 1 | Hippo toy | - | Hippo ASL | Star | - |
| 2 | Donkey toy | - | Donkey ASL | Triangle | - |
| 3 | Hedgehog toy | - | Hedgehog ASL | Square | - |
| 4 | - | Sliding Slide | Sliding ASL | - | Green |
| 5 | - | Spinning Merry-Go-Round | Spinning ASL | - | Yellow |
| 6 | - | Seesawing Seesaw | Seesawing ASL | - | Red |

**Table 2***: Compound Stimuli*

| Stimulus Set | Animal + Accessories (A Stimuli) | ASL (B Stimuli) | Colored Shapes (C Stimuli) |
|---|---|---|---|
| 1 | Hippo + Slide Toys | Hippo Sliding ASL | Green Star |
| 2 | Donkey + Merry-Go- Round Toys | Donkey Spinning ASL | Yellow Triangle |
| 3 | Hedgehog + Seesaw Toys | Hedgehog Spinning ASL | Red Square |
| 4 | Hippo + Merry-Go- Round Toys | Hippo Spinning ASL | Yellow Star |
| 5 | Hippo + Seesaw Toys | Hippo Seesawing ASL | Red Star |
| 6 | Donkey + Slide Toys | Donkey Sliding ASL | Green Triangle |
| 7 | Donkey + Seesaw Toys | Donkey Seesawing ASL | Red Triangle |
| 8 | Hedgehog + Slide Toys | Hedgehog Sliding ASL | Green Square |
| 9 | Hedgehog + Merry-Go-Round Toys | Hedgehog Spinning ASL | Yellow Square |

Note. Shaded compounds received instruction during diagonal training.

## Response Measurement

Data were collected using paper and pencil on performance across individual components as well as corresponding compounds. Components were defined as individual nouns or verb concepts in isolation (e.g., the noun component for A-stimuli is toy hippo). Compounds included responses that included both a noun and verb concept across stimuli (e.g., hippo sliding, green star). The specific study phase determined which relations were targeted.

### *Recombinative Generalization*

The dependent variable to address research questions 1 and 2 was the percentage of correct responses to A-B compounds ($A_NA_V$-$B_NB_V$) during test sessions. A correct response entailed the participant engaging in B stimuli (ASL signs for noun-verb compound) when presented with toys performing actions (A-stimuli). An answer was considered correct when it included signs for both the noun and the verb, with no more than 3 seconds between them, and included no additional signs. Based on our consultation with the ASL interpreter, the order of noun and verb responses did not matter in the rules of ASL, so we accepted either combination (i.e., noun-verb or verb-noun). The percentage of correct responses was calculated by dividing the number of correct responses by the total number of trials within a given trial block (i.e., 9 trials).

Data were also collected during training sessions for A-B components (i.e., $A_N$-$B_N$, $A_V$-$B_V$). For A-B noun components ($A_N$-$B_N$), a correct response was scored when a participant engaged in an ASL sign (B stimuli) that corresponded with a toy animal (A-stimuli). During training sessions for A-B verb components ($A_V$-$B_V$), a correct response was scored when a participant engaged in an ASL sign (B stimuli) that corresponded with a toy accessory (A-stimuli). Training data were reported as trials to mastery. Mastery was defined as a participant answering at least 8/9 (88.89%) trials within a given trial block correctly for three consecutive trial blocks.

### *Component AARRing*

The dependent variable to address research questions 3, was the percentage of correct responding to C-A components ($C_N$-$A_N$, $C_V$-$A_V$) during test sessions. A correct response entailed participants selecting a toy from an array of 6 (3 animals, 3 accessories) that corresponded with a given shape (i.e., $C_N$) or color (i.e., $C_V$). The percentage of correct responses was calculated by dividing the number of correct responses by the total number of trials within a given trial block (i.e., 9 trials).

Training data were also collected on B-C components. A correct response was scored when a participant selected a specific comparison stimulus (among an array of three) in the presence of a sample stimulus that participated in the same relational network. Incorrect responses entailed pointing to a comparison stimulus that did not participate in the same relational network. Specifically, during training sessions for B-C noun components ($B_N$-$C_N$) a correct response was scored when a participant selected the correct shape (C-stimuli) when presented with the ASL sign for a noun (B-stimuli). Likewise, during training sessions for B-C verb components ($B_V$-$C_V$), a correct response was scored when a participant selected a color from an array of three (C stimuli) that corresponded with an ASL sign (B-stimuli). For all training data, the percentage of correct responses was calculated by dividing the number of correct responses by the number of correct and incorrect responses.

### ***Compound AARRing***

There were two dependent variables to evaluate compound AARRing: the emergence of B-C compounds and C-A compounds. During test probes for the emergence of B-C compounds, a correct response was scored when a participant selected a specific comparison stimulus (among an array of three) in the presence of a sample stimulus that participated in the same relational network. Incorrect responses entailed pointing to a comparison stimulus that did not participate in the same relational network. Specifically, a correct response entailed the participant selecting C stimuli (colored shape from an array of three) when presented with ASL for a noun-verb combination (B-stimuli). Comparison stimuli were presented in an array of three pictures of colored shapes. Incorrect comparison stimuli included one correct component (i.e., shape or color). For example, if the correct stimulus is a green star, the incorrect answer choices may have included a green triangle and a red star. All incorrect options were possible combinations of shapes and colors from the matrix; that is, each answer choice was a correct answer during other trials within the block. During test probes for the emergence of C-A compounds, a correct response entailed the participant producing A-stimuli (toys engaging with accessories) when presented with a colored shape (C-stimuli). The percentage of correct responses was calculated by dividing the number of correct responses by the number of correct and incorrect responses.

### *Gameplay*

Finally, researchers monitored participant responses during a gameplay context. During gameplay, the dependent variable was the percentage of correct responses for the A-C compound. A correct response was scored when a participant touched a colored shape (C-stimulus) on the Twister gameboard that corresponded to the toys the therapist showed (A-stimuli). The percentage of correct responses was calculated by dividing the number of correct responses by the number of correct and incorrect responses.

## Reliability

A second observer independently collected in situ data on all participants' performances across study phases. We calculated interobserver agreement using a point-by-point agreement method (Ledford & Gast, 2018). Specifically, we scored agreements when both primary and secondary observers recorded that the same participant response occurred during a given trial and disagreements when they did not. We then divided agreements by the sum of agreements and disagreements and multiplied by 100. Overall agreement across participants and study phases was collected during 40.3% of trial blocks, with an overall agreement of 99.46%. Interobserver agreement for Jerome was measured during 51.05% of trial blocks (243 out of 476) across study phases, maintaining an overall agreement rate of 99.5%. Mean agreement for Zeke was assessed in 45.79% of trial blocks (125 out of 273), with an overall agreement of 99.54%. Mean agreement for Serenity was assessed in 49.22% of trial blocks (65 out of 128), with an overall agreement of 99.10%. Interobserver Agreement across study phases and participants is presented in Table 3.

***Table 3:*** Interobserver Agreement across Study Phases

| | **Jerome** | **Zeke** | **Serenity** |
|---|---|---|---|
| **Recombinative Generalization** | | | |
| Component Training | 95.56% | 100.00% | 96.30% |
| Diagonal Training | 100.00% | 100.00% | N/A |
| Compound Test | 100.00% | 100.00% | 100.00% |
| Compound Remediation | 100.00% | 98.41% | N/A |
| **Component AARRing** | | | |
| Component Training | 99.47% | 99.44% | 97.22% |
| Component Test | 98.80% | 100.00% | 98.89% |
| Component Remediation | 99.57% | 100.00% | 100.00% |
| **Compound AARRing** | | | |
| Compound Test | 98.85% | 100.00% | 100.00% |
| Compound Remediation | 99.50% | 99.79% | 100.00% |
| **Gameplay** | | | |
| Pre-Test | 100.00% | 100.00% | 100.00% |
| Post-Test | 100.00% | 100.00% | 100.00% |

### Procedural Fidelity

We measured procedural fidelity using checklists highlighting critical steps for completing each session. Data collectors scored a “yes” each time a step on the checklist was correctly completed and a “no” when a step was incorrectly completed (or not completed at all) and generated session means by dividing “yes” by the sum of “yes” and “no.” Across participants and study phases, fidelity data were assessed for 42.63% of trial blocks, and mean fidelity was 99.31%. Fidelity data by study phase and participant are presented in Table 4.

***Table 4:*** Procedural Fidelity across Study Phases

| | **Jerome** | **Zeke** | **Serenity** |
|---|---|---|---|
| **Recombinative Generalization** | | | |
| Component Training | 100.00% | 94.10% | 92.59% |
| Diagonal Training | 100.00% | 100.00% | N/A |
| Compound Test | 98.22% | 99.29% | 98.57% |
| Compound Remediation | 100.00% | 97.96% | N/A |
| **Component AARRing** | | | |
| Component Training | 100.00% | 100.00% | 100.00% |
| Component Test | 100.00% | 100.00% | 100.00% |
| Component Remediation | 100.00% | 100.00% | 100.00% |
| **Compound AARRing** | | | |
| Compound Test | 99.62% | 100.00% | 100.00% |
| Compound Remediation | 99.48% | 99.86% | 96.48% |
| **Gameplay** | | | |
| Pre-test | 100.00% | 100.00% | 100.00% |
| Post-test | 100.00% | 100.00% | 100.00% |

## Experimental Design

Our research questions aimed to evaluate generality across two domains: generalization to non-targeted behaviors (i.e., generativity) and generalization across contexts (Lambert et al., 2024). Through a series of non-concurrent multiple baseline across participants designs (Watson & Workman, 1981), we evaluated the extent to which logically organized instruction results in the emergence of untaught relations (i.e., generativity).

A non-concurrent design was selected as participants had a long history of engaging in challenging behavior that could preclude the ability to engage in study procedures. It was important that subsequent tiers were not delayed due to challenging behavior in previous tiers. Participants were assigned to tier, based on severity of problem behavior and tolerance of demands. That is participants that were expected to tolerate (i.e., participate in the absence of challenging behavior) longer periods of baseline, were assigned to later tiers. Across designs, Jerome was assigned to tier 1, Zeke to tier 2, and Serenity to tier 3. The number of data points in each tier was predetermined. For the evaluation of compound relations, tier 1 was assigned 3 baseline data points, tier 2 was assigned 6 and tier 3 assigned 9. For component evaluations, tier 1 was assigned 3, tier 2 was assigned 5, and tier 7 was assigned 7.

Generalization across context was evaluated through gameplay using pre-and post-tests (i.e. research question 5). Study progression is depicted in Table 5.

**Table 5:** *General Study Progression*

| *Phases:* | **Recombinative Generalization** (Figure 2) | **Component AARRing** (Figure 4) | **Compound AARRing** (Figure 5) | **Gameplay** (Figure 6) |
|---|---|---|---|---|
| Gameplay Pre-test | | | | Test A-C |
| Compound Baseline | Test A-B | | Test B-C<br>Test C-A | |
| Component Baseline | Test $A_N$-$B_N$<br>Test $A_V$-$B_V$ | | | |
| Component Training (Table 9) | Train $A_N$-$B_N$<br>Train $A_V$-$B_V$ | | | |
| Compound Baseline | Test A-B | | | |
| Compound Diagonal Training | Train A-B Diag. | | | |
| Compound Baseline | Test A-B | | | |
| Remediation (Table 10) | Train $A_N$-$B_N$<br>Train $A_V$-$B_V$<br>Train A-B Diag.<br>Test A-B | | | |
| Component Baseline | | Test $C_N$-$A_N$<br>Test $C_V$-$A_V$<br>Test $B_N$-$C_N$<br>Test $B_V$-$C_V$ | | |
| Component Training (Table 6; Figure 3) | | Train $B_N$-$C_N$<br>Train $B_V$-$C_V$ | | |
| Component Baseline | | Test $C_N$-$A_N$<br>Test $C_V$-$A_V$ | | |
| Compound Baseline | | | Test B-C<br>Test C-A | |
| Remediation (Table 10) | | | Train $A_N$-$B_N$<br>Train $A_V$-$B_V$<br>Train A-B Diag.<br>Train $B_N$-$C_N$<br>Train $B_V$-$C_V$<br>Test A-B<br>Test B-C<br>Test C-A | |
| Gameplay Post-test | | | | Test C-A |

*Note.* Subscripts indicate noun ($X_N$) and verb components ($X_V$). $A_N$ = toy animals, $A_V$ = toy accessories, $B_N$= ASL animals, $B_V$ = ASL action, $C_N$ = shapes, $C_V$ = colors

## Procedures

Across phases of the study, participants had access to materials consistent with their individual behavior plans. Individual reinforcement schedules, outlined in each participant’s escape from demand treatment, were applied. Specifically, each participant had a signal that reinforcement (i.e., breaks from demands with highly preferred tangibles and access to therapist attention) was unavailable, and it was time to work. Jerome had a visual schedule displaying the activities' order programmed into each appointment. When an activity was completed, Jerome moved the visual depiction of the completed activity from a "to do" column to a "done" column. Following three study activities, Jerome received a break with access to his highly preferred tangibles and access to the therapist's attention. The signal for Zeke and Serenity was a star, indicating that breaks were unavailable. When the star was present, participants were taught to complete demands without challenging behavior. When the therapist removed the star, participants could ask for a break and have access to preferred tangibles and the therapist's attention.

Appointment activities were divided into (roughly) 5-min elements and into three general categories (i.e., instruction, gameplay, break). Both instruction and gameplay components were presented as "work" and were interspersed with "breaks." That is, after participants complete study ("work") sessions, they transitioned to 5-min of child-directed play (with access to highly preferred toys and technology) before returning to complete more study activities (Paranczak et al., 2024).

Each study activity was presented in a nine-trial block. The included targets in each block were dependent on the study phase (see below) and were randomized across trials. However, no more than two of the same targets would be presented consecutively. At the onset of each trial block, researchers presented two to three mastered skills to ensure attending behaviors and provide praise contingent on correct responses. The mastered skills consisted of motor imitation and one-step listener responding tasks. In instances where the participant did not respond correctly to mastered skills, therapists delayed the onset of study procedures and followed procedures outlined in individualized behavior plans to regain instructional control. Additionally, following the completion of an individual trial, the researcher either began the subsequent target trial or interspersed a mastered skill to mitigate the likelihood of extinguishing appropriate responses while delivering demands in a discrete trial format (Carbone et al., 2010). Work sessions (i.e., instruction and gameplay) included two types of trial blocks: train and test. Regardless of the study phase, training and test procedures remained constant.

### *Training Procedures*

During training trial blocks, phase-specific targets were taught using constant time delay (CTD; Ledford et al., 2019). During the first trial block, an immediate (i.e., 0 s delay) controlling prompt was provided for all trials. In subsequent trial blocks, participants were allowed 5 s to respond. Correct responses were reinforced with praise and/or access to preferred tangible items. Incorrect responses or failure to respond resulted in an error correction procedure (Frampton et al., 2016). During the error correction procedure, the researcher represented the trial, providing an immediate controlling prompt. If the participant responded to the prompt, another independent opportunity called a transfer trial was provided. In the transfer trial, the researcher presented the question in the same manner but did not prompt the response, allowing for an evaluation of the transfer of stimulus control from the prompt to the instructional cue. After a correct response to the transfer trial, researchers interspersed between one and three mastered skills. Finally, the researcher presented the target again for another opportunity for independent responding. Correct responses within the error correction procedure did not contribute to mastery criteria.

### *Test Procedures*

During test trial blocks, the participant was given 5 s to respond to each individual trial. No praise or feedback was provided contingent on responding. Neutral statements (e.g., "okay”) were delivered for both correct and incorrect responses.

## Recombinative Generalization

To evaluate if participants engaged in recombinative generalization, researchers evaluated participant responding to A-B compounds (a) prior to any instruction, (b) after component training, (c) after diagonal training, and (d) following remediation, if A-B compounds did not emerge, with previous instruction. Contingency reviews and examples of participant and researcher behavior during each phase for recombinative generalization are presented in Table 6.

### *Test: A-B Compounds*

Baseline data were first collected on A-B compounds. Standard test procedures were used; researchers did not provide reinforcement or feedback for any response throughout these sessions. Researchers presented a toy figurine corresponding to the target noun and modeled the target action (A-Stimuli). Participants were asked to respond in ASL (B-Stimuli). Correct answers included signs for both the corresponding noun and verb. Researchers returned to tests for A-B compounds when following (a) instruction in components, (b) diagonal training and (c) remediation.

### *Test A-B Components*

Baseline data were collected using standard test procedures for both noun ($A_N$-$B_N$) and verb components ($A_V$-$B_V$). In the absence of feedback and reinforcement, participants were presented with a toy animal figurine (noun) or an accessory (verb) and asked to identify the animal in ASL (noun) or identify what they do with the object in ASL (verb). If, during baseline, a participant responded to three trial blocks at 88.89% accuracy or greater, the component was considered mastered, and no additional instruction was provided. Training was subsequently provided for components in which participants did not respond with at least 88.89% accuracy.

**Table 6.** Recombinative Generalization Participant and Researcher Behavior

| | | | *Example* | |
|---|---|---|---|---|
| | Contingency Review | Stimulus Presentation | Participant Behavior | Researcher Behavior |
| **Test:** **$A_NA_V$-$B_NB_V$** | *I am going to show you some toys. Your job will be to tell me what's happening in ASL. This time, I can't tell you if you are right, so just try your best! Do you have any questions?* | Toy Donkey on a Seesaw | ASL Donkey + ASL Seesawing | Neutral Statement<br>No Feedback<br>No Reinforcement |
| **Test:** **$A_N$-$B_N$** | *I am going to show you a toy. Your job will be to tell me what it is in ASL. This time, I can't tell you if you are right, so just try your best! Do you have any questions?* | Toy Donkey | ASL Donkey | Neutral Statement<br>No Feedback<br>No Reinforcement |
| **Test:** **$A_V$-$B_V$** | *I am going to show you a toy. Your job will be to tell me what you do with it in ASL. This time, I can't tell you if you are right, so just try your best! Do you have any questions?* | Toy Seesaw | ASL Seesawing | Neutral Statement<br>No Feedback<br>No Reinforcement |
| **Train:** **$A_N$-$B_V$** | *I am going to show you a toy. Your job will be to tell me what it is in ASL. This time, I can help you, so if you don't know the answer, just wait and I'll show you! Do you have any questions?* | Toy Donkey | ASL Donkey | Constant Time Delay<br>Praise/Tangible Reinforcement<br>Error Correction |
| **Train:** **$A_V$-$B_V$** | *I am going to show you a toy. Your job will be to tell me what you do with it in ASL. This time, I can help you, so if you don't know the answer, just wait and I'll show you! Do you have any questions?* | Toy Seesaw | ASL Seesawing | Constant Time Delay<br>Praise/Tangible Reinforcement<br>Error Correction |
| **Train:** **$A_NA_V$-$B_NB_V$** **Diagonals** | *I am going to show you some toys. Your job will be to tell me what's happening in ASL. This time, I can help you, so if you don't know the answer, just wait and I'll show you! Do you have any questions?* | Toy Hippo on a Slide | ASL Hippo + ASL Sliding | Constant Time Delay<br>Praise/Tangible Reinforcement<br>Error Correction |

### ***Train A-B Components***

Standard training procedures, including CTD, were used to teach participants the relations between A stimuli (toys) and B Stimuli (ASL). Participants were first taught relations between the toys representing noun stimuli (i.e., $A_N$) and their corresponding sign in ASL (i.e., $B_N$). Instruction continued until participants responded correctly to 8 out of 9 (i.e., 88.89%) trials within three consecutive trial blocks. Following mastery of noun components, participants were similarly taught verb components. Researchers taught the relation between toys representing verb components (i.e., $A_V$) and their corresponding sign in ASL (i.e., $B_V$). Instruction continued until participants reached the mastery criterion of at least 88.89% correct responses over three consecutive trial blocks.

### ***Train: Diagonal Targets Procedures***

For participants whose A-B compounds did not emerge following component training, training was conducted for the diagonal targets in the matrix (See Figure 1). Participants received standard instruction for diagonal targets (i.e., hippo sliding, donkey spinning, hedgehog seesawing). That is, when presented with compound A-stimuli (e.g., toy hippo sliding), participants were prompted to engage in B-stimuli (e.g., ASL for hippo + sliding). No instruction was provided on non-diagonal targets within the matrix. Each trial block included three trials of each diagonal target. Diagonal training continued until the participant responded at mastery levels (i.e., three consecutive sessions at 88.89% or higher).

### *Test: A-B Compounds Procedures*

Following diagonal training, participants returned to a series of three test sessions involving all targets from within the matrix targets for A-B compounds. Procedures were identical to previous test conditions. Again, if, during these test probes, a participant responded to three trial blocks at 88.89% accuracy or greater, the compound for A-B relations was considered mastered. If participants did not reach mastery of the compound following component training, they proceeded to remediation.

### *A-B Compound Remediation Procedures*

During remediation, additional instruction was provided on previously mastered relations (i.e., $A_N$-$B_N$, $A_V$-$B_V$, A-B diagonals). Researchers presented trial blocks in the following sequence: (a) noun component training, (b) verb component training, and (c) diagonal training. A test session for A-B compounds followed each remediation sequence. Remediation continued until the participant responded to A-B compounds with at least 89.89% accuracy across three consecutive trials.

## AARRing Components

Following participants' demonstration of responding consistent with recombinative generalization (i.e., when they responded correctly to A-B relations that were not directly taught), additional relata were introduced to the stimulus network. That is, participants were taught noun and verb components for B-C relations. AARRing at the component level was evaluated by testing for the emergence of C-A component relations following instruction in B-C components. Contingency reviews and examples of participant and researcher behavior during each phase for AARRing components are presented in Table 7.

**Table 7.** Component AARRing Participant and Researcher Behavior

| | | | *Example* | |
|---|---|---|---|---|
| | Contingency Review | Stimulus Presentation | Participant Behavior | Researcher Behavior |
| **Test:**<br>**$C_N$-$A_N$** | *I am going to show you a shape. Your job will be to find the toy that goes with my shape. I can't tell you if you are right, so just try your best! Do you have any questions?* | Triangle | Selects<br>Toy Donkey | Neutral Statement<br>No Feedback<br>No Reinforcement |
| **Test:**<br>**$C_V$-$A_V$** | *I am going to show you a color. Your job will be to find the toy that goes with my color. I can't tell you if you are right, so just try your best! Do you have any questions?* | Red | Selects<br>Toy Seesaw | Neutral Statement<br>No Feedback<br>No Reinforcement |
| **Test:**<br>**$B_N$-$C_N$** | *I am going to show you a sign. Your job will be to find the shape that goes with my sign. I can't tell you if you are right, so just try your best! Do you have any questions?* | ASL Donkey | Touches triangle from an array of three (triangle, circle, square) | Neutral Statement<br>No Feedback<br>No Reinforcement |
| **Test:**<br>**$B_V$-$C_V$** | *I am going to show you a sign. Your job will be to find the color that goes with my sign. I can't tell you if you are right, so just try your best! Do you have any questions?* | ASL Seesawing | Touches red from an array of three (green, yellow, red) | Neutral Statement<br>No Feedback<br>No Reinforcement |
| **Train:**<br>**$B_N$-$C_N$** | *I am going to show you a sign. Your job will be to find the shape that goes with my sign. This time, I can help you, so if you don't know the answer, just wait and I'll show you! Do you have any questions?* | ASL Donkey | Touches triangle from an array of three (triangle, circle, square) | Constant Time Delay<br>Praise/Tangible Reinforcement<br>Error Correction |
| **Train:**<br>**$B_V$-$C_V$** | *I am going to show you some toys. Your job will be to tell me what's happening in ASL. This time, I can help you, so if you don't know the answer, just wait and I'll show you! Do you have any questions?* | ASL Seesawing | Touches red from an array of three (green, yellow, red) | Constant Time Delay<br>Praise/Tangible Reinforcement<br>Error Correction |

### *Test C-A Components*

Baseline data were collected using standard test procedures for both noun and verb components of the C-A relation. In the absence of feedback and reinforcement, participants were presented with a shape (noun) or color (verb) and asked to identify the corresponding toy animal (noun) or toy object (verb) from an array of 6 (all available noun and verb toys). Stimuli presentation for noun components is displayed in Appendix D and verb components in Appendix E.

### *Test B-C Components*

Baseline data were collected using standard test procedures for both noun and verb components of the B-C relation. In the absence of feedback and reinforcement, participants were presented with ASL of an animal (noun) or action (verb) and asked to identify the corresponding shape (noun) or color (verb). Stimuli presentation for noun components is displayed in Appendix F and verb components in Appendix G. Subsequently, training was provided in B-C relations for components in which participants did not respond with at least 88.89% accuracy.

### *Train B-C Components*

Standard training procedures, including CTD, were used to teach participants the relations between B stimuli (ASL) and C Stimuli (colors and shapes). Participants were first taught relations between the ASL sign representing noun stimuli (i.e., $B_N$) and their corresponding shape (i.e., $C_N$). Instruction continued until participants responded correctly to 8 out of 9 (i.e., 88.89%) trials within three consecutive trial blocks. Following mastery of noun components, participants were similarly taught verb components. Researchers taught the relation between ASL for verbs (i.e., $B_V$) and their corresponding color (i.e., $C_V$). Instruction continued until participants reached the mastery criterion of at least 88.89% correct responses over three consecutive trial blocks.

### *Test C-A Components*

Following mastery of B-C components ($B_N$-$C_N$, $B_V$-$C_V$), researchers tested for the emergence of C-A component relations ($C_N$-$A_N$, $C_V$-$A_V$). Procedures were identical to the baseline. If, during these test probes, a participant responded to three trial blocks at 88.89% accuracy or greater, the C-A components were considered mastered. compound for A-B relations was considered mastered. If C-A components did not emerge following component training, they received remediation.

### *C-A Component Remediation Procedures*

During remediation, additional instruction was provided on previously mastered steps. Researchers presented trial blocks in the following sequence: (a) noun component training ($B_N$-$C_N$), (b) verb component training ($B_V$-$C_v$). Each remediation sequence was followed by test sessions for C-A components ($C_N$-$A_N$, $C_V$-$A_V$). Remediation continued until the participant responded to both noun and verb C-A components with at least 89.89% accuracy across three consecutive trials.

## AARRing Compounds

Following evidence that participants were able to derive C-A components, the extent to which participants were also able to engage in AARRing at the compound level was evaluated. Specifically, researchers tested for the emergence of B-C and C-A compound relations. Data were collected prior to and following all component training. Contingency reviews and examples of participant and researcher behavior during each phase for AARRing compounds are presented in Table 8.

### *Test B-C Compounds*

Baseline data were collected on B-C compounds. Standard test procedures were used; researchers did not provide reinforcement or feedback for any response throughout these sessions. Researchers engaged in a compound sign in ASL (B-stimuli) and instructed participants to pick the colored shape that matched their sign from an array of three (See Appendix H).

### *Test C-A Compounds*

Baseline data were also collected for C-A (i.e., colored shapes to toys) relations for compound stimuli. Researchers provided access to all animal figures and accessories and asked participants to show the toys (animals engaging with accessories) that matched a given colored shape (See Appendix I). Correct answers included toys corresponding to the corresponding noun and verb represented by the shapes and colors, respectively.

**Table 8.** Compound AARRing Participant and Researcher Behavior

| | | | *Example* | |
|---|---|---|---|---|
| | Contingency Review | Stimulus Presentation | Participant Behavior | Researcher Behavior |
| **Test:** **$B_NB_V$-$C_NC_V$** | *I am going to show you something in ASL. Your job will be to find the secret code that goes with my sign. I can't tell you if you are right, so just try your best! Do you have any questions?* | ASL Hippo + ASL Spinning | Selects yellow star from an array of three (yellow star, yellow triangle and green star) | Neutral Statement<br>No Feedback<br>No Reinforcement |
| **Test:** **$C_NC_V$-$A_NA_V$** | *I am going to show you a secret code. Your job will be to find the toys that go with the secret code. I can't tell you if you are right, so just try your best! Do you have any questions?* | Green Star | Puts hippo on slide from an array of 3 animals and 3 accessories | Neutral Statement<br>No Feedback<br>No Reinforcement |

### *B-C and C-A Compound Remediation Procedures*

During remediation, additional instruction was provided on previously mastered phases. Researchers presented trial blocks in the following sequence: (a) noun component training ($A_N$-$B_N$), (b) verb component training ($A_V$-$B_v$), (c) diagonal training (A-B diagonals), (d) noun component training ($B_N$-$C_N$), (e) verb component training ($B_V$-$C_v$). Test sessions for A-B, B-C, and C-A compounds followed each remediation sequence. Remediation continued until the participant responded to both B-C and C-A compounds with at least 89.89% accuracy across three consecutive trials.

### *Gameplay*

Researchers evaluated the generalization in the context of a game of Twister Shapes. Researchers presented participants with a toy animal (e.g., hippo; $A_N$) going down the slide ($A_v$) and asked participants to place their hand (or foot) on the Twister spot that matches. Each of the nine combinations in the matrix (e.g., hippo sliding) was presented once per trial block. There were three gameplay test probes prior to any instruction in the study and three gameplay test probes following the derivation of all relations at the end of the study. Procedures were identical in pre-and post-tests. Contingency reviews and examples of participant and researcher behavior during each phase for gameplay are presented in Table 9.

|  |  | *Example* |  |  |
|---|---|---|---|---|
|  | Contingency Review | Stimulus Presentation | Participant Behavior | Researcher Behavior |
| **Test:** **$A_NA_V$-$C_NC_V$** | *We are going to play a silly game of Twister. I am going to show you toys, and your job is to put your hand on the spot that goes with my toys.* | Toy Hippo on Slide | Puts foot on green star space | Neutral Statement<br>No Feedback<br>No Reinforcement |

# CHAPTER 3

## Results

### Recombinative Generalization

All participants demonstrated evidence of recombinative generalization. Jerome and Zeke engaged in generative responding following diagonal training procedures, while Serenity engaged in recombinative generalization following component training alone.

#### *A-B Components*

Training sessions for A-B component mastery are presented in Table 9. During baseline for A-B components, no participants responded at mastery levels for $A_N$-$B_N$ relations. Following the introduction of training procedures, all participants engaged in increased correct responding. Jerome required 8 training blocks, Zeke required 12 training blocks, and Serenity required 5 training blocks to reach mastery levels of responding for $A_N$-$B_N$ relations. In baseline for $A_V$-$B_V$ relations, Jerome responded correctly to at least 88.89% of trials across three consecutive baseline sessions, so he did not receive instruction on $A_V$-$B_V$ relations. Zeke and Serenity did not display mastery of $A_V$-$B_V$ relations during baseline. After introducing instruction, Zeke mastered $A_V$-$B_V$ relations in 8 trial blocks and Serenity in 4 trial blocks.

**Table 9:** *Teaching sessions to component mastery*

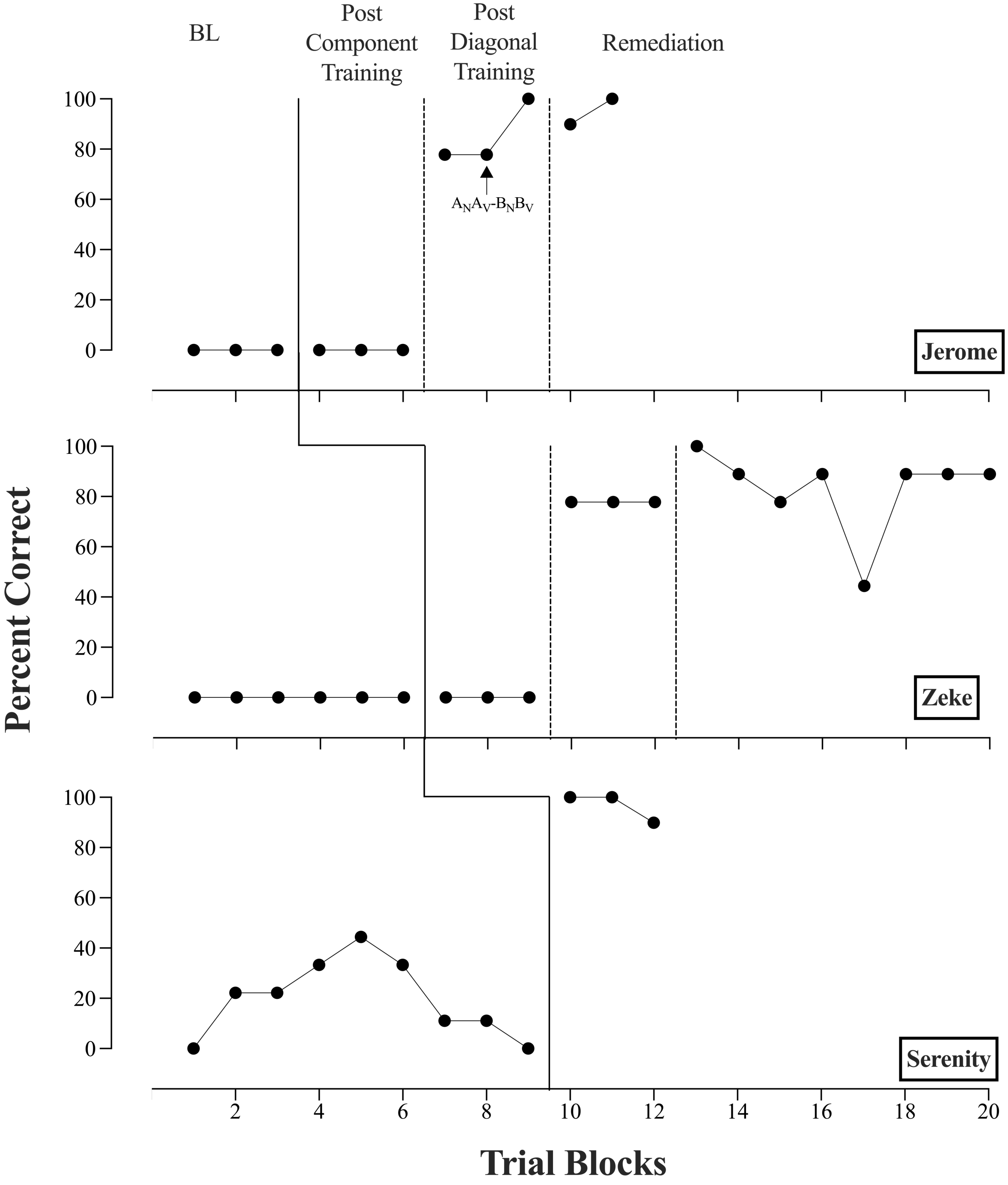
| Participant | $A_N$-$B_N$ | $A_V$-$B_V$ | $B_N$-$C_N$ | $B_V$-$C_V$ |
|---|---|---|---|---|
| Jerome | 8 | N/A | 43 | 30 |
| Zeke | 12 | 8 | 13 | 21 |
| Serenity | 5 | 4 | 4 | 4 |

Note. Jerome responded to $A_V$-$B_V$ correctly during baseline

### ***A-B Compound***

Figure 2 depicts results for A-B compounds across participants. During Baseline, no participants engaged in correct responding for A-B Compounds. Jerome and Zeke responded with 0% accuracy across baseline trials. Serenity engaged in variable but low responding across baseline trials (range: 0-44.44%). Following component training (i.e., $A_N$-$B_N$, $A_V$-$B_V$), Jerome and Zeke continued to respond correctly to 0% of trials. In contrast, Serenity's responding resulted in a distinct level change to mastery levels in the first three trials following component training, providing evidence of recombinative generalization. Following the introduction of diagonal training, Jerome and Zeke responding resulted in a clear level change. Still, they did not reach mastery levels of responding during the first three sessions following mastery of diagonal targets in training. Following a remediation procedure, Jerome and Zeke reached mastery levels. The number of training sessions required to reach mastery during remediation are presented in Table 10.

**Figure 2:** *Recombinative Generalization Results*



Note: BL = baseline

**Table 10:** *Remediation Trial Blocks to Mastery*

| | Jerome | Zeke | Serenity |
|---|---|---|---|
| A-B Compound | | | N/A |
| Train: $A_N$-$B_N$ | 2 | 8 | |
| Train: $A_V$-$B_V$ | 0 | 8 | |
| Train: $A^D$-$B^D$ | 2 | 8 | |
| C-A Component | | | |
| Train: $B_N$-$C_N$ | 45 | 1 | 1 |
| Train: $B_V$-$C_V$ | 10 | 1 | 1 |
| B-C Compound | | | N/A |
| Train: $A_N$-$B_N$ | 21 | 13 | |
| Train: $A_V$-$B_V$ | 21 | 13 | |
| Train: $A^D$-$B^D$ | 21 | 13 | |
| Train: $B_N$-$C_N$ | 25 | 13 | |
| Train: $B_V$-$C_V$ | 21 | 13 | |
| C-A compound | | | N/A |
| Train: $A_N$-$B_N$ | 21 | 11 | |
| Train: $A_V$-$B_V$ | 21 | 11 | |
| Train: $A^D$-$B^D$ | 21 | 11 | |
| Train: $B_N$-$C_N$ | 33 | 11 | |
| Train: $B_V$-$C_V$ | 27 | 11 | |

Note.

## AARRing Components

All participants demonstrated evidence of AARRing at the component level. All three participants required some level of remediation to reach mastery levels.

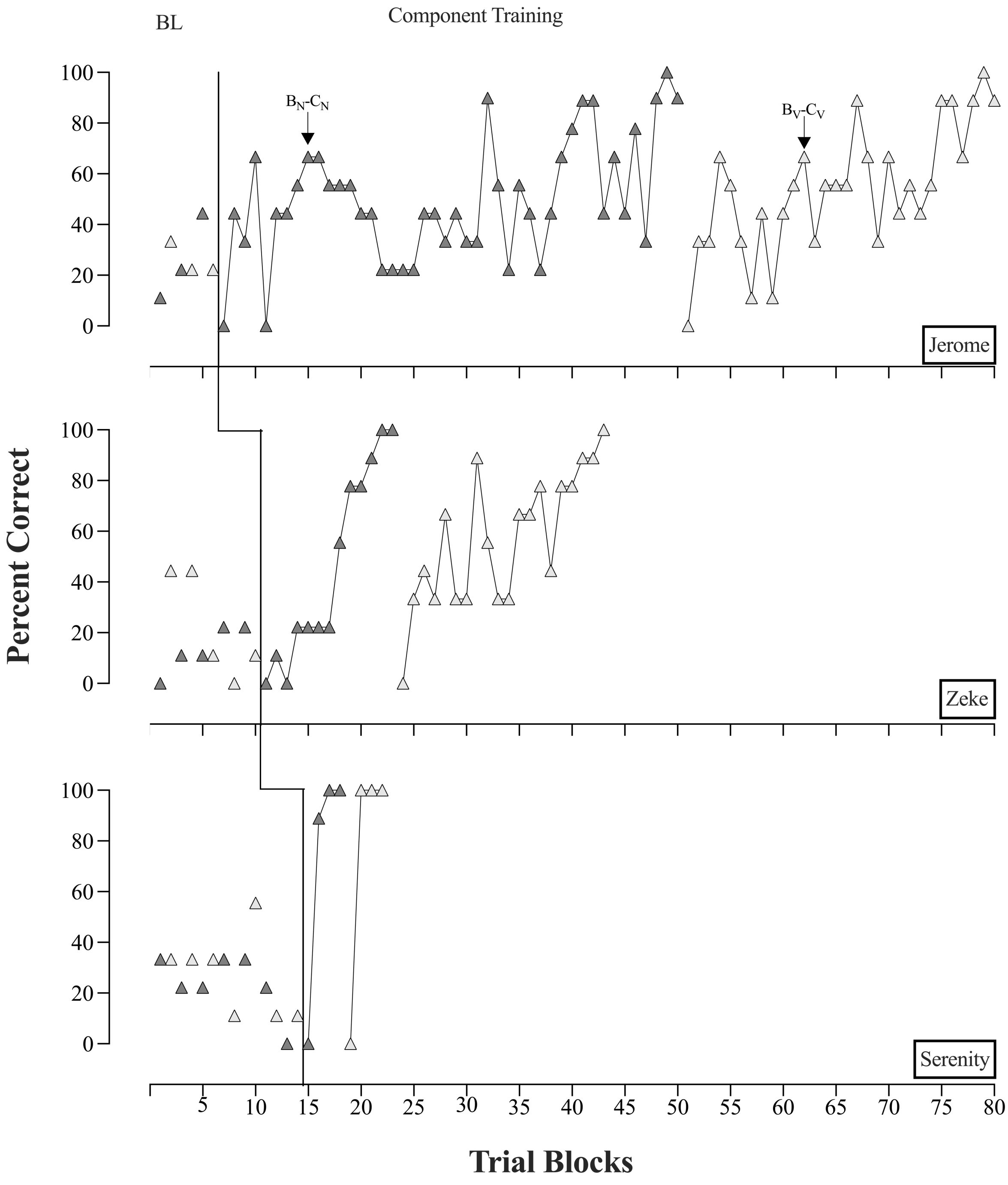
### *B-C Component*

Figure 3 depicts B-C component training data across participants, and Table 4 presents training sessions for mastery. During baseline for B-C components, no participants responded at mastery levels for $B_N$-$C_N$ relations (as expected, given these were arbitrarily assigned). Following the introduction of training procedures, all participants engaged in increased correct responses. Jerome required 43 training blocks, Zeke required 13 training blocks, and Serenity required 4 training blocks to reach mastery levels of responding for $B_N$-$C_N$ relations. Similarly, all participants engaged in low levels of correct responding during baseline for $B_V$-$C_V$ relations. Following the introduction of instruction, Jerome mastered $B_V$-$C_V$ relations in 43 trial blocks, Zeke in 13 trial blocks, and Serenity in 4 trial blocks

### *C-A Component*

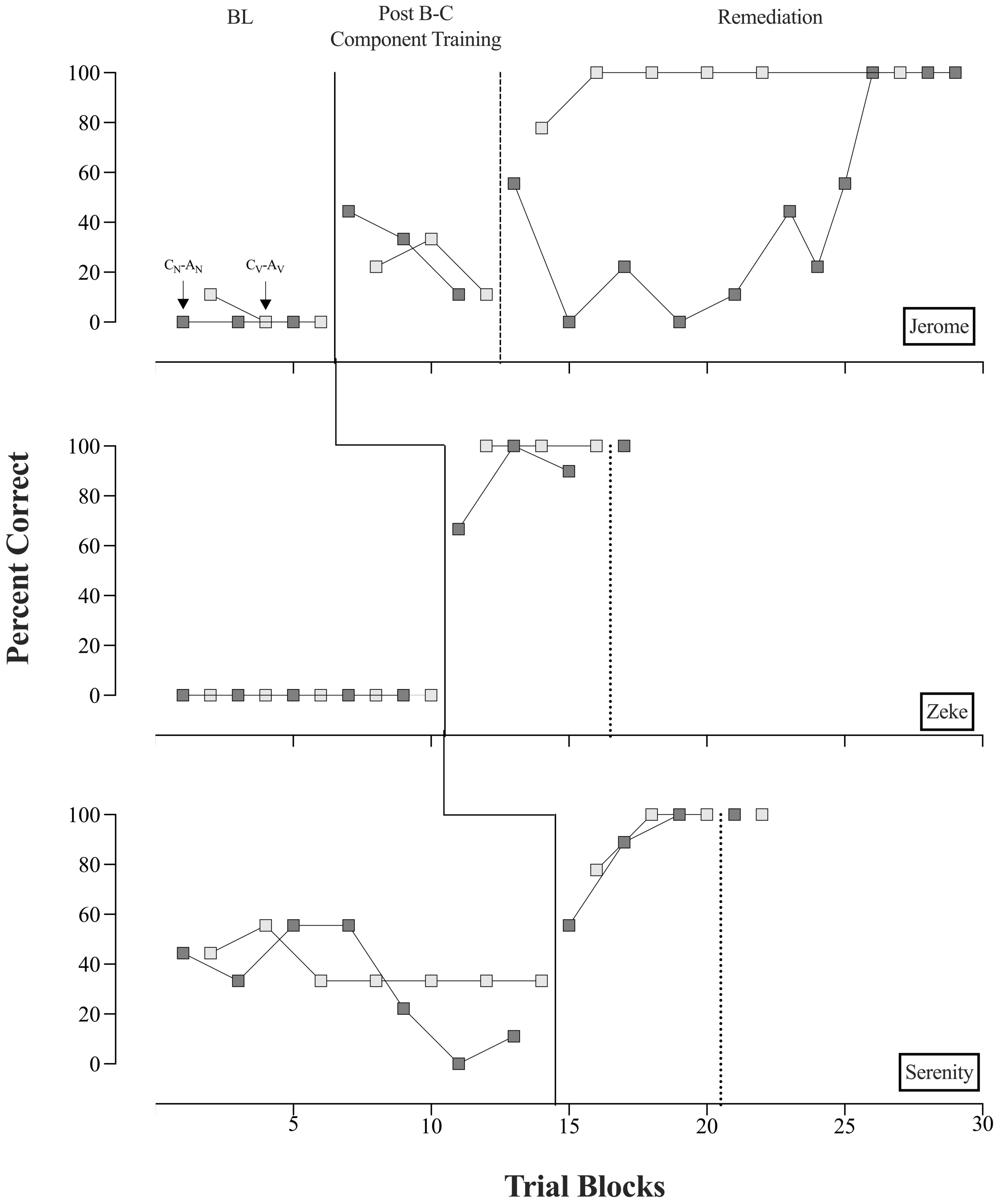
Figure 4 depicts results for the emergence of C-A components across participants. During baseline for C-A components, no participants engaged in responding at mastery levels for $C_N$-$A_N$ or $C_V$-$A_V$ relations. All participants engaged in increased correct responding following training in A-B and B-C components, but no participants reached mastery in the first three sessions following training. Following a remediation procedure, all participants derived C-A components. The number of training sessions required to reach mastery during remediation is presented in Table 10.

**Figure 3:** *B-C Component Training Results*



Note: BL = baseline

**Figure 4:** *AARRing Components Results*



Note: BL = baseline

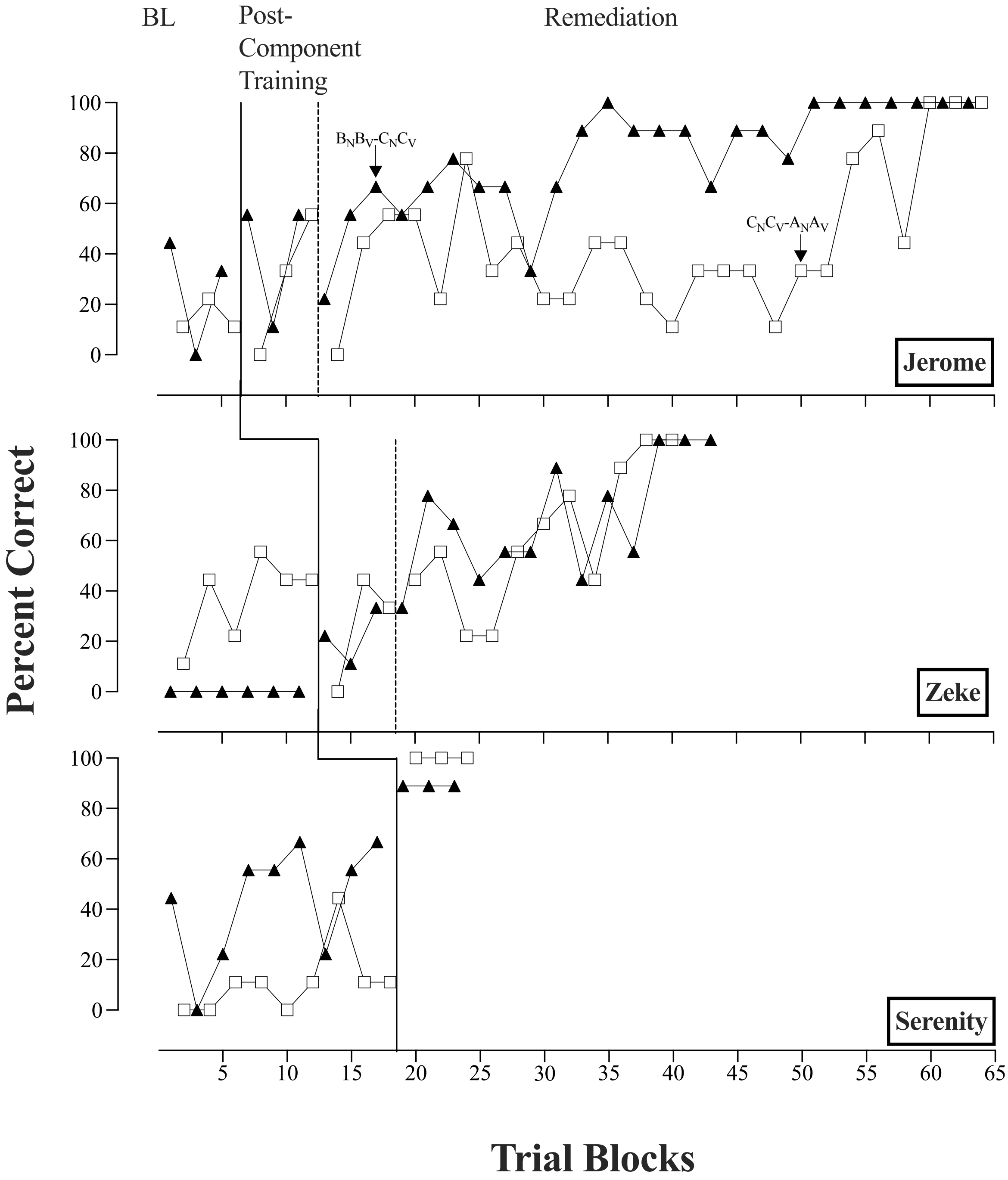
## AARRing Compounds

All participants demonstrated evidence of AARRing at the compound level. Jerome and Zeke required a remediation procedure to engage in AARRing for B-C and C-A compounds. While compound AARRing emerged for Serenity following emergence of C-A components.

### *B-C and C-A Compounds*

Figure 5 depicts results for B-C and C-A compounds across participants. During baseline, responding to B-C and C-A compounds, responding was at variable low levels. Following the emergence of C-A components, Jerome and Zeke continued to respond at levels similar to baseline, while Serenity's responding resulted in a distinct level change to mastery levels. Following a remediation procedure, Jerome and Zeke reached mastery levels. The number of training sessions required to reach mastery during remediation are presented in Table 10.

**Figure 5:** *AARRing Compounds*

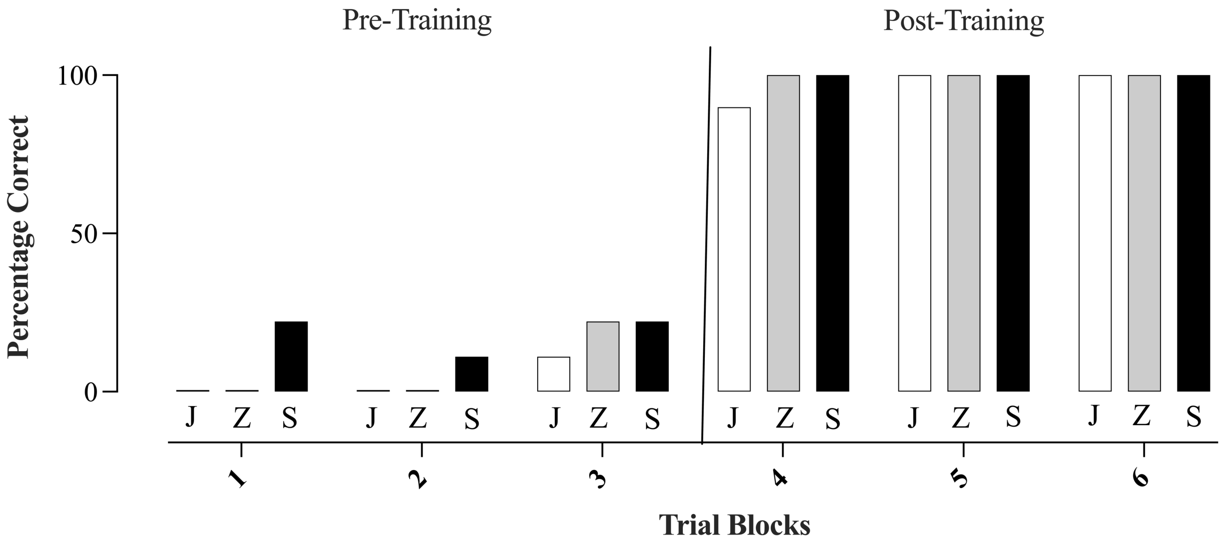


Note: BL = baseline

## Gameplay

Figure 6 shows the results from A-C Compounds in the gameplay context. All participants engaged in near-zero levels of correct responding during the gameplay baseline condition. After participants derived C-A Compounds in the instructional context, all participants engaged in consistent and high levels of responding in the post-test gameplay condition.

**Figure 6:** *Gameplay Results*



Note: J=Jerome, Z=Zeke, S=Serenity

# CHAPTER 4

## Discussion

Through this study, we sought to demonstrate how decontextualized and logically organized instruction would lead to derived and contextually appropriate recombinative generalization and AARRing in multiple contexts. We first sought to determine whether recombinative generalization would occur for noun-verb combinations following decontextualized discrete trial teaching of individual noun and verb components. For one participant, Serenity, this strategy was sufficient training for her to engage in A-B compound responses (i.e., using two signs to describe the action of a toy). For the other two participants, Jerome and Zeke, diagonal training was required for evidence of recombination generalization (Figure 3).

We then sought to determine whether derivative relations, AARRing at the component level (i.e., $C_N$-$A_N$, $C_V$-$A_V$) emerged following MTS instruction designed to reinforce coordination between (a) $A_N$-$B_N$, and $B_N$-$C_N$ and (b) $A_V$-$B_V$ and $B_V$-$C_V$ stimuli. For all participants, it did. There was, however, variance in how quickly these relations emerged. For Zeke and Serenity, these C-A component relations emerged quickly (Figure 4). Jerome required much more remediation (i.e., 45 additional trial blocks of $B_N$-$C_N$ training; Table 5). After Jerome mastered $B_N$-$C_N$ relations, there was an extended break (~2 weeks) for the winter holidays. Upon returning from break, Jerome no longer responded to this relation at mastery levels and required additional instruction.

Following the emergence of C-A components, we evaluated the extent to which AARRing at a compound level would emerge as evidenced by derivative B-C and C-A compounds responding. For Serenity, B-C and C-A Compounds emerged immediately following derivative responses for C-A components. Jerome and Zeke each required remediation to derive B-C and C-A compounds (Figure 5). Finally, Following the emergence of C-A compound relations in a decontextualized format, we evaluated to what extent A-C relations emerge in a gameplay context (Figure 6). All participants demonstrated generalized responding to the new context.

These results extend the existing matrix training (Kemmerer et al., 2021) and equivalence-based instruction (Tullis & Gibbs, 2022) literature bases and provide evidence that there may be utility in integrating components from matrix training and EBI. All participants displayed evidence of generative language typical of both instructional procedures. All participants responded to diagonal and non-diagonal targets within the matrix (A-B compounds), a common result following matrix training. Additionally, through supplemental MTS instruction rooted in the EBI tradition, they were also able to respond correctly to C-A components ($C_N$-$A_N$, $C_V$-$A_V$) and B-C and C-A Compounds.

The gameplay results also provide a demonstration that logically organized decontextualized instruction may lead to responding in generalized responding to alternate contexts. While lack of generalization is a common criticism of behaviorally based teaching methodologies (e.g., discrete trial teaching; Cowan and Allen 2007), this outcome adds to existing evidence that there can be socially valid benefits to decontextualized discrete-trial instruction (e.g., Paranczak et al., 2024; Paliliunas et al., 2022).

Researchers interested in generalized equivalence classes have long considered intersecting mechanisms responsible for generative responding (Adams et al., 1994).

Generalized equivalence classes often include both stimuli that share formal properties, as well as stimuli that are physically dissimilar (Adams et al., 1993). Through the mechanisms responsible for equivalence class formation, primary generalization and discrimination training, researchers have shown responses trained to one member of the equivalence class may transfer to the other stimuli in the class (e.g., Fields et al., 1996). While we included elements in line with these conceptualizations (e.g., game play performance with physically similar stimuli, related to physically dissimilar stimuli), we also programmed for higher-order generalization (i.e., recombinative generalization) within a given equivalence class.

While it is an empirical question whether the participants would have acquired the derived relations more efficiently if taught directly, there may be implications of the remediation method used in this study to discrete trial teaching. Through the remediation procedure, participants received repeated instruction in previously mastered skills to strengthen relations between stimuli. Despite a history of persistent and severe challenging behavior, participants engaged in no instances of problem behavior during additional training sessions provided through remediation. It is well-documented that reducing errors during instructional sessions results in lower rates of problem behavior (Carbone et al., 2010). It is possible that our remediation procedure served both to facilitate generative language, but also to abolish escape-motivated problem behavior. This could be impactful as patterns of persistent problem behavior often negatively impact educational experiences (Dworschak et al., 2016; Simó-Pinatella et al., 2019). For example, children who engage in externalizing problem behaviors often disrupt not only their own learning but also the learning of their peers (Watson et al., 2016). In turn, these students are more likely to have limited access to academic instruction (Carr et al., 1991). While not the purpose of this investigation, future research should consider the interaction between skill-acquisition programming and problem behavior reduction.

In summary, despite their shared history of engaging in challenging behavior, the study participants represented a heterogeneous group of children in terms of language ability and disability status. Our results demonstrate that despite the varying amount of instruction required, all of these children had the capacity for language generativity, as demonstrated by AARRing across multiple behaviors and contexts when provided with logically organized instruction through matrix training and EBI. While this study represents a small sample size, it adds to the initial evidence that organizing instruction promotes interaction between recombinative generalization and AARRing may have utility.

## Limitations

A couple of limitations should be noted. First, we did not conduct baseline assessments of each participant's AARRing ability before the study, which may limit the ability to replicate our findings. Second, our participants required different instructional procedures to engage in A-B compounds, so we did not have experimental control over responding. Neither Jerome nor Zeke combined nouns and verbs during probes for A-B compounds following instruction in components relations. However, Serenity correctly responded to A-B compounds following A-B component training alone. Jerome and Zeke's responding contribute to two demonstrations of effect for the use of diagonal training on recombinative generalization. A third limitation is the extent to which Jerome required remediation to derive relations. It is possible that we did not consider the correct mechanisms responsible for his performance. Anecdotally, Jerome engaged in more accurate responses when he supplemented researcher behavior with additional verbal behavior. For instance, during $B_N$-$C_N$ training, the researcher signed an animal in ASL, and participants were asked to match the sign to an array of three pictures. Jerome often copied the researcher's sign, engaged in the sign name for the shape, and then selected the shape from the array. Occasionally, Jerome would also point to the correct shape and state that the shape name matches the animal name. Since Jerome saw the most success when engaging in both expressive (i.e. tact) and receptive (i.e., listener responding) behavior, a more efficient remediation procedure may have incorporated procedures common to the naming tradition (Horne & Lowe, 1996) and included multiple exemplar instruction (i.e., rapidly rotating through verbal operants when providing instructions to learners; LaFrance & Tarbox, 2020).

## Conclusion

Despite limitations, this work holds value because it serves as a “proof of concept” by demonstrating that integrating components of matrix training and equivalence-based instruction can lead to derivative responding in multiple contexts (across relations and gameplay). For many individuals, generativity occurs naturally, such as in typical language development. That is, without direct instruction, humans often produce sentences they have never produced previously and understand sentences they have never heard before (Hayes et al., 2001). In contrast, learners with developmental and learning disabilities often have difficulty engaging in generative responses in the absence of systematic instruction (Suchowierska, 2006). Given that generativity is not guaranteed, understanding instructional methods to promote generativity is imperative. Future research should continue to examine methods of incorporating matrix training into ongoing, intensive instructional programming designed to promote language generativity for children with IDD. The efficacy of the integration of matrix training and EBI compared to generativity produced by other evidenced-based teaching strategies, or each alone, may be of particular interest to practitioners who are faced with clients with significant language deficits and finite resources. Critics of behavior analysis methodology (e.g., discrete trial teaching) have raised concerns about developing rote and inflexible language in children (Peterson et al., 2018); programming with language generativity in mind may serve as a practical antidote to this criticism.

# References

Adams, B. J., Fields, L., & Verhave, T. (1993). Formation of generalized equivalence classes: Stimulus equivalence. The Psychological Record, 43(4), 553–566. https://doi.org/10.1007/BF03395899

Alessi, G. (1987). The analysis of verbal behavior generative strategies and teaching for generalization. *The Analysis of Verbal Behavior*, *5*(1), 15–27. http://doi.org/10.1007/BF03392816

Axe, J. B. (2015). Combining concepts from verbal behavior and derived relational responding produces efficient language instruction for children with autism. *Evidence-Based Communication Assessment and Intervention*, *9*(3), 106-112. http://doi.org/10.1080/17489539.2016.1153813

Axe, J. B., & Sainato, D. M. (2010). Matrix training of preliteracy skills with preschoolers with autism. *Journal of Applied Behavior Analysis*, *43*(4), 635–652. https://doi.org/10.1901/jaba.2010.43-635

Baer, D. M., Wolf, M. M., & Risley, T. R. (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis*, *1*(1), 91–98. https://doi.org/10.1901/jaba.1968.1-91

Carbone, V. J., Morgenstern, B., Zecchin-Tirri, G., & Kolberg, L. (2010). The role of the reflexive-conditioned motivating operation (CMO-R) during discrete trial instruction of children with autism. *Focus on Autism and Other Developmental Disabilities*, *25*(2), 110–124. https://doi.org/10.1177/1088357610364393

Carr, E. G., Taylor, J. C., & Robinson, S. (1991). The effects of severe behavior problems in children on the teaching behavior of adults. Journal of Applied Behavior Analysis, 24(3), 523–535. https://doi.org/10.1901/jaba.1991.24-523

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

Cowan, R. J., & Allen, K. D. (2007). Using naturalistic procedures to enhance learning in individuals with autism: A focus on generalized teaching within the school setting. *Psychology in the Schools*, *44*(7), 701-715. https://doi.org/ 10.1002/pits.20259

Curiel, E. S. L., Axe, J. B., Sainato, D. M., & Goldstein, H. (2020). Systematic review of matrix training for individuals with autism spectrum disorder. F*ocus on Autism and Other Developmental Disabilities 35*(1), 55–64. https://doi.org/10.1177/1088357619881216

Dixon, M. R. (2014a). PEAK relational training system—Direct training module. Carbondale, IL: Shawnee Scientific Press.

Dixon, M. R. (2014b). PEAK relational training system—Generalization module. Carbondale, IL: Shawnee Scientific Press.

Dixon, M. R. (2015). PEAK relational training system—Equivalence module. Carbondale, IL: Shawnee Scientific Press.

Dixon, M. R. (2016). PEAK relational training system—Transformation module. Carbondale, IL: Shawnee Scientific Press.

Dworschak, W., Ratz, C., & Wagner, M. (2016). Prevalence and putative risk markers of challenging behavior in students with intellectual disabilities. *Research in Developmental Disabilities*, *58*, 94-103. https://doi.org/10.1016/j.ridd.2016.08.006

Fields, L., Adams, B. J., Buffington, D. M., Yang, W., & Verhave, T. (1996). Response transfer between stimuli in generalized equivalence classes: A model for the establishment of natural kind and fuzzy superordinate categories. *The Psychological Record*, *46*, 665-684. https://doi.org/10.1007/BF03395191

Paliliunas, D., Belisle, J., Barron, B. F., & Dixon, M. R. (2022). Transformation of hierarchical multiply controlled verbal relations in children with autism in a game of I Spy. *Behavior Analysis in Practice.*, *15*(4), 1383–1389. https://doi.org/10.1007/s40617-019-00391-0

Engelmann, S., & Carnine, D. (2016). *Theory of instruction: Principle sand applications (Rev. ed.).* NIFDI Press.

Frampton, S. E., Wymer, S. C., Hansen, B., & Shillingsburg, M. A. (2016). The use of matrix training to promote generative language with children with autism. *Journal of Applied Behavior Analysis*, *49*(4), 869–883. https://doi.org/10.1002/jaba.340

Gibbs, A. R., Tullis, C. A., Conine, D. E., & Fulton, A. A. (2023). A Systematic Review of Derived Relational Responding Beyond Coordination in Individuals with Autism and Intellectual and Developmental Disabilities. *Journal of Developmental and Physical Disabilities*, 1-36. https://doi.org/10.1007/s10882-023-09901-z

Goldstein, H., & Mousetis, L. (1989). Generalized language learning by children with severe mental retardation: Effects of peer' expressive modeling. *Journal of Applied Behavior Analysis*, *22*(3), 245–259. https://doi.org/10.1901/jaba.1989.22-245

Greer, R. D., & Longano, J. (2010). A rose by naming: How we may learn how to do it. *The Analysis of Verbal Behavior*, *26*(1), 73-106. http://doi.org/10.1007/BF03393085

Hanna, E. S., De Souza, D. G., De Rose, J. C., & Fonseca, M. (2004). Effects of delayed constructed-response identity matching on spelling of dictated words. *Journal of Applied Behavior Analysis, 37*(2), 223–227. https://doi.org/ 10.1901/jaba.2004.37-223

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

Horne, P. J., & Lowe, C. F. (1996). On the origins of naming and other symbolic behavior. *Journal of the Experimental Analysis of Behavior, 65*(1), 185–241. https://doi.org/10.1901/jeab.1996.65-185

Johnston, J. M. (2014). *Radical behaviorism for ABA practitioners*. Cornwall-on-Hudson, NY: Sloan Publishing

Johnson, K. R., & Layng, T. V. J. (1992). Breaking the structuralist barrier: Literacy and numeracy with fluency. Am*erican Psychologist, 47*(11), 1475–1490. https://doi.org/10.1037/0003-066X.47.11.1475

Kazdin, A. E. (1994). Behavior modification in applied settings (5th ed.). Brooks-Cole.

Kemmerer, A. R., Vladescu, J. C., Carrow, J. N., Sidener, T. M., & Deshais, M. A. (2021). A systematic review of the matrix training literature. *Behavioral Interventions 36*(2), 473–495. https://doi.org/10.1002/bin.1780

Kohler, K. T., & Malott, R. W. (2014). Matrix training and verbal generativity in children with autism. *The Analysis of Verbal Behavior, 30*(2), 170–177. https://doi.org/10.1007/s40616-014-0016-9

Lambert, J. M., Copeland, B. A., Paranczak, J. L., Macdonald, M. J., Torelli, J. N., & Houchins-Juarez, N. J. (2022). Description and evaluation of a function-informed and mechanisms-based framework for treating challenging behavior. *Journal of Applied Behavior Analysis, 55*(4), 1193–1219. https://doi.org/10.1002/jaba.940

Lambert, J. M., Meadan-Kaplansky, H, & Ledford, J. R. (2024). Measuring generality and social validity in single case research. In J. R. Ledford & D. L. Gast (Eds.), *Single case research methodology* (4th ed.)

Lambert, J. M., Paranczak, J. L., Copeland, B. A., Macdonald, M. J. M., & Houchins-Juarez, N. J. (2022). Exploring the validity of university-based practicum tailored to develop expertise in addressing challenging behavior. Journal of Applied Behavior Analysis. *55*(4), 1172-1192. https://doi.org/10.1002/jaba.941

LaFrance, D. L., & Tarbox, J. (2020). The importance of multiple exemplar instruction in the establishment of novel verbal behavior. *Journal of Applied Behavior Analysis*, *53*(1), 10-24.Click or tap here to enter text. http://doi.org/ 10.1002/jaba.611

Ledford, J., & Gast, D. (Eds.). (2018). Single case research methodology: Applications in special education and behavioral sciences (3rd ed.). Routledge.

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

Lovaas, I. O. (2003). Teaching individuals with developmental delays: Basic intervention techniques. PRO-ED.

Malott, R. W. (2003). Behavior analysis and linguistic productivity. *The Analysis of Verbal Behavior, 19*(1). 11-18. https://doi.org/10.1007/BF03392978

Mark L. Sundberg. (2008). VB-MAPP Verbal Behavior Milestones Assessment and Placement Program: a language and social skills assessment program for children with autism or other developmental disabilities: guide. AVB Press.

May, R. J., Stewart, I., Baez, L., Freegard, G., & Dymond, S. (2017). Arbitrarily applicable spatial relational responding. *Journal of the Experimental Analysis of Behavior, 107*(2). https://doi.org/10.1002/jeab.250

Mayer, G. R., Sulzer-Azaroff, B., & Wallace, M. (2011). Behavior analysis for lasting change. Sloan Educational Publishing.

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*(2). 418–431. https://doi.org/10.1016/j.rasd.2012.11.002

Miguel, C. F. (2016). Common and intraverbal bidirectional naming. *The Analysis of Verbal Behavior, 32*(2), 125–138. https://doi.org/10.1007/s40616-016-0066-2.

Miguel, C. F. (2018). Problem-solving, bidirectional naming, and the development of verbal repertoires. *Behavior Analysis Research and Practice, 18*(4), 340–353. https://doi.org/10.1037/bar0000110.

O'Hora, D., Barnes-Holmes, D., Roche, B., & Smeets, P. (2004). Derived relational networks and control by novel instructions: A possible model of generative verbal responding. *The Psychological Record, 54*(3), 437-460 http://doi.org/10.1007/BF33995484

Paranczak, J. L., Lambert, J. M., Ledford, J. R., Copeland, B. A., & Macdonald, M. J. (2024). Deriving relations at multiple levels of complexity following minimal instruction: A demonstration. *Journal of Applied Behavior Analysis,*1-18. https://doi.org/10.1002/jaba.1067

Pauwels, A. A., Ahearn, W. H., & Cohen, S. J. (2015). Recombinative generalization of tacts through matrix training with individuals with autism spectrum disorder. *The Analysis of Verbal Behavior, 31*(2), 200–214. https://doi.org/10.1007/s40616-015-0038-y

Peterson, S., Rodriguez, N., & Pawich, T. (2018). Effects of modeling rote versus varied responses on response variability and skill acquisition during discrete-trial instruction. *Journal of Applied Behavior Analysis, 52*(2), 370-385. https://doi.org/10.1002/jaba.528

Pilgrim, C. (2020). Equivalence-based instruction. In J. O. Cooper, T. E. Heron, & W. L. Heward (Eds.), Applied Behavior Analysis (3rd ed., pp. 452–496). Pearson Education.

Ribeiro, D. M., Elias, N. C., Goyos, C., & Miguel, C. F. (2010). The effects of listener training on the emergence of tact and mand signs by individuals with intellectual disabilities. *The Analysis of Verbal Behavior, 26*(1), 65–72. https://doi.org/10.1007/bf03393084

Sidman, M. (1971). Reading and auditory-visual equivalences. *Journal of Speech and Hearing Research*, 14(1), 5–13. https://doi.org/10.1044/jshr.1401.05

Sidman, M. (2000). Equivalence relations and the reinforcement contingency. Journal of the *Experimental Analysis of Behavior, 74*(1), 127-146. http://doi.org/10.1901/jeab.2000.74-127

Sidman, M., & Cresson, O. (1973). Reading and crossmodal transfer of stimulus equivalences in severe retardation. *American Journal of Mental Deficiency, 77*(5), 515–523.

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*(1), 5-22. https://doi.org/ 10.1901/jeab.1982.37-5

Simó-Pinatella, D., Mumbardó-Adam, C., Alomar-Kurz, E., Sugai, G., & Simonsen, B. (2019). Prevalence of challenging behaviors exhibited by children with disabilities: Mapping the literature. Journal of behavioral education, 28, 323-343. https://doi.org/ 10.1007/s10864-019-09326-9

Skinner, B. F. (1953). Science and human behavior. MacMillan.

Skinner, B. F. (1957). Verbal behavior. Appleton-Century-Crofts.

Steele, D. and Hayes, S. (1991). Stimulus equivalence and arbitrarily applicable relational responding. *Journal of the Experimental Analysis of Behavior, 56*(3), 519-555. https://doi.org/10.1901/jeab.1991.56-519

Stewart, I., McElwee, J., & Ming, S. (2013). Language generativity, response generalization, and derived relational responding. *The Analysis of Verbal Behavior, 29(1),* 137-155. https://doi.org/10.1007/BF03393131

Suchowierska, M. (2006). Recombinative generalization: Some theoretical and practical remarks. International Journal of Psychology, 41(6), 514–522. https://doi.org/10.1080/00207590500492534

Sundberg, M. L. (2008). VB-MAPP Verbal Behavior Milestones Assessment and Placement Program: a language and social skills assessment program for children with autism or other developmental disabilities: guide. Mark Sundberg.

Stokes, T. F., & Baer, D. M. (1977). An implicit technology of generalization. *Journal of Applied Behavior Analysis, 10*(2), 349–367. https://doi.org/10.1901/jaba.1977.10-349

Tullis, C. A., & Gibbs, A. R. (2022). Equivalence-Based Instruction for People with Autism Spectrum Disorder. In Handbook of Autism and Pervasive Developmental Disorder: Assessment, Diagnosis, and Treatment (pp. 919-935). Cham: Springer International Publishing.

Watson, T. L., Skinner, C. H., Skinner, A. L., Cazzell, S., Aspiranti, K. B., Moore, T., & Coleman, M. (2016). Preventing Disruptive Behavior via Classroom Management: Validating the Color Wheel System in Kindergarten Classrooms. *Behavior Modification, 40*(4), 518–540. https://doi.org/ doi/10.1177/0145445515626890

Watson, P. J., & Workman, E. A. (1981). The non-concurrent multiple baseline across-individuals design: An extension of the traditional multiple baseline design. *Journal of Behavior Therapy and Experimental Psychiatry*, 12(3), 257-259. https://doi.org/10.1016/0005-7916(81)90055-0

# Appendix A

## Example Data Sheets for Test Sessions for A-B, B-C, and C-A Compounds

## Probe: A-B, B-C, C-A

Participant:_________________________

| | |
|---|---|
| **Date:**_________<br>**Session #:**_______ | **Prim/Reli:** ______<br>**Therapist:**_______ |
| **Session Duration:** _____ **s**<br>**Latency to PB:** ________ **s** | **Intensity:** Below Average Above<br>**Duration:** Below Average Above |
| **A Stimuli (Action with Toys)** | **B Stimuli (ASL)** |
| Hippo Sliding | + - |
| Donkey Spinning | + - |
| Hedgehog Seesawing | + - |
| Donkey Seesawing | + - |
| Hippo Spinning | + - |
| Hedgehog Spinning | + - |
| Donkey Sliding | + - |
| Hedgehog Sliding | + - |
| Hippo Seesawing | + - |

| | | |
|---|---|---|
| **Date:**_________<br>**Session #:**_______ | **Prim/Reli:** ______<br>**Therapist:**_______ | |
| **Session Duration:** _____ **s**<br>**Latency to PB:** ________ **s** | **Intensity:** Below Average Above<br>**Duration:** Below Average Above | |
| **B Stimuli (ASL)** | **C Stimuli (Shapes)** | **Response** |
| Hippo Sliding | Green Star | + - |
| Donkey Spinning | Yellow Triangle | + - |
| Hedgehog Seesawing | Red Square | + - |
| Donkey Seesawing | Red Triangle | + - |
| Hippo Spinning | Yellow Star | + - |
| Hedgehog Spinning | Yellow Square | + - |
| Donkey Sliding | Green Triangle | + - |
| Hedgehog Sliding | Green Square | + - |
| Hippo Seesawing | Red Star | + - |

| | | |
|---|---|---|
| **Date:**_________<br>**Session #:**_______ | **Prim/Reli:** ______<br>**Therapist:**_______ | |
| **Session Duration:** _____ **s**<br>**Latency to PB:** ________ **s** | **Intensity:** Below Average Above<br>**Duration:** Below Average Above | |
| **C Stimuli (Shapes)** | **A Stimuli (Action with Toys)** | **Response** |
| Green Star | Hippo Sliding | + - |
| Yellow Triangle | Donkey Spinning | + - |
| Red Square | Hedgehog Seesawing | + - |
| Red Triangle | Donkey Seesawing | + - |
| Yellow Star | Hippo Spinning | + - |
| Yellow Square | Hedgehog Spinning | + - |
| Green Triangle | Donkey Sliding | + - |
| Green Square | Hedgehog Sliding | + - |
| Red Star | Hippo Seesawing | + - |

***A-B:*** We are going to play a signing game. I am going to show you some toys. Your job will be to tell me what's happening in ASL. This time, I can't tell you if you are right, so just try your best! Do you have any questions?

| | | |
|---|---|---|
| **Date:** | | |
| **Session #:** | **Prim Reli Initials:** | |
| Were the correct materials present? | YES | NO |
| Did the therapist provide a contingency review? | YES | NO |
| Did the therapist start the session with a signaled delta? | YES | NO |
| Did present 2-3 known tasks prior to study tasks? | YES | NO |
| Did the therapist provide reinforcement consistent with individual behavior plans? | YES | NO |
| | Yes | No |
| Did the therapist refrain from providing feedback for correct or incorrect responses? (Tally) | | |
| **Percent Fidelity**<br>Y/(Y+N)*100 | | |

***B-C:*** We are going to play another signing game. This time I will show you a sign, and I want you to pick the secret shape that goes with it. I can't tell you if you are right, so just try your best! Do you have any questions?

| | | |
|---|---|---|
| **Date:** | | |
| **Session #:** | **Prim Reli Initials:** | |
| Were the correct materials present? | YES | NO |
| Did the therapist provide a contingency review? | YES | NO |
| Did the therapist start the session with a signaled delta? | YES | NO |
| Did present 2-3 known tasks prior to study tasks? | YES | NO |
| Did the therapist provide reinforcement consistent with individual behavior plans? | YES | NO |
| | Yes | No |
| Did the therapist refrain from providing feedback for correct or incorrect responses? (Tally) | | |
| **Percent Fidelity**<br>Y/(Y+N)*100 | | |

***C-A:*** We are going to play another game with our secret shapes. I will show you a shape. Your job will be to show me that shape with the toys. This time, I can't tell you if you are right, so just try your best! Do you have any questions?

| | | |
|---|---|---|
| **Date:** | | |
| **Session #:** | **Prim Reli Initials:** | |
| Were the correct materials present? | YES | NO |
| Did the therapist provide a contingency review? | YES | NO |
| Did the therapist start the session with a signaled delta? | YES | NO |
| Did present 2-3 known tasks prior to study tasks? | YES | NO |
| Did the therapist provide reinforcement consistent with individual behavior plans? | YES | NO |
| | Yes | No |
| Did the therapist refrain from providing feedback for correct or incorrect responses? (Tally) | | |
| **Percent Fidelity**<br>Y/(Y+N)*100 | | |

# Appendix B

## Example Data Sheets for Training Sessions BN-CN (Noun Components)

## Train: $B_N – C_N$

Participant:_________________________

We are going to play a signing game. I am going to show you a sign. Your job will be to find the secret shape that matches with my sign. If you don't know, just wait and I will show you! Do you have any questions?

| | | | |
|---|---|---|---|
| **Date:__________**<br>**Session #:________** | | **Prim/Reli: _______**<br>**Therapist:________** | |
| **Session Duration: _____ s**<br>**Latency to PB: ________ s** | | **Intensity: Below Average Above**<br>**Duration: Below Average Above** | |
| **Slide** | **Sign** | **Shape** | **Response** |
| 1 | Hippo | Star | + - |
| 2 | Donkey | Triangle | + - |
| 3 | Hedgehog | Square | + - |
| 4 | Donkey | Triangle | + - |
| 5 | Hedgehog | Square | + - |
| 6 | Hippo | Star | + - |
| 7 | Hippo | Star | + - |
| 8 | Hedgehog | Square | + - |
| 9 | Donkey | Triangle | + - |

| | | | |
|---|---|---|---|
| **Date:** | | | |
| **Session #:** | **Prim Reli Initials:** | | |
| Were the correct materials present? | YES | NO | |
| Did the therapist provide a contingency review? | YES | NO | |
| Did the therapist start the session with a signaled delta? | YES | NO | |
| Did present 2-3 known tasks prior to study tasks? | YES | NO | |
| Did the therapist provide reinforcement consistent with individual behavior plans? | YES | NO | |
| | Yes | No | N/A |
| Did the therapist provide a prompt at the correct delay?<br>*Delay:* 0 s 5 s | | | |
| Did the therapist provide immediate reinforcement for correct responses? | | | |
| Did the therapist engage in the error correction procedure for incorrect responses? | | | |
| **Percent Fidelity**<br>Y/(Y+N)*100 | | | |

| | | | |
|---|---|---|---|
| **Date:__________**<br>**Session #:________** | | **Prim/Reli: _______**<br>**Therapist:________** | |
| **Session Duration: _____ s**<br>**Latency to PB: ________ s** | | **Intensity: Below Average Above**<br>**Duration: Below Average Above** | |
| **Slide** | **Sign** | **Shape** | **Response** |
| 1 | Hippo | Star | + - |
| 2 | Donkey | Triangle | + - |
| 3 | Hedgehog | Square | + - |
| 4 | Donkey | Triangle | + - |
| 5 | Hedgehog | Square | + - |
| 6 | Hippo | Star | + - |
| 7 | Hippo | Star | + - |
| 8 | Hedgehog | Square | + - |
| 9 | Donkey | Triangle | + - |

| | | | |
|---|---|---|---|
| **Date:** | | | |
| **Session #:** | **Prim Reli Initials:** | | |
| Were the correct materials present? | YES | NO | |
| Did the therapist provide a contingency review? | YES | NO | |
| Did the therapist start the session with a signaled delta? | YES | NO | |
| Did present 2-3 known tasks prior to study tasks? | YES | NO | |
| Did the therapist provide reinforcement consistent with individual behavior plans? | YES | NO | |
| | Yes | No | N/A |
| Did the therapist provide a prompt at the correct delay?<br>*Delay:* 0 s 5 s | | | |
| Did the therapist provide immediate reinforcement for correct responses? | | | |
| Did the therapist engage in the error correction procedure for incorrect responses? | | | |
| **Percent Fidelity**<br>Y/(Y+N)*100 | | | |

| | | | |
|---|---|---|---|
| **Date:__________**<br>**Session #:________** | | **Prim/Reli: _______**<br>**Therapist:________** | |
| **Session Duration: _____ s**<br>**Latency to PB: ________ s** | | **Intensity: Below Average Above**<br>**Duration: Below Average Above** | |
| **Slide** | **Sign** | **Shape** | **Response** |
| 1 | Hippo | Star | + - |
| 2 | Donkey | Triangle | + - |
| 3 | Hedgehog | Square | + - |
| 4 | Donkey | Triangle | + - |
| 5 | Hedgehog | Square | + - |
| 6 | Hippo | Star | + - |
| 7 | Hippo | Star | + - |
| 8 | Hedgehog | Square | + - |
| 9 | Donkey | Triangle | + - |

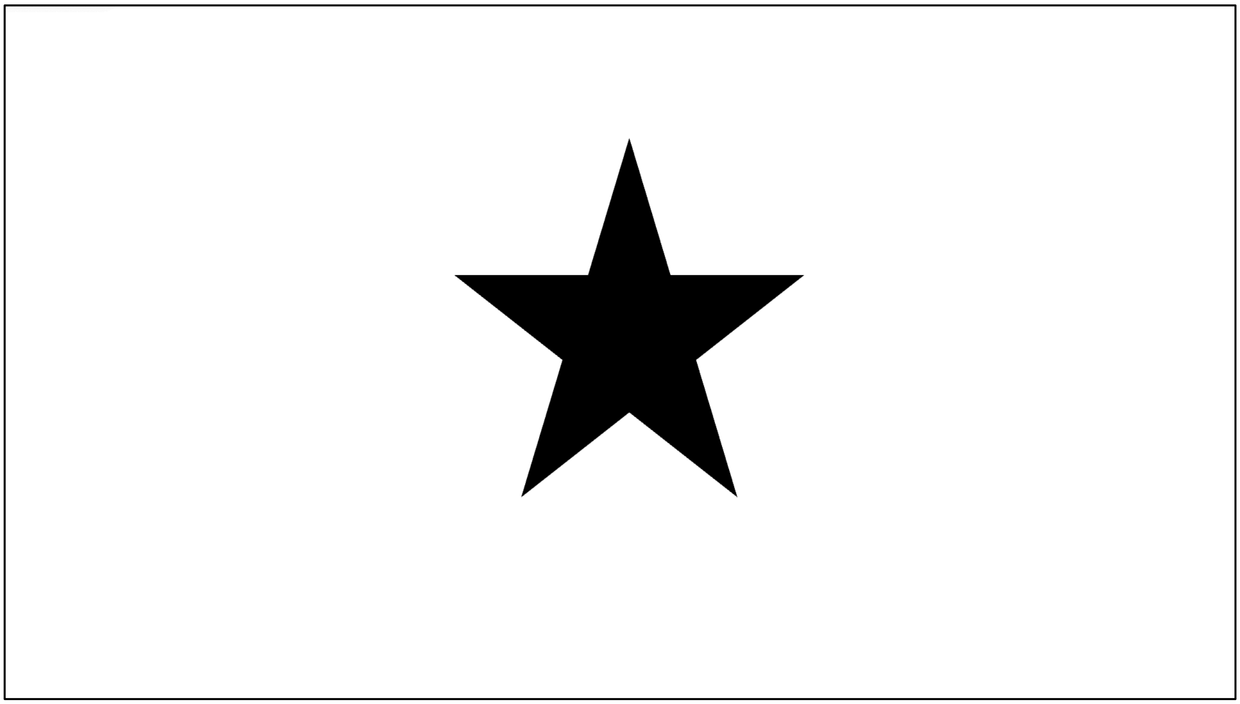
| | | | |
|---|---|---|---|
| **Date:** | | | |
| **Session #:** | **Prim Reli Initials:** | | |
| Were the correct materials present? | YES | NO | |
| Did the therapist provide a contingency review? | YES | NO | |
| Did the therapist start the session with a signaled delta? | YES | NO | |
| Did present 2-3 known tasks prior to study tasks? | YES | NO | |
| Did the therapist provide reinforcement consistent with individual behavior plans? | YES | NO | |
| | Yes | No | N/A |
| Did the therapist provide a prompt at the correct delay?<br>*Delay:* 0 s 5 s | | | |
| Did the therapist provide immediate reinforcement for correct responses? | | | |
| Did the therapist engage in the error correction procedure for incorrect responses? | | | |
| **Percent Fidelity**<br>Y/(Y+N)*100 | | | |

# Appendix C

Gameplay Materials: Twister Shapes Gameboard

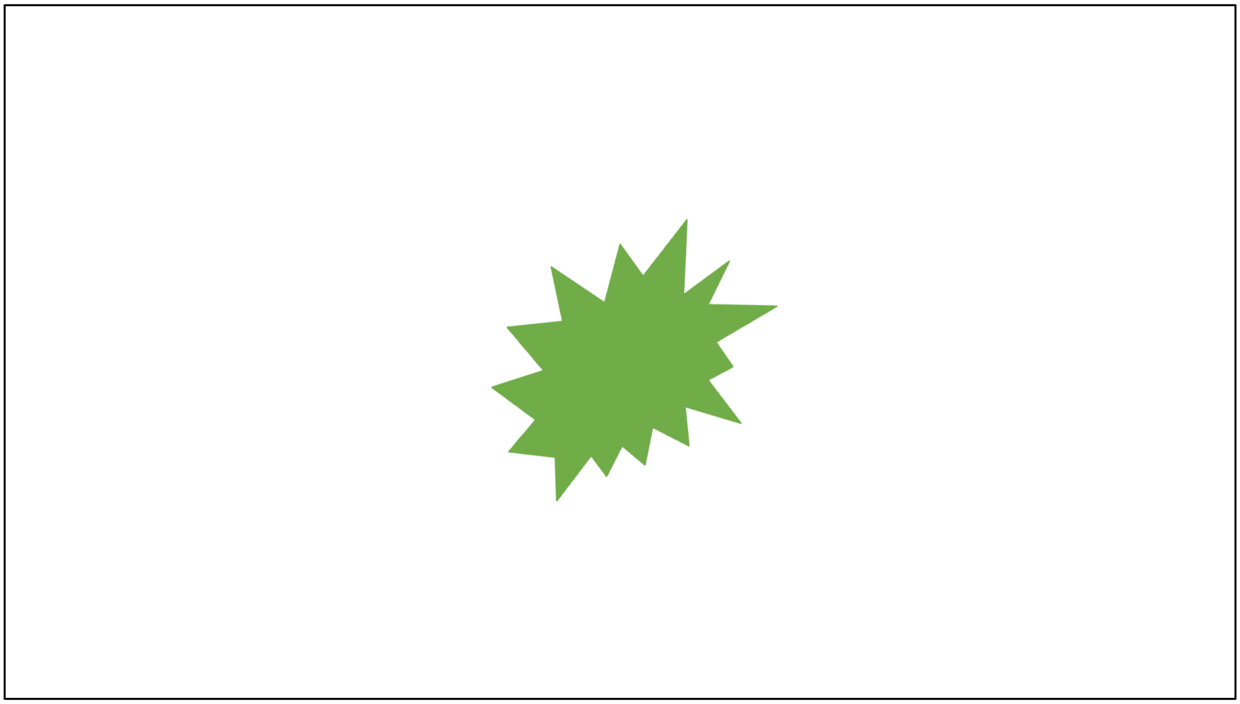
# Appendix D:

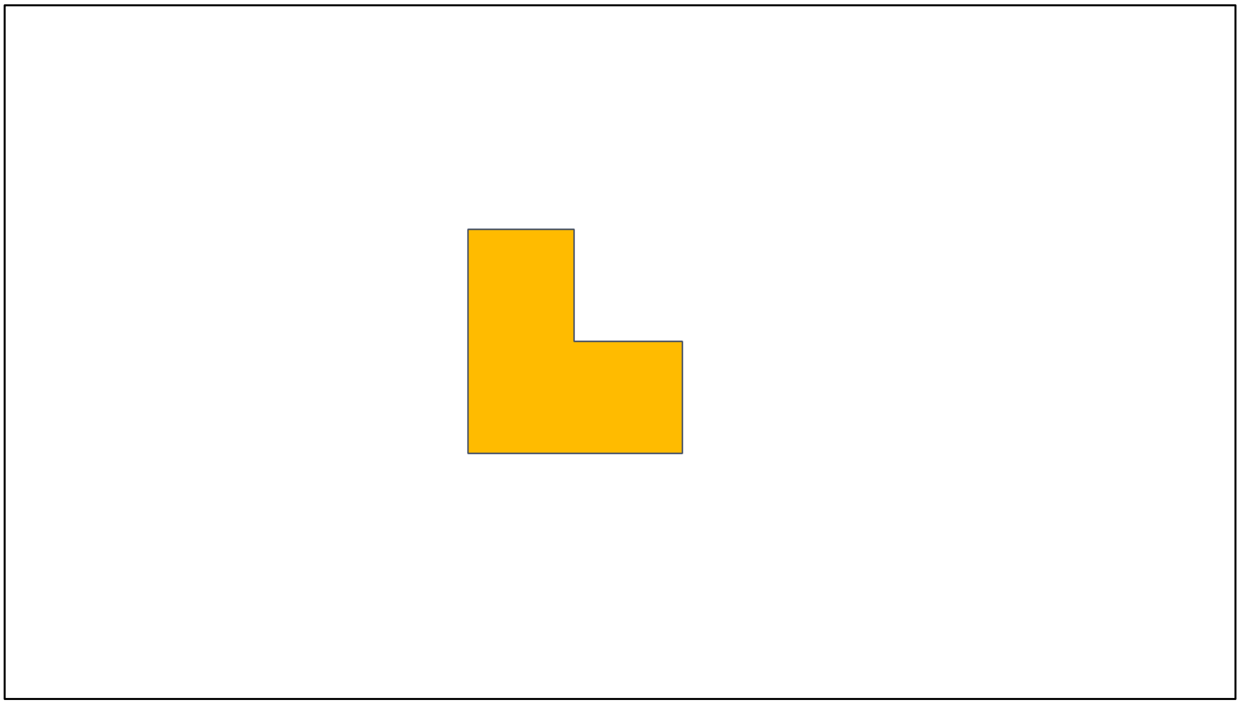
Stimulus Presentation: $C_N$-$A_N$



# Appendix E

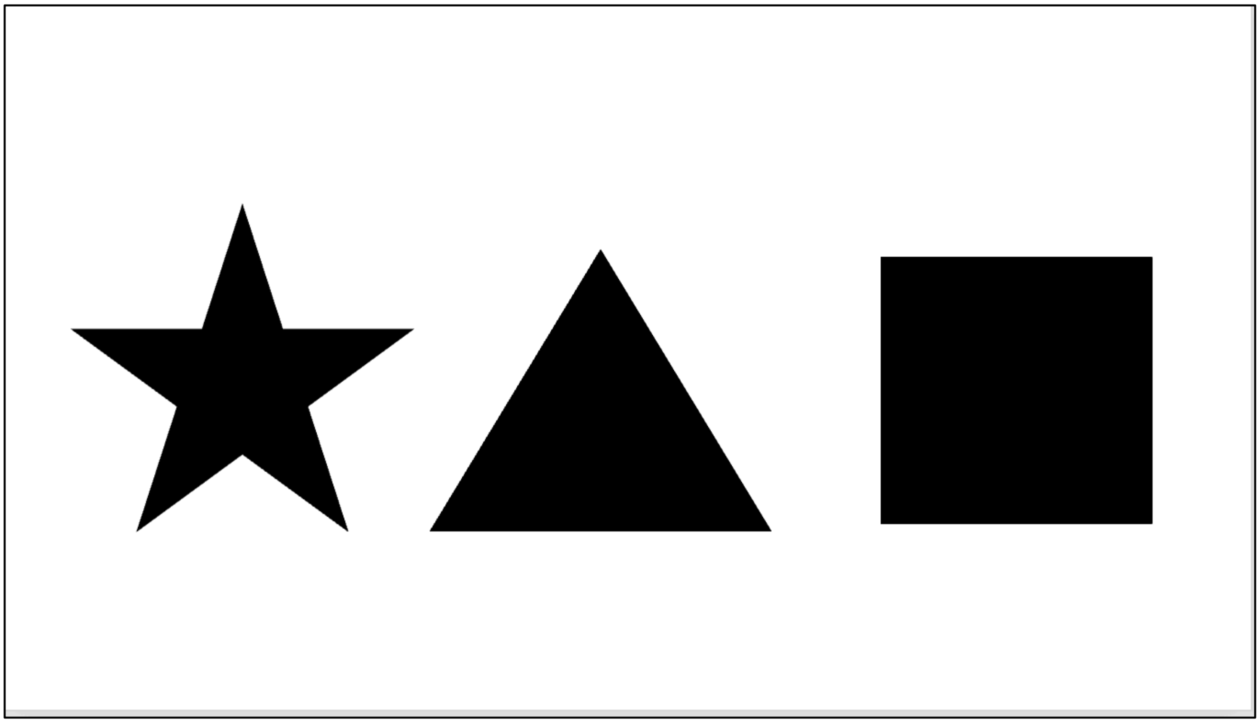
Stimulus Presentation: $C_V$-$A_V$





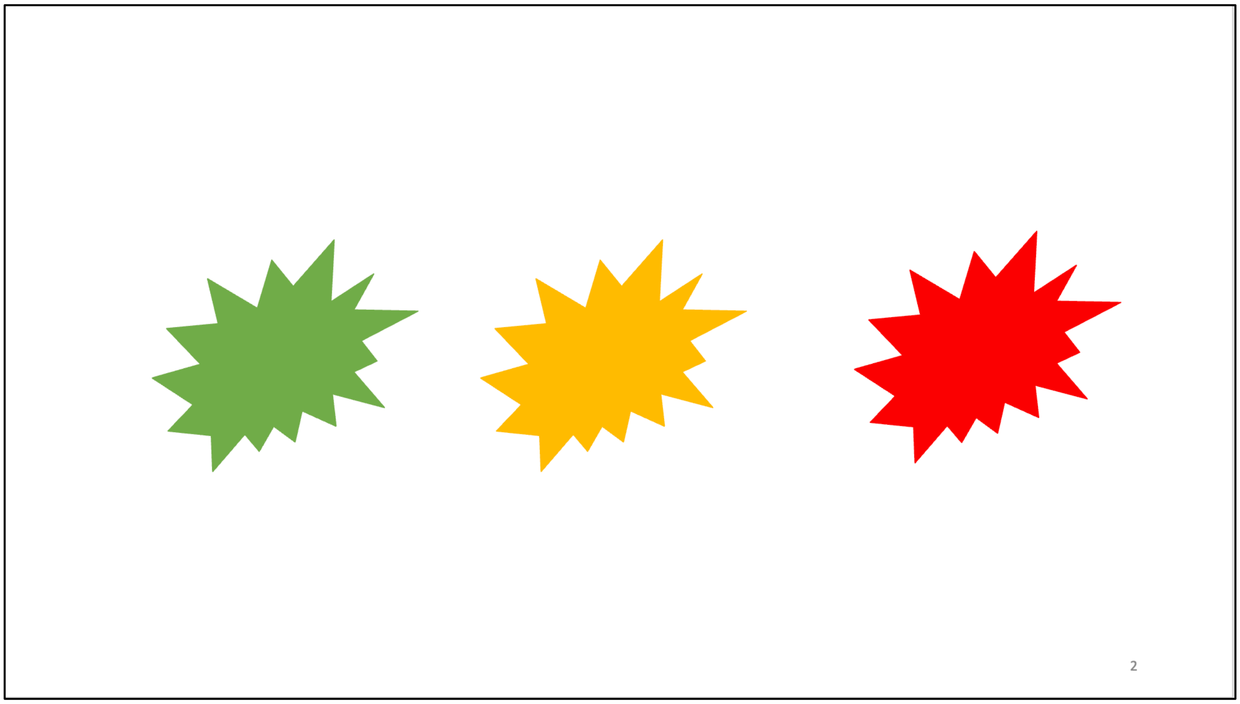
# Appendix F

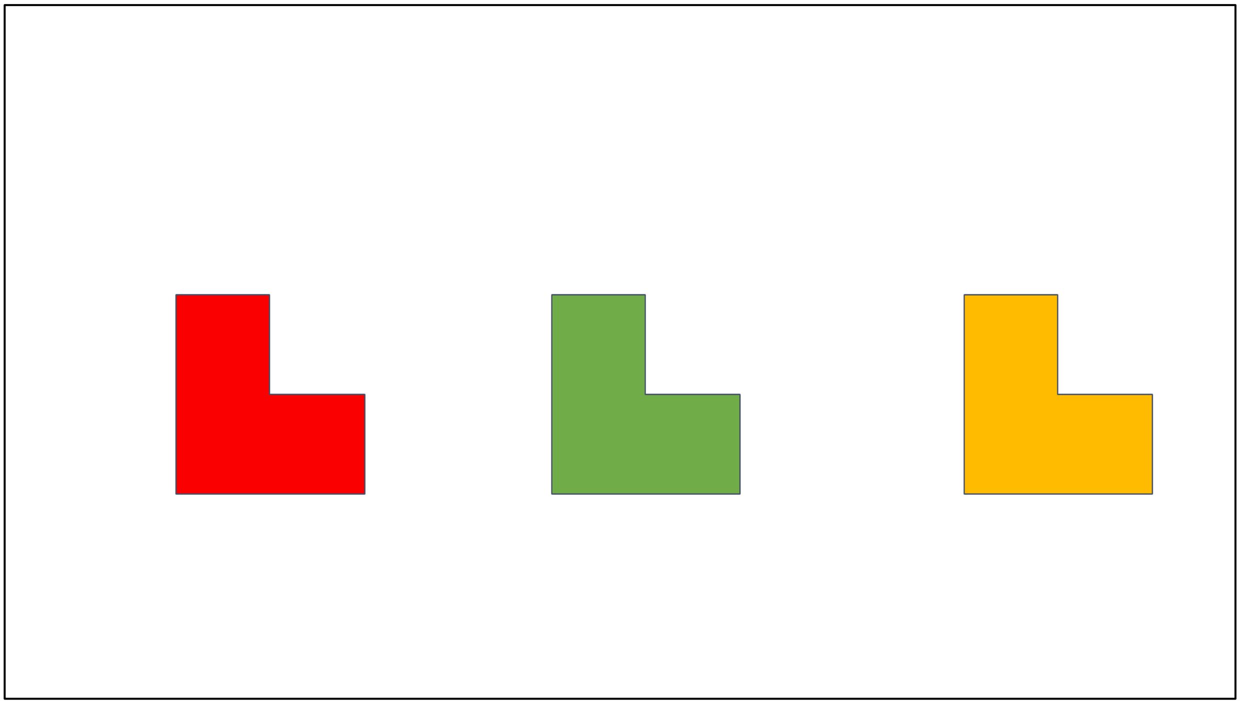
Stimulus Presentation: $B_N$-$C_N$



# Appendix G

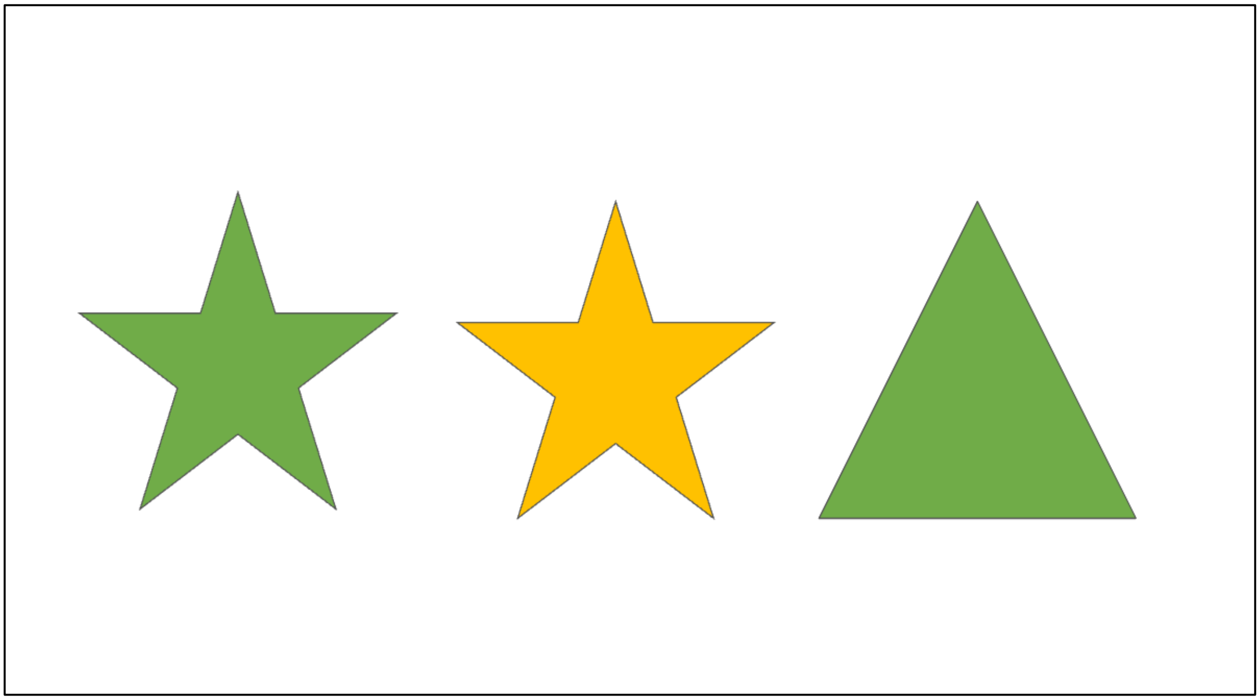
Stimulus Presentation: $B_V$-$C_V$

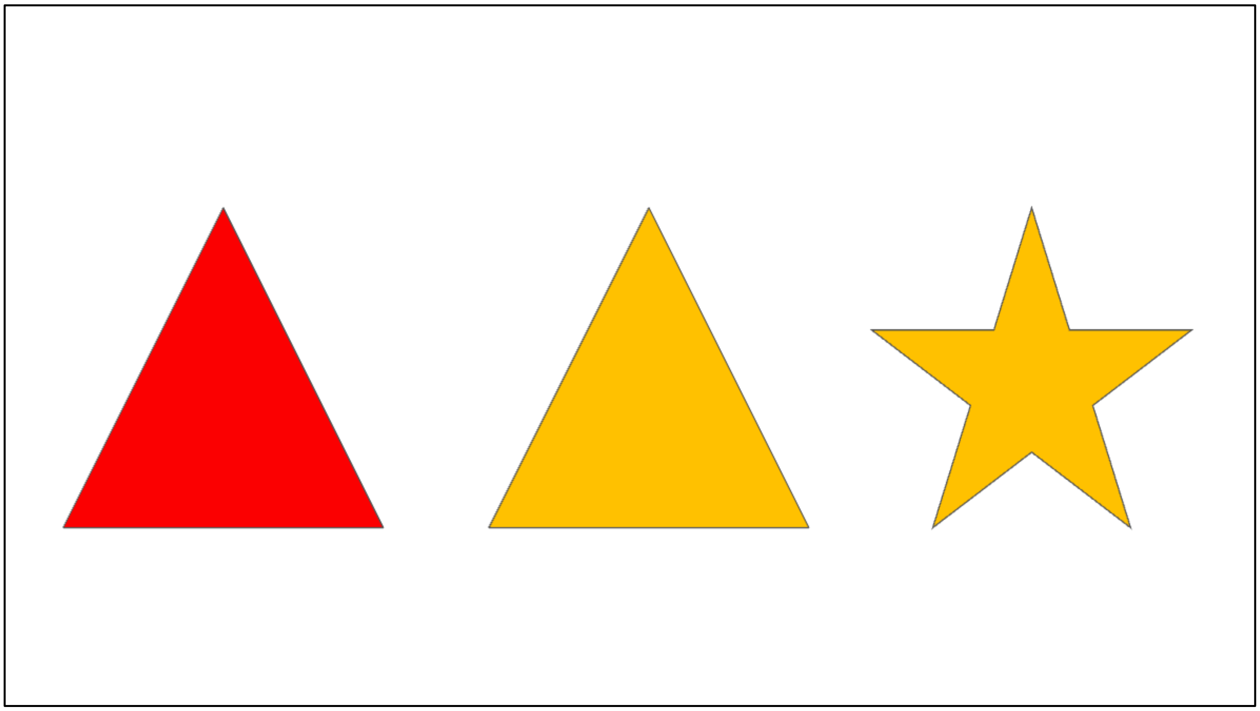




# Appendix H

Stimulus Presentation: B-C





# Appendix I

## Stimulus Presentation: C-A