MEDIATION BY INTRAVERBAL NAMING IN

CHILDREN'S EQUIVALENCE TEST PERFORMANCE

by

CHARLOTTE LYNN CARP

Bachelor of Science, 2004
Texas Christian University
Fort Worth, Texas

Master of Science, 2010
Texas Christian University
Fort Worth, Texas

Submitted to the Graduate Faculty of the
College of Science and Engineering
Texas Christian University
in partial fulfillment of the requirements
for the degree of

Doctor of Philosophy

May 2012

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

MEDIATION BY INTRAVERBAL NAMING IN

CHILDREN'S EQUIVALENCE TEST PERFORMANCE

Words often take on a symbolic nature when used by speakers in verbal communities. That is, persons often use words to *refer* to objects, events, or relations (Premack, 1977). For example, the spoken word "cat," the written word CAT, and pictures of cats, may enter into a complex network of relations in which each stimulus, in a sense, stands for all the others. Therefore, the spoken word "cat" may symbolize, or *refer* to, a picture of a cat. What is interesting about these complex networks, in which the stimuli become substitutable elements, is that many of the relations between the stimuli emerge without being taught explicitly. For example, after a person learns to select a picture of a cat and the written word CAT after hearing the spoken word "cat," then that person can often the select the picture of the cat when presented with the written word CAT, and vice versa, without any direct training; thus, providing the basis for reading comprehension (Sidman, 1994). This implies that stimuli can enter into these complex networks and expand a person's repertoire by simply forming a stimulus-stimulus relation with one stimulus in a previously formed network. For example, in addition to the previous example, if a person now learns a relationship between the spoken English word "cat" and the spoken Japanese word for cat "neko," then that person may now be able to select a picture of a cat and the written word CAT after hearing the spoken word "neko" without any direct training. This is an important phenomenon to understand because one characteristic of complex human behavior requires responding appropriately to novel stimuli on the basis of a previous history with other related stimuli, and is involved in language, reading, and cognition (Spradlin, Cotter, & Baxley, 1973).

Stimulus Equivalence

One framework that has been suggested to explain these complex networks of relations is stimulus equivalence (Sidman, 1994). Stimulus equivalence is a paradigm in which a person is taught a small set of baseline relations between stimuli, and then those stimuli become related in ways that have not been directly taught. A simple example is when a person learns to relate stimulus A to stimulus B, and stimulus B to stimulus C, and is subsequently able to relate stimuli A to C without any explicit training (e.g., Sidman, Cresson, & Wilson-Morris, 1974). Persons are taught to relate specific stimuli together via conditional discrimination training. In a conditional discrimination, a stimulus guides one response in one specific stimulus context, and the same stimulus guides a different response in a different stimulus context (Donahoe & Palmer, 2004). That is, the specific stimulus-stimulus relation that is selected by the reinforcer depends on the context in which the same stimulus appears.

In the laboratory, conditional discriminations are typically studied with a procedure called matching-to-sample (MTS). In a MTS procedure, the participant is presented with a sample stimulus (most commonly visual or auditory) and responds by selecting a comparison stimulus (typically visual) from an array of concurrently available stimuli; for example, by touching or with a mouse click. This task involves a conditional discrimination because the comparison selection that will be reinforced depends upon (or is conditional upon) which sample stimulus is presented (Catania, 2007). For example, if the sample stimulus is one of three A stimuli (i.e., A1, A2, or A3) and the comparison stimuli are three B stimuli (i.e., B1, B2, and B3), the participant will learn to select the comparison stimulus B1 and not B2 or B3 in the presence of the sample stimulus A1, B2 but not B1 or B3 in the presence of A2, and

B3 but not B1 or B2 in the presence of A3. Subsequently, if the comparison stimuli are now three C stimuli (i.e., C1, C2, and C3) in place of the three B stimuli, then the participant will learn to select the comparison stimulus C1, not C2 or C3, in the presence of the sample stimulus A1, etc. MTS tasks in which the sample stimulus is auditory and the comparison stimuli are visual are referred to as auditory-visual conditional discriminations (AVCD), and MTS tasks in which both the sample and comparison stimuli are visual are referred to as visual-visual conditional discriminations (VVCD). When two baseline stimulus-stimulus relations (i.e., A-B and A-C) have been established, persons can often perform with great accuracy when presented novel relations between the stimuli (i.e., the B-A, C-A, B-C and C-B relations). These novel stimulus-stimulus relations, which are strengthened by a training procedure, but not directly reinforced within the procedure, are referred to as derived stimulus relations (Donahoe & Palmer, 2004).

Stimuli are said to enter into equivalence classes when the trained stimulus-stimulus relations demonstrate the three defining derived stimulus relations of reflexivity, symmetry, and transitivity (Sidman & Tailby, 1982). For clarity, when describing the three defining derived stimulus relations with regard to MTS procedures, training and testing will be described for only one set of sample stimuli (A1, B1, C1); however, in typical studies, the need for conditionality among the trained relations requires simultaneous training with two or more sets of sample stimuli (e.g., A1, B1, C1, A2, B2, C2, A3, B3, C3). *Reflexivity* refers to generalized identity matching in that the selection of the comparison stimulus that is reinforced as correct is the same as (i.e., matches) the sample stimulus. An example of this is selecting A1, but not A2 or A3, from an array of comparison stimuli when A1 is the sample stimulus. *Symmetry* refers to the functional reversibility of trained stimulus-stimulus

relations. That is, training a person to select B1, but not B2 or B3, in the presence of A1, then evaluating if selecting A1, but not A2 or A3, in the presence of B1 will emerge in the absence of direct reinforcement. To evaluate *transitivity*, at least three stimuli with two directly trained stimulus-stimulus relations that share a common stimulus are required. Therefore, if training contingencies establish selecting B1, not B2 or B3, in the presence of A1 and selecting C1, but not C2 or C3, in the presence of B1, then transitivity would be demonstrated if C1, but not C2 or C3, is selected in the presence of A1 and if A1, but not A2 or A3, is selected in the presence of C1. These defining properties of stimulus equivalence are shown in Figure 1.
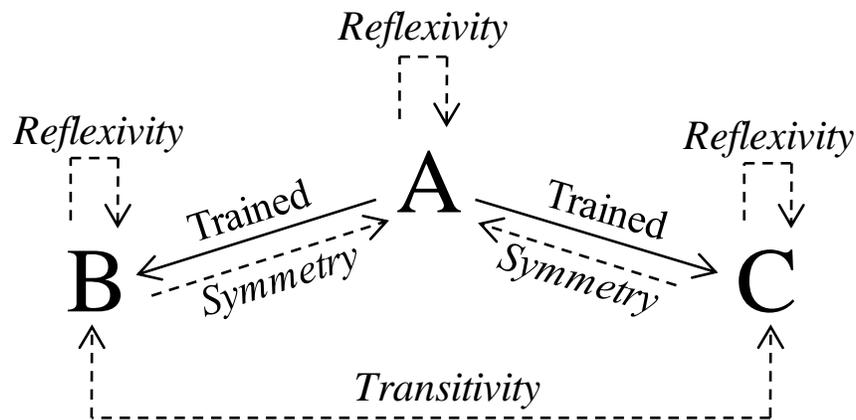
Figure 1. A diagrammatic presentation of the three defining properties of stimulus equivalence. The solid arrows indicate trained relations, and the dotted arrows indicate derived relations. Figure from Hayes and Hayes (1986).

Stimulus equivalence is important at both basic and applied levels of research. At the more basic level of research, stimulus equivalence is important because it provides researchers with a model to study novel behavior using behavioral mechanisms. This idea advances a behavioral conceptualization of human language and cognition, in that instances

of behavior can be governed by a small set of basic principles (Hayes, Barnes-Holmes, & Roche, 2001; Sidman, 1994). In a more applied sense, the efficiency with which stimulus equivalence procedures can expand a person's repertoire, as the training of a small set of baseline stimulus-stimulus relations yields a multitude of untrained relations, has led to applications ranging from language interventions for children with developmental disabilities (Greer & Ross, 2008; Rehfeldt & Barnes-Holmes, 2009) to instruction of college-level course material (e.g., Fields et al., 2009; Fienup, Covey, & Critchfield, 2010; Fienup & Critchfeild, 2010; Ninness et al., 2009; Walker, Redfeldt, & Ninness, 2010).

Stimulus Equivalence and Verbal Behavior

A major source of interest in identifying critical variables that affect stimulus equivalence is its relationship to verbal behavior. To date, humans are the only species in which multiple conditional discriminations have consistently led to the emergence of all derived stimulus relations (Donahoe & Palmer, 2004). Thus, experimental studies have yet to find reliable evidence of all basic relations required for an equivalence class (i.e., reflexivity, symmetry, and transitivity) in non-humans. Despite positive reports of reflexivity (Barros, Galvao, & McIlvane, 2002; Herman & Gordon, 1974; Herman, Hovancik, Gory, & Bradshaw, 1989; Oden, Thompson, & Premack, 1988; Pena, Pitts, & Galizio, 2006), and transitivity (D'Amato, Salmon, Loukas, & Tomie, 1985; Kuno, Kitadate, & Iwamoto, 1994), symmetry has proven to be more difficult to demonstrate (see review, Lionello-DeNolf, 2009). Despite considerable effort, positive reports of symmetry have only been shown in one study with pigeons (Frank & Wasserman, 2005). The most convincing studies demonstrating equivalence in nonhumans have been studies with 2 sea lions (Kastak & Schusterman, 2002; Kastak, Schusterman, & Kastak, 2001; Schusterman & Kastak, 1993);

however, these studies are controversial due to certain issues. The fact that equivalence has not been reliably shown in organisms absent of verbal repertoires is an important issue when considering the theory of stimulus equivalence, in which a major point distinguishing different theoretical accounts is the role of verbal behavior. Three major theoretical accounts have been proposed in the literature. That is, Sidman's (1994 and 2000) perspective, the Naming Hypothesis developed by Horne and Lowe (1996), and Relational Frame Theory (RFT) discussed in Hayes et al. (2001). Although both the Naming Hypothesis and RFT differ in the proposed role that verbal behavior plays in equivalence classes, they both suggest that the ability to form equivalence classes is related to the verbal capabilities of the organism (Lionello-DeNolf, 2009). Sidman's perspective, however, takes a different viewpoint, in positing that verbal behavior is not necessary for the formation of equivalence classes. The present study focuses on the Naming Hypothesis due to specific research questions concerning this study; therefore, Sidman's perspective and RFT will only receive brief discussion.

Sidman (1994) proposes that equivalence is a primitive function, and that just like other primitive functions (e.g., reinforcement and stimulus discrimination), it does not emerge from other behavioral processes. Specifically, equivalence classes are a direct outcome of reinforcement contingencies, and the environment in which these reinforcement contingencies occur (i.e., conditional discrimination training) creates the pre-requisite conditions for the basic equivalence relations to emerge. Therefore, contingencies of survival throughout our evolutionary history have provided us with the necessary skills for equivalence classes to emerge from conditional discriminations, in the same manner that behavior can be reinforced and controlled by specific stimuli. In suggesting that all species

have the potential to form equivalence classes, additional factors (e.g., testing conditions, context, and history) will play a role in realizing that potential. Sidman takes the position that equivalence classes arise first, and are later constrained by verbal behavior. For example, to a newborn infant, the world may be one large stimulus class, but during his interactions with verbal communities he learns that certain objects are equivalent to other objects in some situations but not others. So he may learn that in the context of the word "animal" that "dog" and "elephant" are in an equivalence class, but in the context of the word "zoo" or "pet" they are not.

RFT (Hayes et al., 2001) does not address the abilities of non-human animals, but suggests that equivalence is the result of a prolonged history of reinforcement operating with in a verbal community in which stimuli are related together via relational frames, and the equivalence relation is just one type of relational frame. Relational frames are defined in terms of three properties: mutual entailment, combinatorial entailment, and the transformation of function. These three properties come about via multiple exemplar training. A relational frame is conceptualized as a three-term contingency in which the third term is a contextual cue, the second term is arbitrary applicable relational responding (e.g., selecting B in the presence of A having learned to select A in the presence of B), and the first term is a history of differential reinforcement (i.e., the consequence contingent on the specific relational response in the presence of the contextual stimulus). Therefore, any stimuli can be involved in relational frames if given the appropriate context. RFT states that any organism that can demonstrate complex types of learning is capable of learning to respond to non-arbitrary relations between stimuli, but focuses on humans because of their history of learning to respond to the relations between stimuli due to contextual cues from the verbal

community (e.g., a child speaking English around his friends and Spanish around his parents) and not by the physical form of the stimulus (Barnes, 1994).

Horne and Lowe's (1996) Naming Hypothesis

Horne and Lowe's (1996) Naming Hypothesis builds upon Skinner's (1957) functional account of verbal behavior. Skinner's basic assumption was that human verbal behavior is operant behavior that is controlled by specific antecedent and consequent stimuli. An antecedent stimulus precedes the target behavior (e.g., seeing a cat before emitting the verbal response "cat") and may function, for example, as a discriminative stimulus. A consequent stimulus occurs immediately after the target behavior (e.g., receiving praise for emitting the correct verbal response "cat" after seeing a cat) and may function, for example, as a positive reinforcer (Catania, 2007).

Skinner (1957) classified verbal behavior into several types of verbal operants. A verbal operant consist of a verbal response and the antecedent and consequence stimuli to which it is functionally related. One operant is the mand, in which a state of deprivation or an aversive stimulus evokes a verbal response due to a history of response-specific reinforcement. The need for something to drink may evoke the vocal response "water," which is reinforced by receiving water. Another operant is the tact, in which a nonverbal stimulus evokes a verbal response due to a history of generalized or nonspecific reinforcement. Because a child has received praise in the past for emitting the response "cat" in the presence of a specific stimulus, the sight of a cat may evoke the verbal response "cat." Other verbal operants include echoic, textual, and intraverbal behavior, all of which are under discriminative control of a prior verbal stimulus. In echoic and textual behavior, there is point-to-point correspondence between stimulus and response. In the echoic relation, the

antecedent stimulus is an auditory verbal stimulus that evokes a vocal response that produces the same sound pattern as the antecedent stimulus (e.g., saying "Good morning" in response to hearing someone else say "Good morning"). In the textual relation, a written verbal stimulus is the antecedent stimulus that evokes the same vocal verbal stimulus as the response (e.g., reading). In intraverbal behavior, there is no point-to-point correspondence between stimulus and response, and the antecedent stimulus is an auditory verbal stimulus that evokes a vocal response that differs from the antecedent stimulus (e.g., saying "You're welcome" in response to hearing someone else say "Thank you").

Horne and Lowe (1996; see also Dugdale & Lowe, 1990) argue that a person's naming skills are a necessity for the formation of equivalence classes, and therefore, nonhumans or humans lacking verbal behavior skills will not show equivalence. They propose an account in behavior-analytic terms that defines the *naming relation*, which involves bi-directionality of speaker and listener behaviors, and is acquired within the first 24 months via natural reinforcement processes. When naming has fully developed, it functions as a higher-order bidirectional relation, interlocking the tact, echoic, and listener repertoires, making possible the formation of equivalence relations. Horne and Lowe propose that naming involved in the verbal mediation of equivalences classes can be either overt, which allows naming to be assessed through a vocal naming test, or covert, which is un-measurable except through self-report or supplementary talk-aloud procedures. Regardless of whether naming is covert or overt, Horne and Lowe view it as the critical variable in the mediation of equivalence classes.

Horne and Lowe (1996) propose two ways in which naming may mediate the formation of equivalence classes: common naming and intraverbal naming. In common naming, all stimuli in an equivalence class are given the same name. According to Horne and Lowe, on a MTS trial, the participant tacts the sample stimulus (i.e., either covertly or overtly), hears himself tact the sample stimulus, and then selects the corresponding comparison stimulus as a listener. For example, if a child has learned to name a set of pictures as animals and a different set of pictures as fruits, on a MTS trial in which the sample stimulus is a cat and the comparison stimuli are a dog and an apple, the child will tact the sample stimulus of the cat by saying "animal." The child will then hear himself say "animal," and respond as a listener by selecting the specific comparison stimulus that also belongs to the category "animal" (e.g., the picture of the dog), even though these two stimuli have never previously been presented together in the same MTS trial.

In intraverbal naming, all stimuli in an equivalence class are given a different name and are connected by a verbal link or rule (e.g., "psi goes with green" or "psi, green"). On a MTS trial, the participant will tact the sample stimulus (e.g., "psi"), hear himself say "psi," then engage in the verbal link/rule (e.g., "psi goes with green"), hear himself engage in the verbal link/rule, and select a comparison stimulus as a listener (e.g., select the green stimulus). During MTS training, Horne and Lowe suggest that the training contingencies cause the verbal link or rule to become bi-directional via self-echoic skills repeating the verbal link or rule (e.g., "psi, green, psi, green, psi, green, etc."). The participant now not only knows that psi goes with green, but also that green goes with psi due to green being emitted prior to psi. Therefore, participants are now able to respond correctly on symmetry tests. Figure 2 demonstrates intraverbal naming for an equivalence class when an additional

stimulus is trained (i.e., the Greek letter psi, a green stimulus, and a triangle); thus,

demonstrating how equivalence classes form (i.e., how reflexivity, symmetry, and transitivity

can be performed with accuracy) as a function of intraverbal naming.
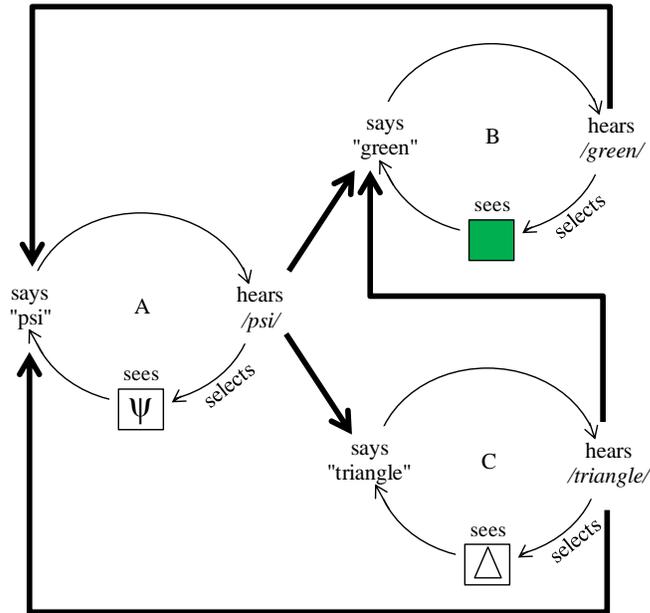
Figure 2. A diagrammatic presentation of the formation of equivalence classes via conditional discrimination tasks through intraverbal naming. The bold solid arrows indicate the bi-directional intraverbal name relations between the stimuli, and the thin solid arrows indicate the process of the naming relation. Figure from Horne and Lowe (1996).

Stimulus equivalence and Naming

In the stimulus equivalence literature, it is unclear what role naming plays in the

formation of equivalence classes. According to Horne and Lowe (1996), equivalence classes

will not form if participants do not have a prior history of naming. Support for this position

has been shown in studies in which common naming facilitates the formation of equivalence

classes (Dugdale & Lowe, 1990; Eikeseth & Smith, 1992; Goyos, 2000; Saunders, Saunder,

Williams, & Spradlin, 1993; Wulfert, Dougher, & Greenway, 1991). In particular, Eikeseth

and Smith (1992) evaluated the effects of naming for four participants diagnosed with autism who failed to demonstrate equivalence classes without naming. Through visual-visual conditional discrimination training, participants were taught an A-B and A-C relation, followed by equivalence testing. No participants demonstrated the emergence of equivalence classes; however, following explicit training of vocal naming responses (i.e., learning common names) to each of the stimuli, the formation of equivalence classes was demonstrated in 2 of the 4 participants.

Subsequently, Goyos (2000) evaluated the effects of common naming on equivalence class formation by using class-specific reinforcers. In Experiment 1, four typically developing children were first exposed to identity-matching visual-visual conditional discrimination training, in which two different reinforcers were used: yellow tokens (R1) and red tokens (R2), both of which could be exchanged for two different reinforcers. R1 followed correct selections of A1, B1, C1, and D1, and R2 followed correct selection of A2, B2, C2, and D2. For example, if children selected A1 and not A2 in the presence of A1, then the reinforcer delivered was R1, and if children selected A2 and not A1 in the presence of A2, then the reinforcer delivered was R2. Following identity-matching training, arbitrary visual-visual conditional discrimination training (i.e., A-B and B-C) commenced with the respective reinforcers (e.g., if participants selected B1 and not B2 in the presence of A1, then R1 was delivered), followed by testing for two equivalence classes (i.e., A1B1C1D1 and A2B2C2D2). Three out of the 4 children emitted vocalizations to the sample and/or comparison stimuli (i.e., labeling the colors), and results showed that these participants demonstrated equivalence classes. One participant did not emit vocalizations and did not demonstrate the emergence of equivalence classes. This participant was then taught to emit

vocal naming responses to the stimuli (e.g., emitting the vocal response "yellow" in the presence of A1, B1, C1, and D1), which resulted in the formation of equivalence classes. In Experiment 2, two participants from Experiment 1 were exposed to identity-matching trials with the D stimuli, in which the class-specific reinforcers were reversed so that the selection of D1 and not D2 in the presence of D1 resulted in the delivery of R2 and the selection of D2 and not D1 in the presence of D2 resulted in the delivery of R1. Following identity-matching training, equivalence tests were administered. For both participants, equivalence classes remained consistent with the original classes, that is, they did not reverse due to the reinforcer reversal identity-matching trials. Goyos stated that although the reinforcer contingency had been reversed, one participant continued to name the stimuli according to the initial training. This could be one possible explanation to why the equivalence classes remained constant with the original classes, even though the reinforcement contingency had been reversed. Therefore, both participants were exposed to vocal name training (i.e., this time the stimuli A1,B1,C1, and D1 set the occasion for the vocal response "red" and the stimuli A2, B2, C2, and D2 set the occasion for the vocal response "yellow"), which resulted in the formation of the reverse reinforcer equivalence classes. Furthermore, support has also been shown in studies demonstrating that equivalence classes more easily form using words that rhyme (Randell & Remington, 1999; Randell & Remington, 2006), words that are easier to pronounce for the participants (Holth & Arntzen, 1998), stimuli that are more familiar to the participant (Arntzen, 2004), and pictures rather than abstract stimuli (Arntzen & Lian, 2010).

Another group of studies consistent with naming demonstrates that participants' with limited verbal skills and no prior history of naming frequently do not show formation of

equivalence classes (Barnes, McCullagh, & Keenan, 1990; Devany, Hayes, Nelson, 1986; Eikeseth & Smith, 1992; Goyos, 2000). Devany et al. (1986) evaluated the formation of equivalence classes in three groups of participants with different verbal capabilities. The groups included one group of typically developing preschoolers, one group of children with intellectual disabilities who used speech or signs spontaneously and appropriately, and one group of children with intellectual disabilities who did not. All participants were exposed to visual-visual conditional discrimination training in which four conditional discriminations were taught (i.e., A-B, A-C, D-E, and D-F), followed by equivalence testing. Results showed that participants with a verbal repertoire (i.e., the typically developing children and children with intellectual disabilities) demonstrated the formation of equivalence classes, whereas none of the children with intellectual disabilities without verbal repertoires did. However, this study has been criticized for several reasons. For example, Saunders and Green (1996) questioned the validity and reliability of the assessments of the participants' verbal repertoires. Moreover, they questioned whether or not baseline performances during visual-visual conditional discrimination training were maintained due to immediate testing following meeting the training criterion. Participants were given nonreinforced tests for equivalence immediately following criterion performance for the baseline conditional relations. During the unreinforced equivalence tests, participants failed the baseline relations, which are prerequisites for equivalence. Therefore, if the prerequisite relations for equivalence (baseline conditional relations) were not well established during training, it would seem unlikely that they would be maintained during testing. Finally, they questioned whether the verbal instructions given to participants actually controlled responding by participants with very limited verbal repertoires.

Barnes et al. (1990) replicated and extended the Devany et al. (1986) study with typically developing preschool children, severely to profoundly deaf children who had verbal ages greater than two years, and severe to profoundly deaf children who had verbal ages less than two years. Thus, they evaluated the effects of verbal repertoires on the formation of equivalence classes in participants with no intellectual disabilities. Results were consistent with Devany et al. in that all participants with verbal repertoires showed equivalence classes and the only participant who had a verbal age of less than 2 and differed from the others by not having an expressive or receptive verbal repertoire of common nouns, did not demonstrate equivalence classes. These studies show that visual-visual conditional discrimination training may not set the occasion for the emergence of equivalence classes in the absence of the corresponding speaker behavior (i.e., naming). Although these studies do not demonstrate that verbal behavior (e.g., naming) is necessary for the formation of equivalence classes, they do suggest that the formation of equivalence classes and verbal behavior are closely related.

Finally, the fact that equivalence classes have not reliably been shown in non-humans thus far, provides additional support for the role of verbal behavior (e.g., naming) in the formation of equivalence classes. That is, if equivalence classes were to be shown in non-humans, this would render Horne and Lowe's (1996) Naming Hypothesis, with regard to the formation of equivalence classes, implausible. Despite considerable effort to show equivalence in nonhumans, unsuccessful attempts have been demonstrated with pigeons, monkeys, baboons, and other primates (e.g., D'Amato et al., 1985; Sidman et al., 1982). The most successful attempts to demonstrate the formation of equivalence classes in non-humans are studies with two sea lions (Kastak & Schusterman, 2002; Kastak et al., 2001;

Schusterman & Kastak, 1993), though these studies are controversial. Schusterman and Kastak (1993) first trained sea lions on VVCDs with 30 stimulus sets (i.e., A1-A30, B1-B30, C1-C30) to establish multiple three-member classes (i.e., A1B1C1, A2B2C2, etc.). Initial training consisted of A-B visual-visual conditional discrimination training on all 30 sets. Then, symmetry was tested on 12 of the sets, testing 6 at a time. That is, the sea lions were presented with symmetry test for 6 stimulus sets, and if symmetry did not emerge, then symmetry was explicitly trained on those 6 sets and tested again with the remaining 6 sets. Then the B-C relation was trained and tested in the same manner, followed by A-C testing (because the pre-requisite A-B and B-C relations had already been trained). Finally, additional A-C training was conducted followed by C-A testing, which was done in a similar manner as the B-A and C-B relations. Following these training and testing procedures, equivalence relations emerged. Schusterman and Kastak proposed that equivalence was demonstrated because of direct training of the relations with other stimulus sets (i.e., multiple exemplar training), and randomly presented multiple incorrect comparison stimuli during training. This methodology, according to Schusterman and Kastak, ensured that the incorrect comparisons did not become part of the equivalence class due to reject control. Horne and Lowe (1996) argued that procedural artefacts led to the sea lions to respond to the stimuli as compounds, possibly explaining their apparent successes; however, some researchers disagree (e.g., see Saunders & Green, 1996). Horne and Lowe (1996) also state that even if equivalence is shown in nonhumans, it may emerge through different previous histories and different processes to those in humans; therefore, equivalence in nonhumans is not of great interest to them. Because their theory describes processes developed through previous verbal histories, species with verbal behavior (i.e., humans) are of primary interest.

Research supporting the naming hypothesis suggests that naming may be a key variable in the formation of equivalence classes. However, evidence against the naming hypothesis suggests that although naming may facilitate stimulus class formation, it is not necessary for the formation of equivalence classes. The evidence against the naming theory gives support for Sidman's perspective that the formation of equivalence classes involve a primitive function that exist prior to verbal behavior. Evidence against the naming hypothesis is demonstrated in studies where verbal participants can demonstrate the formation of equivalence classes, but cannot emit the appropriate common vocal naming response to the stimuli in a follow-up naming test (Green, 1990; Lazar, Davis-Lang, & Sanchez, 1984; Sidman, Cresson, & Willson-Morris, 1974; Sidman, Kirk, & Willson-Morris, 1985; Sidman & Tailby, 1982; Sidman, Willison-Morris, & Kirk, 1986). If naming mediates the formation of equivalence classes, then one would expect verbal participants to emit common vocal naming responses to the stimuli. For example, 4 participants in Sidman et al. (1986) and 5 participants in Green (1990), all with intellectual disabilities, were taught AVCDs and VVCDs, then tested for the formation of equivalence classes. Following the formation of equivalence classes, vocal naming tests were given to evaluate whether participants gave consistent names to all stimuli in a class. All participants were able to form two types of equivalence classes, one type containing auditory-visual and VVCDs and another type containing only VVCDs. Out of these nine participants, vocal naming responses were not consistently emitted to the visual stimuli for 6 participants in auditory-visual classes and for 8 participants in visual-visual classes. Therefore, for the majority of participants with intellectual disabilities in these two studies, equivalence classes were demonstrated in the absence of corresponding vocal responses.

Lazar et al. (1984) showed similar results in a vocal naming test after equivalence classes had been taught via visual-visual conditional discrimination training with typically developing children. Following the formation of equivalence classes with five sets of stimuli (i.e., A, B, C, D, and E), no participants assigned a common name to all stimuli in a class. The authors point out that many of the vocal naming responses emitted by the participants consisted of letters and numbers that shared common physical properties with the stimuli. For example, the word "triangle" was emitted when participants were presented with the upper-case delta. As some of these studies assess vocal naming responses following the formation of equivalence classes in participants with intellectual disabilities, it can be argued that vocal responses may not have been in the participants' repertoires (Stromer, MacKay, & Remington, 1996). On the other hand, as participants were verbal, it is possible that although the participants did not emit overt names for the stimuli during the follow-up naming test, they were still covertly naming the stimuli during conditional discrimination training.

Horne and Lowe (1996; also see Dugdale & Lowe, 1990) take issue with the way the vocal naming tests are conducted, and argue that this is not an appropriate test to assess whether or not naming is mediating equivalence classes. In these vocal naming tests, stimuli are presented to the participants and the experimenter asks, "What is this?" or ""What is the name for this?" Horne and Lowe argue that since vocal naming is assessed following the training and testing of equivalence classes, it may not correspond with the verbal behavior emitted by the participant in the actual experimental conditions. They suggest that naming should be assessed by recording spontaneous vocalizations that children emitted during conditional discrimination tasks.

Horne and Lowe (1996) discussed unpublished studies conducted in their laboratory that supported the naming hypothesis when spontaneous vocalizations were recorded. They taught 29 children of differing ages to select the green comparison stimulus when a vertical line was presented as the sample, and to select the red comparison stimulus when a horizontal line was presented as the sample (i.e., the A-B relation) via visual-visual conditional discrimination training. Then, children were taught to select the triangle comparison stimulus when a vertical line was presented as the sample stimulus, and to select the cross comparison stimulus when the horizontal line was presented as a sample stimulus (i.e., the A-C relation). Out of the 29 children, 17 passed the equivalence test (i.e., with success being related to age), all of who emitted correct vocal intraverbal names for the sample-comparison pairs during visual-visual conditional discrimination training. For example, during training of the A-B relation, when participants were presented with the vertical line as a sample stimulus and the green stimulus as a comparison stimulus, participants would respond with "up green."

Following the suggestion to assess vocal naming responses spontaneously throughout conditional discriminations tasks, other studies have found differing results (Boelens, Van der Broek, & Klarenbosch, 2000; Goyos, 2000; Smeets & Barnes-Holmes, 2005), in that participants' vocal naming responses were absent or not consistent. Boelens et al. (2000) demonstrated symmetry in 2-year-old children via visual-visual conditional discrimination training. Spontaneous naming was assessed, and results showed that very few vocal naming responses were emitted throughout experimental trials. That is, across 7 participants, only 19 out of 2,813 total trials consisted of a common vocal naming response being emitted during that trial. Smeets and Barnes-Holmes (2005) compared spontaneous common vocal naming

responses emitted for auditory-visual equivalence classes and visual-visual equivalence classes for typically developing children. Results showed that 1 child never emitted vocal naming responses for auditory-visual classes and 5 participants never emitted vocal naming responses for visual-visual equivalence classes, although all 6 participants met criterion in conditional discrimination training in the same number of trials as other participants and passed the equivalence test.

It is worth noting, however, that O'Donnell and Saunders (2003) propose that many studies involving participants with developmental disabilities, who show equivalence formation in the absence of corresponding vocal naming responses, may not be representative of their populations. Many of these studies do not explicitly test participants' verbal skills, and therefore may be subject to implicit screening. That is, participants may have been included in studies on the basis of known arbitrary matching abilities, or excluded from studies when they did not learn arbitrary matching (because it did not matter for the questions asked in those studies). O'Donnell and Saunders (2003) suggest that this is possible for two reasons. First, several studies have shown that it is very difficult to teach conditional discriminations to persons with intellectual disabilities (Saunders & Spradlin, 1989, 1990, 1993), yet these difficulties are not shown in the literature in that participants involved in the literature with intellectual disabilities often acquire conditional discriminations with relatively rudimentary procedures. Second, in studies involving conditional discriminations, when training failures are explicitly reported, the failure rate is high; however, in studies of stimulus equivalence, very few teaching failures are reported. This suggests that studies aiming to evaluate the formation of equivalence classes in verbally impaired participants should explicitly test participants' verbal skills to ensure their verbal capabilities, as well as

be prepared to use alternate methods to teach conditional discriminations if the typical

teaching procedures used in stimulus equivalence research are not sufficient (e.g., Saunders

& Spradlin, 1989, 1990, 1993; Perez-Gonzalez & Williams, 2002; Pilgrim, Jackson, &

Galizio, 2000).

To date, there have been two studies that have used standardized tests to assess the

verbal capabilities of their participants and ensure limited verbal skills (Brady & McLean,

2000; Carr, Wilkinson, Blackman, & McIlvane, 2000). Equivalence classes formed for 4 out

of 5 participants for Carr et al. (2000), and for 2 out of 4 participants for Brady and McLean

(2000). In Experiment 1, Carr et al. (2000) evaluated the formation of equivalence classes for

three severely intellectually disabled participants without a vocal naming repertoire. All

participants entered into the study with AVCD skills, and were taught two VVCDs followed

by equivalence tests. Results showed that all participants demonstrated the formation of

equivalence classes. Experiment 2 extended the previous experiment by testing two

participants, both with diagnoses of intellectual disabilities and autism; however, one of the

participants had a highly developed echoic repertoire. Results showed that the formation of

equivalence classes was only demonstrated in the participant without the echoic repertoire.

This is interesting, as an echoic repertoire is an essential component of Horne and Lowe's

account. Since Horne and Lowe (1996) questioned whether or not the formation of

equivalence classes in nonhumans (e.g., sea lions) emerged from the same prior histories and

processes as humans, these two latter studies may have provided the most convincing

evidence against naming as a verbal mediator (although further replication is needed). That

is, if humans with very limited verbal repertoires can form equivalence classes, then naming

may not be necessary for the formation of equivalence classes. Nevertheless, further research

is needed to determine the necessity of verbal behavior for the formation of equivalence classes.

Intraverbal naming

Previous research on assessing naming (e.g., Boelens et al., 2000; Green, 1990; Sidman et al., 1986) during or after equivalence testing has only focused on the common naming mechanism of Horne and Lowe's (1996) naming hypothesis. However, according to Horne and Lowe, failure to find evidence that equivalence class formation was mediated by common names does not necessarily mean that verbal mediation did not occur. Horne and Lowe propose that equivalence class formation could also be mediated by intraverbal naming. As previously noted, intraverbal naming is the second mechanism proposed by Horne and Lowe to mediate equivalence formation, and it involves the establishment of intraverbal relations between specific names of stimuli. The specific names may be the conventional names associated with the stimuli, or in the case of arbitrary stimuli, the participant may assign names to the stimuli based on their similarity with other stimuli. The intraverbal relations between the specific names are proposed to be acquired incidentally as participants name the stimuli during visual-visual conditional discrimination training. Once established, they are available to mediate performance during testing.

Horne and Lowe's description of intraverbal naming has received very little attention in the literature. To date, there is only one study that attempts to address the intraverbal naming mechanism (Smeets & Barnes-Holmes, 2005). In addition to assessing spontaneous common vocal naming responses emitted for auditory-visual and visual-visual equivalence classes for typically developing children, Smeets and Barnes-Holmes also assessed whether equivalence could have resulted from intraverbal naming. According to the participants'

naming responses during training and testing, they were divided into two groups: (a) a "consistent" group of 7 children who consistently gave each stimulus a different name (e.g., consistently calling the A1 stimulus an "elephant" and the A2 stimulus a "car") and (b) an "inconsistent" group of 8 children whom gave the same stimuli different names (e.g., sometimes calling the A1 stimulus an "elephant" and sometimes calling it a "car"). If intraverbal naming was mediating performance on the equivalence tests, then it would be expected that participants in the consistent group would demonstrate equivalence and participants in the inconsistent group would not. However, the results showed that only 3 participants in the consistent group demonstrated equivalence, and 5 participants in the inconsistent group demonstrated equivalence. Smeets and Barnes-Holmes concluded that naming and equivalence class formation were independent processes.

As stated previously, Smeets and Barnes-Holmes (2005) is the only study that has assessed the intraverbal naming mechanism, and this has only been in the context of evaluating if equivalence class formation was correlated with participants' providing a unique name for each stimulus. Therefore, more research is needed on the intraverbal naming mechanism to experimentally investigate if it could be a possible mechanism for the formation of equivalence classes.

Purpose of the present study

The current study was intended to investigate Horne and Lowe's (1996) intraverbal naming hypothesis as a possible mechanism for the formation of equivalence classes. The primary purpose of the study was to evaluate the effects of visual equivalence class formation, in which all visual stimuli have experimenter-supplied names, on emergent intraverbal relations in kindergarteners ages 5 to 6. First, children were exposed to tact

training in which they learned to emit vocal tacts for (i.e., name) pictures of states, birds, and flowers. Following tact training, they were exposed to visual-visual conditional discrimination training, in which a relation between the pictures of states and birds (i.e., the A-B relation), and states and flowers (i.e., the A-C relation) was taught, followed by equivalence testing. After equivalence testing, the children were exposed to an intraverbal test consisting of vocal-verbal conditional discriminations with the corresponding vocal names of the visual stimuli.

There are at least two reasons why intraverbals might emerge based on such training. The first is that according Horne and Lowe's (1996) naming hypothesis, if participants passed an exclusively visual-visual equivalence test, they must have done so because during training they either (a) applied a common name to the stimuli involved in each equivalence class, or (b) acquired intraverbal relations between the specific names of stimuli in each class. In the present study, participants were supplied with the specific names, which should have promoted use of intraverbal naming. As a result, the naming hypothesis should predict that participants who passed the equivalence test should also pass a subsequent test for the emergence of at least those intraverbal relations that are proposed to have mediated their test performance. Alternatively, these same intraverbal responses could have emerged as a by-product of visual-visual conditional discrimination training, equivalence testing, or a combination of the two; that is, the names could have participated in the equivalence class or relational frame (Hayes et al., 2001) with the visual stimuli, without having a functional role in mediating equivalence test performance. In either case, the emergence of intraverbals in this situation might have practical implications for establishing novel intraverbals. However, if any participants pass the equivalence test but fail the intraverbal test, Horne and Lowe's

hypothesis becomes implausible, as nonexisting intraverbal relations cannot mediate performance. This would provide a very strong argument against Horne and Lowe's hypothesis.

A second purpose of the study was to evaluate whether requiring participants to name the visual stimuli during a second equivalence test might allow the formation of equivalence classes. If participants failed the equivalence tests, then additional equivalence tests were conducted in which the participant was required to name the sample stimulus during testing. Previous research has shown that equivalence classes can form after requiring participants to learn common names for the stimuli in each class (e.g., Dugale & Lowe, 1990; Eikeseth & Smith, 1992; Goyos, 2000; Saunders et al., 1993; Wulfert et al., 1991). In addition, participants who failed equivalence tests after learning common names for the stimuli in each class passed the tests when prompted to name sample stimuli (using the common name) during testing (Horne, Hughes, & Lowe, 2006; Horne, Lowe, & Randle, 2004; Lowe, Horne, Harris, & Randle, 2002; Lowe, Horne, & Hughes, 2005). It is possible that these naming responses enhanced test performance simply because they increased attention to the stimuli. However, Horne and Lowe (1996) proposed that they did so because common naming of stimuli is one of the mechanisms that can mediate equivalence. Subsequently, if instead of acquiring common names for the stimuli, a participant were to acquire intraverbal relations between their specific names, and then that participant failed the equivalence test. Requiring the participant to name the sample stimuli during the test might prompt intraverbal responses, and thereby result in enhanced performance on both the equivalence and subsequent intraverbal tests. On the other hand, if naming of the sample stimuli enhanced a participant's performance on the equivalence test, but the participant did not have the intraverbal relations

in his or her repertoire; it is more plausible that the naming response functioned to increase attention to the stimuli.

Method

Participants

With their parents' permission, seven children between the ages of 5 and 7 years without any known developmental delays began participation in the study. One participant was dropped from the study due to a failure to maintain and meet criterion in tact training; thus, all figures and tables contain data only from the six participants who completed equivalence testing. All participants spoke English as their first language. Table 1 shows each participant's gender, chronological age, and mental-age scores on the Peabody Picture Vocabulary Test (PPVT), which is a standardized test of receptive language skills. Sessions took place either in an empty preschool library or an empty elementary school library, where the experimenter and child sat beside one another at a table. Sessions lasted approximately 20 to 30 min and were conducted up to 4 times a week. During sessions, the participant earned tokens that were placed on a token board. When the token board was full, the session ended and the participant received 5 min of access to a box that contained a variety of toys and games.

Table 1

Participant Information

Participant	Sex	Age	PPVT Mental Age Equivalent
Niels	M	7 yr	6 yr, 3 mo
Rita	F	6 yr, 9 mo	9 yr, 4 mo
Emilie	F	6 yr, 4 mo	6 yr, 9 mo
Charles	M	5 yr, 6 mo	6 yr, 2 mo
Marie	F	5 yr, 5 mo	6 yr, 5 mo
Leonardo	M	6 yr, 3 mo	7 yr, 7 mo

Stimuli

Figure 3 shows the visual stimuli presented during tact training, visual-visual

conditional discrimination (VVCD) training, and equivalence testing. These stimuli were 5.5

A1	A2	A3	a1 - "Virginia" a2 - "Alabama" a3 - "California
B1	B2	B3	b1 - "dogwood" b2 - "camellia" b3 - "poppy"
C1	C2	C3	c1 - "cardinal" c2 - "yellowhammer" c3 - "quail"
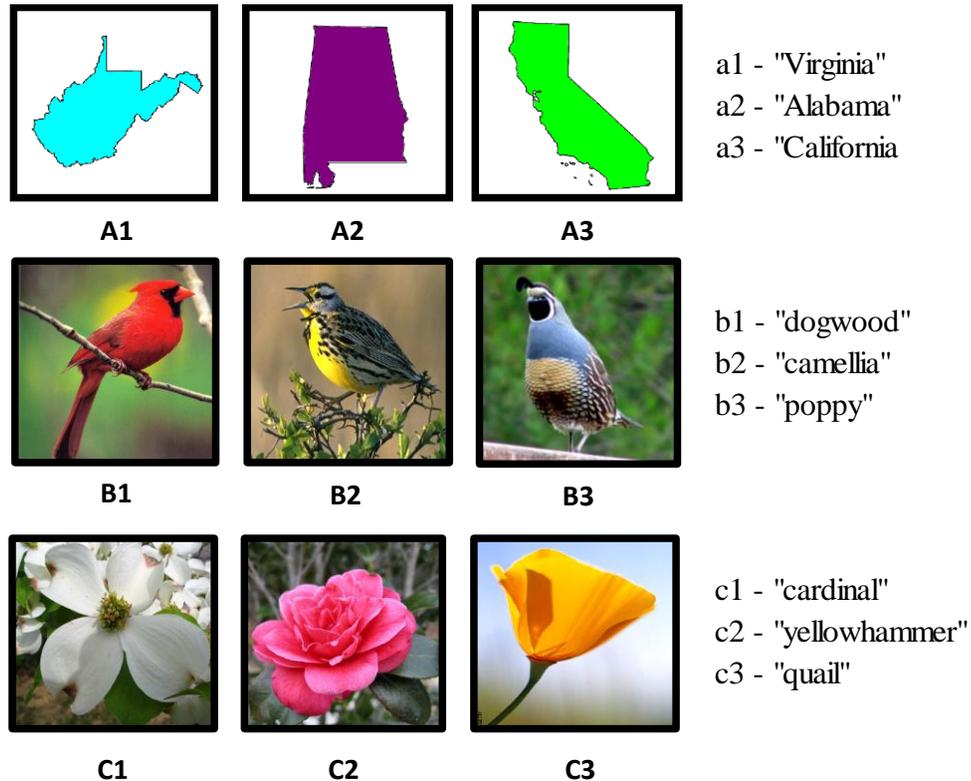
Figure 3. Visual and vocal stimuli used during tact training, VVCD training, and equivalence

testing.

cm x 5.7 cm color photographs consisting of three outlines of states (set A), three pictures of

the state's corresponding birds (set B), and three pictures of the state's corresponding flowers

(set C). During tact training, one picture was presented during a trial. The picture was

centered on a white 20.3 cm x 27.9 cm sheet of paper that was inserted into a transparent

sheet protector (i.e., the *trial page*). Trial pages were contained in a 5.1 cm three-ring binder

and at the completion of each trial, the experimenter flipped the trial page so that the next page became visible. During listener-category pre-training, a trial page consisted of a 20.3 cm x 27.9 cm sheet of paper in which three pictures (i.e., one picture of a state, one picture of a bird, and one picture of a flower) were centered and presented horizontally. During the listener test, a trial page consisted of a white 20.3 cm x 27.9 cm sheet of paper in which nine pictures (i.e., three pictures of states, three pictures of birds, and three pictures of flowers) were presented. VVCD training and equivalence testing were controlled by a 29.0 cm x 19.8 cm laptop screen. The computer program Mestre Libras (Elias, Goyos, Saunders, & Saunders, 2008) was used to organize and show stimuli and consequences, and to record training and testing data from the MTS tasks. The computer stimuli were presented against a blue background.

Table 2 shows the antecedent verbal stimuli that were presented in intraverbal test trials, along with correct vocal responses. The labels a1, a2, and a3 refer to the vocal names of states, regardless of whether the names were presented as antecedent stimuli by the experimenter (as in a-b relations) or produced as correct responses by the participants (as in b-a relations). Similarly, b1, b2, and b3 were vocal names of the state's corresponding birds (e.g., b1 is the state bird for a1), and c1, c2, and c3 were vocal names of the state's corresponding flowers (e.g., c1 is the state flower for a1). All vocal stimuli (i.e., a1, a2, a3, b1, b2, b3, c1, c2, and c3) corresponded to their respective visual stimuli (e.g., a1 is the vocal stimulus for the corresponding visual stimulus A1). Since the a-b and a-c relations were never directly trained but are analogous to the trained A-B and A-C relations, they are referred to as the baseline-like relations. Because the b-a, c-a, b-c, and c-a relations are not truly symmetry and transitivity relations (due to the participants response being topography-

Table 2

Antecedent Stimuli and Correct Responses in Intraverbal Test Trials

	Antecedent Stimuli		Correct Response
Baseline-like			
a–b State–Bird			
[a1] goes with which bird?	[a1] Virginia	[b1] Cardinal	
[a2] goes with which bird?	[a2] Alabama	[b2] Yellowhammer	
[a3] goes with which bird?	[a3] California	[b3] Quail	
a-c State-Flower			
[a1] goes with which flower?	[a1] Virginia	[c1] Dogwood	
[a2] goes with which flower?	[a2] Alabama	[c2] Camellia	
[a3] goes with which flower?	[a3] California	[c3] Poppy	
Symmetry-like			
b-a Bird-State			
[b1] goes with which state?	[b1] Cardinal	[a1] Virginia	
[b2] goes with which state?	[b2] Yellowhammer	[a2] Alabama	
[b3] goes with which state?	[b3] Quail	[a3] California	
c-a Flower-State			
[c1] goes with which state?	[c1] Dogwood	[a1] Virginia	
[c2] goes with which state?	[c2] Camellia	[a2] Alabama	
[c3] goes with which state?	[c3] Poppy	[a3] California	
Transitivity-like			
b-c Bird-Flower			
[b1] goes with which flower?	[b1] Cardinal	[c1] Dogwood	
[b2] goes with which flower?	[b2] Yellowhammer	[c2] Camellia	
[b3] goes with which flower?	[b3] Quail	[c3] Poppy	
c-b Flower-Bird			
[c1] goes with which bird?	[c1] Dogwood	[b1] Cardinal	
[c2] goes with which bird?	[c2] Camellia	[b2] Yellowhammer	
[c3] goes with which bird?	[c3] Poppy	[b3] Quail	

based such as "California" or "Poppy" rather than selection-based such as clicking on a comparison stimulus), those relations are referred to as symmetry-like and transitivity-like, respectively. In all test trials, the instruction "… goes with which ____" was presented along with the contextual stimulus "state", "bird" or "flower" (see Table 2).

Measurement

Data collection. During VVCD training and equivalence testing, the computer recorded correct and incorrect selection responses. The first comparison stimulus that the participant selected (i.e., clicked on) was scored as correct or incorrect. The computer also recorded response latency data in milliseconds (ms), starting when the participant clicked on the sample stimulus and terminating when the participant clicked on a comparison stimulus. During tact training and intraverbal test trials, the experimenter recorded correct and incorrect responses using paper and pencil. The first state, bird, or flower name that the participant vocalized was scored as correct or incorrect. If the participant did not vocalize a state, bird, or flower name within 5 s for tact training and 10 s for intraverbal testing, the trial was scored as incorrect.

Interobserver agreement. A second observer independently recorded data during at least 30% of all tact and 90% of all intraverbal test sessions, either live or from video. For each trial, an agreement was scored if both the experimenter/computer and the second observer scored the trial as correct, or both scored it as incorrect. Interobserver agreement (IOA) was calculated by dividing the number of agreements by the number of trials and multiplying by 100. Niels's mean agreement was 96% (range, 67% to 100%) for tact training and 99 % (range, 96% to 100%) for intraverbal tests. Rita's mean agreement was 98% (range, 83%-100%) for tact training and 99% (range, 96% to 100%) for intraverbal tests.

Emilie's mean agreement was 100% (range, 100% to 100%) for tact training and 96% (range, 93% to 100%) for intraverbal tests. Charles's mean agreement was 100% (range, 100% to 100%) for tact training and 98% (range, 94% to 100%) for intraverbal tests. Marie's mean agreement was 100% (range, 100% to 100%) for tact training and 94% (range, 92% to 100%) for intraverbal tests. Leonardo's mean agreement was 100% (range, 100% to 100%) for tact training and 99% (range, 97% to 100%)for intraverbal tests.

Procedure

Experimental design and overview. A concurrent (for three participants) and non-concurrent (for three participants) multiple-probe design across participants was used to evaluate the effects of visual-visual conditional discrimination training on derived intraverbal relations. As shown in Figure 4, each participant was first exposed to pre-training with familiar visual stimuli to ensure that the participant could perform the tasks in the experiment as well as category-listener pre-training to ensure that the participant could identify pictures of state, bird, and flower categories. Each participant then received tact training in which participants learned to tact each visual stimulus. After criterion was met in tact training, participants were exposed to a brief intraverbal pre-test to ensure that the relations shown in Table 2 were not already in their repertoires. Each participant was then exposed to visual-visual conditional discrimination (VVCD) pre-testing to ensure that the participant could perform a visual discrimination task on the computer. Then VVCD training commenced, followed by an equivalence test. Following equivalence testing, participants were exposed to an intraverbal test. That is, the 18 intraverbals from Table 2 were tested. If participants passed the equivalence and intraverbal tests, then (a) a listener test was conducted to ensure that all three basic relations (i.e., the tact, echoic, and

```
                    ┌─────────────────────────────────────┐
                    │ Pre-Training: Tact and Category-Listener │
                    └─────────────────────────────────────┘
                                     │
                                     ▼
                          ┌──────────────────┐
                          │  Tact Training   │
                          └──────────────────┘
                                     │
                                     ▼
                          ┌──────────────────┐
                          │ Intraverbal Test │
                          └──────────────────┘
                                     │
                                     ▼
                          ┌──────────────────┐
                          │ VVCD Pre-Training │
                          └──────────────────┘
                                     │
                                     ▼
                          ┌──────────────────┐
                          │  VVCD Training   │
                          └──────────────────┘
```

PASS   Equivalence   FAIL

Test

B-C
Training

PASS   Intraverbal   FAIL   PASS   Intraverbal

Test   Test

FAIL

Listener Test
Alternate tact test

Additional
Verbal Relations Tests
Listener test
Alternate tact test
Category-name intraverbal test
Exemplar-name intraverbal
test

Equivalence
test with vocal
naming response to
the sample
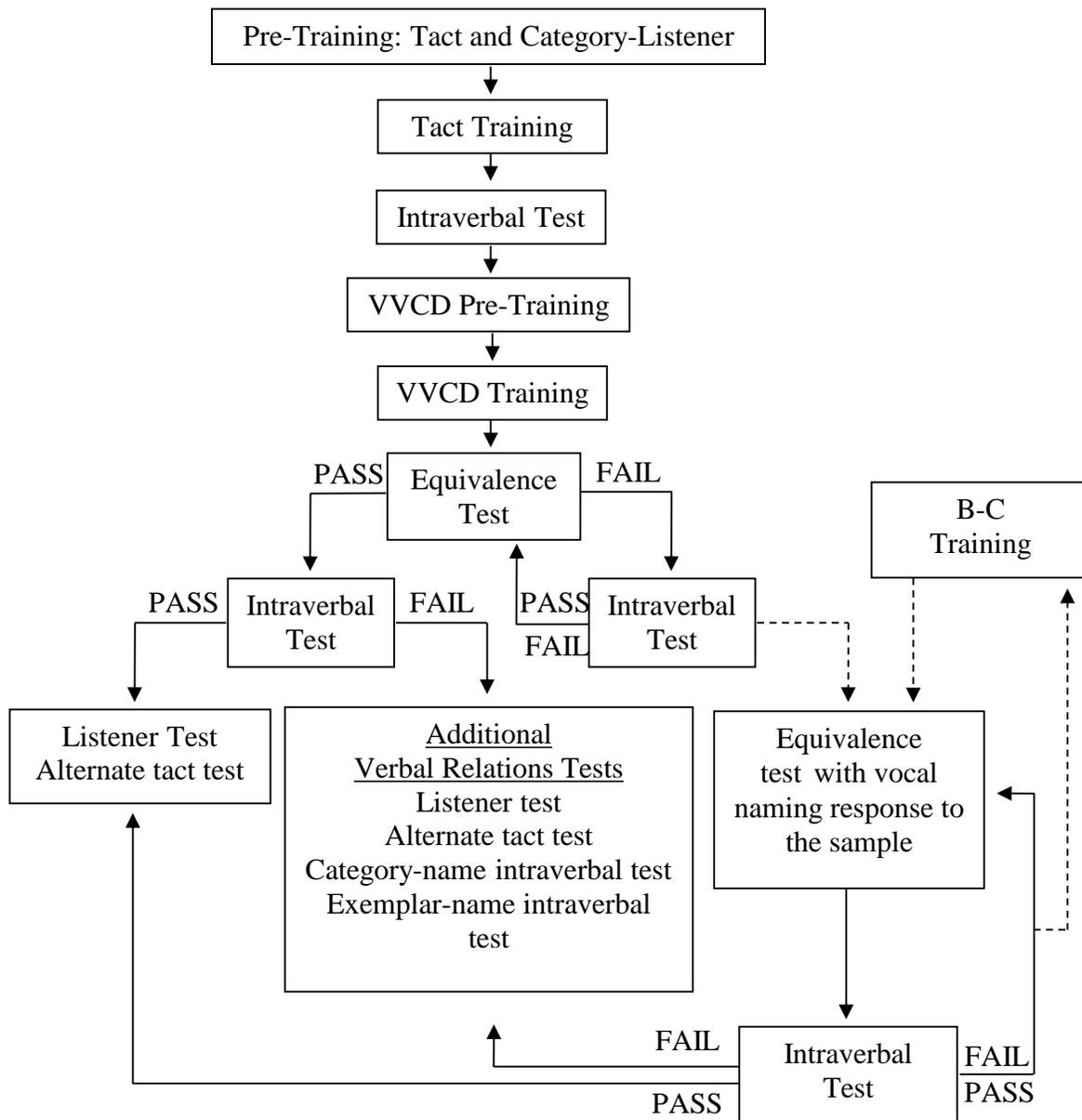
FAIL   Intraverbal   FAIL

Test   PASS

PASS

Figure 4. Sequence of training and testing.

listener relations) of the *naming relation* are in the participants' repertoires and (b) an alternate tact test was conducted to evaluate if the participants had generated their own names for the stimuli other than the experimenter-provided names. If participants passed the equivalence test but intraverbals did not emerge, additional verbal relations tests were conducted to clarify the results. If equivalence classes did not emerge, regardless of performance on the intraverbal test, participants received additional rounds of equivalence and intraverbal testing. If performance on the equivalence and/or intraverbal test did not improve, then the participants were exposed to an equivalence test in which a vocal naming response to the sample stimulus was required followed by an intraverbal test. That is, the introduction of the vocal naming response to the sample stimulus during the equivalence test was staggered across differing amounts of exposure to the equivalence tests, in order to control for the effects of repeated testing and acquisition outside of the experiment. If equivalence classes still did not emerge, then participants were exposed to B-C training, followed by another round of equivalence and intraverbal testing.

Category-listener pre-training. For each trial in category-listener pre-training, the experimenter presented a trial page consisting of one picture of a state, one picture of a bird, and one picture of a flower. The experimenter presented 3 trials for each trial page consisting of the three instructions "Point to the state", Point to the bird", and "Point to the flower." These three instructions were randomly presented for each trial page. Correct responses resulted in praise and a token. If the participant made an incorrect response or did not respond within 5 s, the experimenter pointed to the correct stimulus and said, for example, "This is a state," and proceeded to the next trial or trial page. Criterion was met when the participant responded with 100% accuracy for three consecutive trial pages.

Tact training. Tact training was conducted in five phases. In each trial, the experimenter showed a picture and asked, "What is this?" Table 3 shows the stimulus sets targeted in each phase, along with the consequences of correct and incorrect responses, and the performance criteria for moving on to the next phase or returning to a previous phase.

Table 3

Tact Training Procedures

Phase	Stimulus Sets	Consequence of Correct Response	Consequence of Incorrect Response	Consequence of Prompted Response	Criterion for Moving to Next Phase	Criterion for Returning to a Previous Phase
1	A, B	Praise + token	Vocal prompt and repeat trial	Praise	1 block with 100% correct responses (Phase 2)	
2	B, C	Praise + token	Vocal prompt and repeat trial	Praise	1 block with 100% correct responses (Phase 3)	
3	A, C	Praise + token	Vocal prompt and repeat trial	Praise	1 block with 100% correct responses (Phase 4)	
4	A, B, C	Praise + token	Vocal prompt and repeat trial	Praise	3 consecutive blocks with 100% correct responses (Phase 5)	2 consecutive blocks with less than 67% correct responses for any of the stimuli sets (Phase 1)
5	A, B, C	Praise + token under VR3	Vocal Prompt		3 consecutive blocks with 100% correct responses (VVCD Training)	3 consecutive blocks with less than 100% correct responses (Phase 4)

In phases 1, 2, and 3, pictures were presented in 6-trial blocks containing one presentation of each picture, with presentation order varied across blocks. An immediate vocal prompt was provided in the first 2 blocks. The next two blocks contained a 2-s prompt delay, and the remaining blocks contained a 5-s prompt delay. If the participant responded with a correct

tact, the participant received praise (e.g., "Good job," or "You are so smart," etc.) and a token. If the participant did not respond with the correct tact or did not respond within 5 s, the experimenter provided a vocal prompt (e.g., "No, it's a ____") and repeated the question "What is this?" A correct response following a prompt was praised; however, no token was awarded and the trial was scored as incorrect. In phase 4, all sets of stimuli (i.e., sets A, B, and C) were presented. Pictures were presented in 9-trial blocks containing one presentation of each picture, with presentation order counterbalanced across blocks.

Phase 5 was conducted in the same manner as phase 4, except for the differences in consequences following correct or incorrect responses (see Table 2). Phase 5 was implemented so that participants could experience responding under conditions similar to those during testing, in which consequences were provided for correct responses under a VR3 schedule of reinforcement.

Tact pre-training. Tact pre-training was implemented prior to tact training and conducted in the same manner as tact training, with three exceptions: (a) six familiar stimuli were used that were not presented in the experiment, (b) phase 4 was the only phase that was implemented, and (c) criterion was one 6-trial block of 100% correct responding. Tact pre-training was conducted to ensure that participants understood the instructions and task.

Visual-visual conditional discrimination pre-training. Visual-visual conditional discrimination pre-training was conducted immediately prior to VVCD training with 6 familiar stimuli (i.e., 3 sets of 2 stimuli pairs) that were not presented in other phases of the experiment, and consisted of two phases. The participant was presented with 6-trial blocks in which each picture was a sample stimulus once, with presentation order varied across blocks. Phase 1 was conducted in the same manner as phase 3 in visual-visual conditional

35

Table 4

Stimuli Presented During Visual-Visual Conditional Discrimination Training

Relation	Sample	Comparisons		
		Correct	Incorrect	
A-B (state-bird)	A1	B1	B2	B3
	A2	B2	B1	B3
	A3	B3	B1	B2
A-C (state-flower)	A1	C1	C2	C3
	A2	C2	C1	C3
	A3	C3	C1	C2

Table 5

Visual-Visual Conditional Discrimination Training Procedures

Phase	Relations	Consequence of Correct Response	Consequence of Incorrect Response	Criterion for Moving to Next Phase	Criterion for Returning to a Previous Phase
1	A-B	Animation screen, praise, and a token	Black screen	3 consecutive blocks with 100% correct responses (Phase 2)	
2	A-C	Animation screen, praise, and a token	Black screen	3 consecutive blocks with 100% correct responses (Phase 3)	
3	A-B and A-C	Animation screen, praise, and a token	Black screen	3 consecutive blocks with 100% correct responses (Phase 4)	3 consecutive blocks with less than 83% correct responses (Phase 1)
4	A-B and A-C	Animation screen, praise, and a token under VR3	Black screen	3 consecutive blocks with 100% correct responses (Equivalence Testing)	3 consecutive blocks with less than 100% correct responses (Phase 4)

discrimination training, and phase 2 was conducted in the same manner as phase 4. For all phases, criterion was one block with 100% correct responding.

   Visual-visual conditional discrimination training. During VVCD training, the laptop computer that controlled stimulus presentations was placed on the table in front of the participant. VVCD training was conducted in four phases. Table 4 shows the specific stimuli presented as sample and comparison stimuli during VVCD training trials. Table 5 shows the relations targeted in each phase, along with the consequences of correct and incorrect responses, and the performance criteria for moving on to the next phase or returning to a previous phase. At the beginning of every trial, one of the three sample stimuli (i.e., A1, A2, or A3), was presented in the horizontal center of the top third of the computer screen. Participants clicked on the stimuli by using a mouse. Once the participant clicked on the sample stimulus, the three comparison stimuli were presented below the sample. The three comparison stimuli were presented equidistant from one another across the lower third of the screen with the middle comparison located in the horizontal center. Participants selected a comparison stimulus by clicking on a picture. Depending upon the participant's selection, the computer delivered the programmed consequences for correct and incorrect responses. Correct selection responses produced an animation on the computer screen and verbal praise plus a token provided by the experimenter; incorrect selection responses produced a black screen for 2 s. The animation showed a coin falling into a piggy bank. Verbal praise for correct selections consisted of short sentences such as "Good job!", "Awesome!", "Right!", or "You got it!." An intertrial interval (ITI) of 2 s preceded the next trial. During the ITI, the computer screen was blue. Phases 1, 2, and 3 were conducted in an indentical manner. In phases 1 and 2 pictures were presented in 3-trial blocks containing one presentation of each

stimulus in the specific stimulus set being trained for that phase, with presentation order counterbalanced across blocks. Phase 3 combined the A-B (state-bird) and A-C (state-flower) relations in 6-trial blocks. Each 6-trial block presented each set A stimulus twice, once with set B stimuli as comparison stimuli and once with set C stimuli as comparison stimuli, with presentation order varied across blocks. Phase 4 was conducted in the same manner as phase 3, except for the differences in consequences following correct or incorrect responses (see Table 5). Phase 4 was implemented so that participants could experience responding under conditions similar to those during testing, in which consequences were provided for correct responses under a VR3 schedule of reinforcement. If participants responded correctly on a trial, but that specific trial was not scheduled for reinforcement, then the correct response produced the ITI.

Tact Maintenance. One 9-trial tact block was presented at the beginning of each session of VVCD training to ensure that tacts were maintained. Tact maintenance blocks were conducted in the same manner as phase 5 in tact training. If responding was not 100% correct on the tact maintenance block, the participant returned to phase 5 of tact training; therefore, the participant had to meet the criterion for phase 5 in tact training again before proceeding with VVCD training.

Equivalence testing. Each test consisted of 18 trials. There were three trials for each of the three stimulus sets in which each stimulus in the set was presented as the sample stimulus once. All A-B, A-C, B-A, C-A, B-C, and C-B trials were randomly presented. Table 6 demonstrates all stimulus relations that were tested. Testing was conducted in the same manner as phase 4 of visual-visual conditional discrimination training, that is, under a VR3 schedule of reinforcement in which selection responses on trained trials (A-B and A-C) were

Table 6

Derived Stimulus Relations Presented During Equivalence Tests

Derived Relation		Sample	Comparisons		
			Correct	Incorrect	
B-A	Symmetry	B1	A1	A2	A3
		B2	A2	A1	A3
		B3	A3	A1	A2
C-A	Symmetry	C1	A1	A2	A3
		C2	A2	A1	A3
		C3	A3	A1	A2
B-C	Transitivity	B1	C1	C2	C3
		B2	C2	C1	C3
		B3	C3	C1	C2
C-B	Transitivity	C1	B1	B2	B3
		C2	B2	B1	B3
		C3	B3	B1	B2

Note. Unreinforced trials of the trained A-B and A-C relations will be randomly presented with the derived relation trials during equivalence tests.

reinforced (i.e., the animation screen, praise, and a token). Selection responses on symmetry and transitivity trials produced the ITI, regardless of whether the selection response was correct or incorrect. Because there were 6 trials for each type of relation, equivalence classes were said to emerge if 5 out of 6 correct responses were made for each relation. If participants did not met the criterion on the trained trials during the equivalence test, then another round of testing was administered. If participants continued to fail the trained relations then participants went back to the last phase of VVCD training.

Equivalence testing with a vocal naming response to the sample stimulus. Each test was conducted in the same manner as the original equivalence test, except that participants were required to name the sample stimulus before clicking on the sample stimulus to produce the comparison stimuli. For each trial, the sample stimulus was presented at the top of the computer screen and the experimenter would say, "What is this?" If the

participant did not emit the correct name or did not respond within 5 s, the correct name was provided by the experimenter and repeated by the participant. Following a correct response from the participant (i.e., either unprompted or prompted), the participant could click the sample stimulus to produce the comparison stimuli; no other consequences for correct naming were given. The participant then selected a comparison stimulus. During the equivalence test, each stimulus was presented as a sample stimulus two times; therefore, if the percentage of unprompted correct naming responses fell below 50 for any of the stimuli, the participant returned to phase 5 of tact training. Following the completion of phase 5 of tact training, the equivalence test was re-administered. This did not happen for any participant exposed to this version of the equivalence test.

Intraverbal test. Each test consisted of 27 trials; 18 intraverbal trials and 9 tact maintenance trials. Tact maintenance trials were conducted as in tact training, such that a correct response resulted in praise and a token and an incorrect response resulted in a vocal prompt. In each intraverbal trial, the experimenter vocally presented the instructions and antecedent stimuli shown in Table 2, and waited up to 10 s for a response consisting of the participant's vocalization of a state, bird, or flower name. Following a response or 10 s without a response, the next trial was presented immediately. No feedback or consequences were provided following either correct or incorrect responses. All 18 intraverbal trials (i.e., a-b, a-c, b-a, c-a, b-c, and c-b) were presented in random order, with the 9 tact trials interspersed between them; therefore, each intraverbal test was conducted under a VR3 schedule of reinforcement. The acquisition criterion for the intraverbal test was based only on the 18 intraverbal trials. Criterion was defined as correct responses on 16 out of the 18 intraverbals, or 89% accuracy.

B-C training. B-C training was conducted only for participants who repeatedly failed the intraverbal test both with and without naming. B-C training was conducted in two phases in a similar manner as VVCD training, except that the B stimuli were presented as sample stimuli and the C stimuli were presented as comparison stimuli during trials. Table 7 shows the relations targeted in each phase, along with the consequences of correct and incorrect responses, and the performance criteria for moving on to the next phase or returning to a previous phase. In phase 1, pictures were presented in 3-trial blocks containing one presentation of each stimulus in the B stimuli set, with presentation order varied across blocks. Phase 2 was conducted in the same manner as phase 1, except for the differences in consequences following correct or incorrect responses (see Table 7). Phase 2 was implemented so that participants could experience responding under conditions similar to that during testing, in which consequences were provided for correct responses under a VR3 schedule of reinforcement. If participants responded correctly to a trial, but that specific trial was not scheduled for reinforcement, then the correct response produced the ITI.

Table 7

B-C Training Procedures

Phase	Relations	Consequence of Correct Response	Consequence of Incorrect Response	Criterion for Moving to Next Phase	Criterion for Returning to a Previous Phase
1	B-C	Animation screen, praise, and a token	Black screen	3 consecutive blocks with 100% correct responses (Phase 2)	
2	B-C	Animation screen, praise, and a token under VR3	Black screen	3 consecutive blocks with 100% correct responses (Equivalence Testing)	3 consecutive blocks with less than 100% correct responses (Phase 1)

Symmetry Testing. The symmetry test was conducted for participants who received B-C training to investigate if symmetry could be established. The symmetry test was exactly like the equivalence test; however, because participants had completed training on the B-C relation, the only untrained relations were the symmetry relations B-A, C-A, and C-B. Because there were now 9 trials for each type of relation (i.e., trained and symmetry), symmetry was said to emerge if 8 out of 9 correct responses were made for each relation.

Listener test. The listener test was conducted to ensure that participants were engaging in the naming relation proposed by Horne and Lowe (1996). This test consisted of 36 trials presented in four 9-trial blocks. Each block consisted of one trial page with nine visual stimuli randomly presented and nine randomly presented vocal trial instructions. The experimenter presented the trial page to the participant and then presented the trial instructions (e.g., "Point to Virginia"). Trial instructions consisted of the vocal stimuli "Point to" along with the vocally presented name of one of the nine stimuli. A trial ended when the participant pointed to a specific stimulus. No feedback was given for correct or incorrect responses. Following the child's response, the experimenter presented a new trial. Criterion to pass the listener test was 8 out of 9 correct responses, or 89% accuracy.

Alternate tact test. This test consisted of one 9-trial tact block in which each picture was randomly presented once. Each trial consisted of the experimenter showing a picture and asking, "What is this?" Following a vocal response from the child, the experimenter asked, "Do you call it anything else?" No feedback was given for any vocal responses emitted by the child. This test helped evaluate whether participants were using alternate specific names for the individual stimuli or possibly a common name for the stimuli in an equivalence class.

Category-name intraverbal test. This test consisted of 9 trials in which each of the nine category-name intraverbals shown in Table 8 was randomly presented once. No feedback was given for correct or incorrect responses. The passing criterion was 8 out of 9 correct responses, or 89% accuracy.

Exemplar-name intraverbal test. This test was similar to and had the same passing criterion as the category-name intraverbal test, except that the intraverbals presented were the nine exemplar-name intraverbals shown in Table 8. The category-name and exemplar-names tests were conducted to ensure that the participants' responses on the intraverbal tests were under control of the category names "state," "bird," and "flower."

Table 8

Antecedent Stimuli and Correct Responses in Category-Name and Exemaple-Name Intraverbal Testing Trials

Antecedent Stimuli	Correct Response
Category-Name Intraverbals	
"What is Virginia?"	"A state"
"What is Alabama?"	"A state"
"What is California?"	"A state"
"What is a Cardinal?"	"A bird"
"What is a Yellowhammer?"	"A bird"
"What is a Quail?"	"A bird"
"What is a Dogwood?"	"A flower"
"What is a Camellia?"	"A flower"
"What is a Poppy?"	"A flower"
Exemplar-Name Intraverbals	
"Name a state."	Virgina/Alabama/California
"Name another state."	Virgina/Alabama/California
"Name another state."	Virgina/Alabama/California
"Name a bird."	Cardinal/Yellowhammer/Quail
"Name another bird."	Cardinal/Yellowhammer/Quail
"Name another bird."	Cardinal/Yellowhammer/Quail
"Name a flower."	Dogwood/Camellia/Poppy
"Name another flower."	Dogwood/Camellia/Poppy
"Name another flower."	Dogwood/Camellia/Poppy

Intraverbal test with visual stimuli. This test addressed the possibility that the relevant intraverbal relations were in the participants' repertoires, but that they were partially under the control of the visual stimuli presented in VVCD trials. This test consisted of 9 trials and was conducted in a similar manner as the intraverbal test, except that a visual stimulus was presented to the participant. For example, in the original intraverbal test, one of the intraverbals presented was "Virginia goes with which bird?" In the intraverbal test with a visual stimulus, the experimenter held up a picture of Virginia and said, "If you see THIS, what do you pick?" There were two possible vocal responses that the child could emit (e.g., if the experimenter held up Virginia when presenting the vocal antecedent stimulus, then the child could respond with either Cardinal or Dogwood). Following the first vocal response from the child, the experimenter asked, "What else do you pick?" and waited for another response. Passing criterion was the same as in the original intraverbal test.

Results

Figure 5 shows data from testing sessions (i.e., equivalence and intraverbal tests) in a concurrent multiple probe design for the three participants who passed the equivalence test; Niels, Rita, and Emilie. During the intraverbal pre-test, none of the participants' performances were above chance level. Niels was exposed to one intraverbal pre-test and Rita and Emilie were exposed to three intraverbal pre-tests, thus controlling for history effects (e.g., a curriculum or personal change in a classroom setting), maturation, and/or exposure to the experimental protocol and environment. Following VVCD training, Niels passed equivalence testing but only made correct responses on 67% of the trials; therefore, criterion was not met on the intraverbal test. Additional rounds of testing were conducted in which performance on both tests continued to improve. Following the third round of
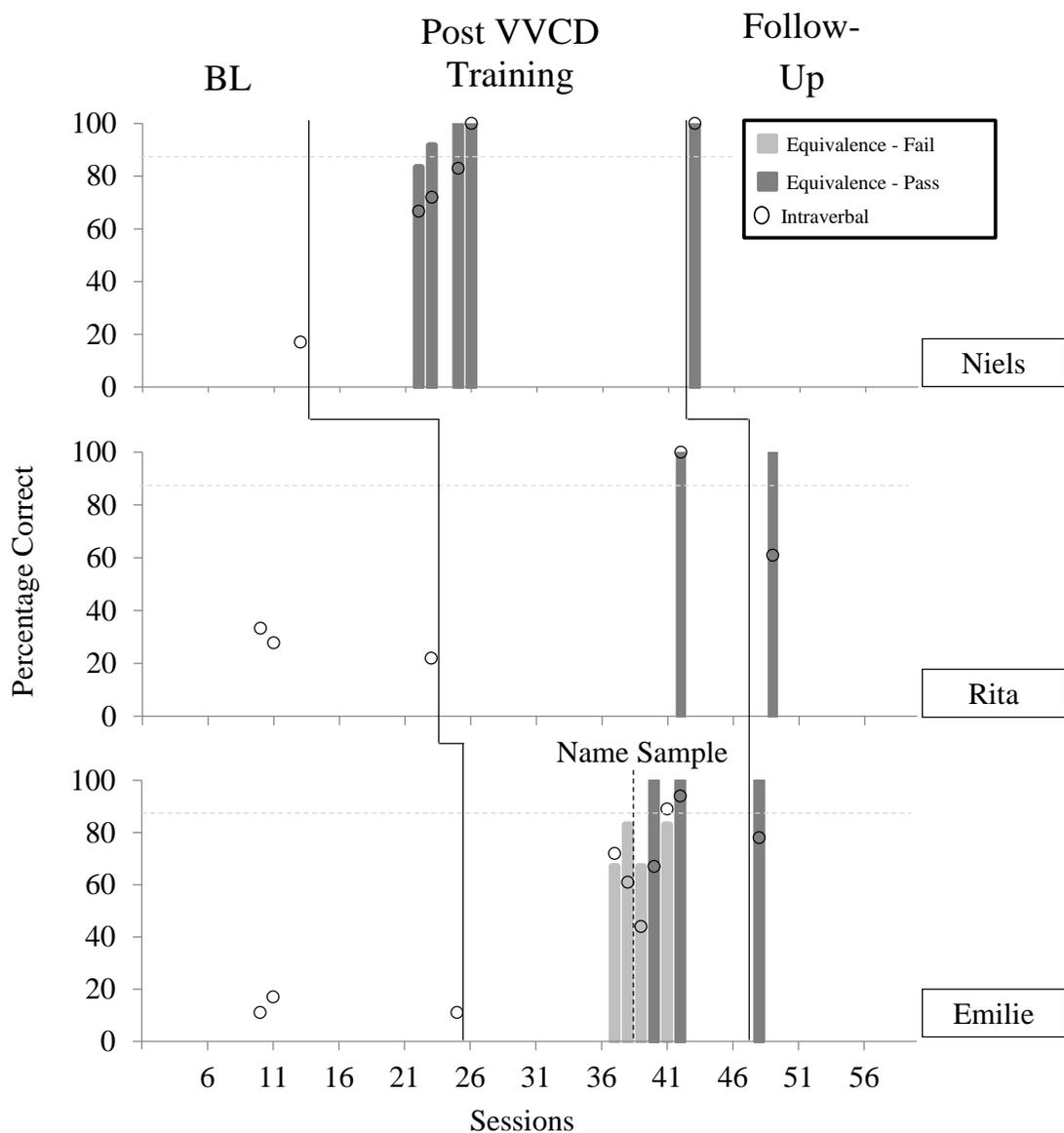
Figure 5. This figure shows the percentage of correct responses during test sessions for Niels, Rita, and Emilie. "BL" indicates baseline intraverbal pre-testing data, "Post VVCD Training" indicates equivalence and intraverbal test data following VVCD training, and "Follow-Up" indicates follow-up equivalence and intraverbal test data. The black line indicates the start of VVCD training. The initial "BL" testing session was preceded by tact training to criterion, and all subsequent "BL" testing sessions were preceded by tact maintenance blocks. The initial "Post VVCD Training" testing session was preceded by VVCD training to criterion. Open circles represent performance on the intraverbal test, light gray bars represent failing performance on the equivalence test, dark gray bars represent passing performance on the equivalence test, light gray horizontal dashed lines represent criterion performance for intraverbal tests, and the dark gray vertical dashed line represents when participants were required to name the sample stimulus during equivalence tests. With the exception of the pre-test, data are shown only for tests in which criterion was met for the baseline relations during equivalence testing.

additional testing, Niels made correct responses on 100% of the trials for both equivalence and intraverbal tests, and performance on both tests was maintained during follow-up testing. During the first round of testing following VVCD training, Rita met criterion on both equivalence and intraverbal tests. During follow-up testing, performance on the equivalence test was maintained, but Rita failed the intraverbal tests with correct responses on 67% of the trials. Emilie did not pass equivalence testing following VVCD training and only made correct responses on 72% of trials for the intraverbal test; therefore, criterion was not met on either test. An additional round of testing was conducted in which equivalence performance increased but did not met criterion, and intraverbal performance decreased to 61%. Because criterion was not met on either test, more additional rounds of testing were conducted in which Emilie was required to name the sample stimulus during the equivalence test. Following the requirement to name the sample stimulus, performance on both tests decreased. That is, performance on the equivalence test decreased to previous levels, and performance on the intraverbal test decreased below previous levels to 41%. Following additional rounds of testing however, performance on both tests continued to increase and criterion was met for both tests following the third round of additional testing when Emilie was required to name the sample stimulus. During follow-up testing, performance on the equivalence test was maintained, but Emilie failed the intraverbal tests with correct responses on 77% of the trials.

Figure 6 shows data from testing sessions (i.e., equivalence, symmetry, and intraverbal tests) in a non-concurrent multiple probe design for the participants who never passed the equivalence test; Charles, Marie, and Leonardo. In the intraverbal pre-test, none of the participants' performances were above chance level. Charles was exposed to one
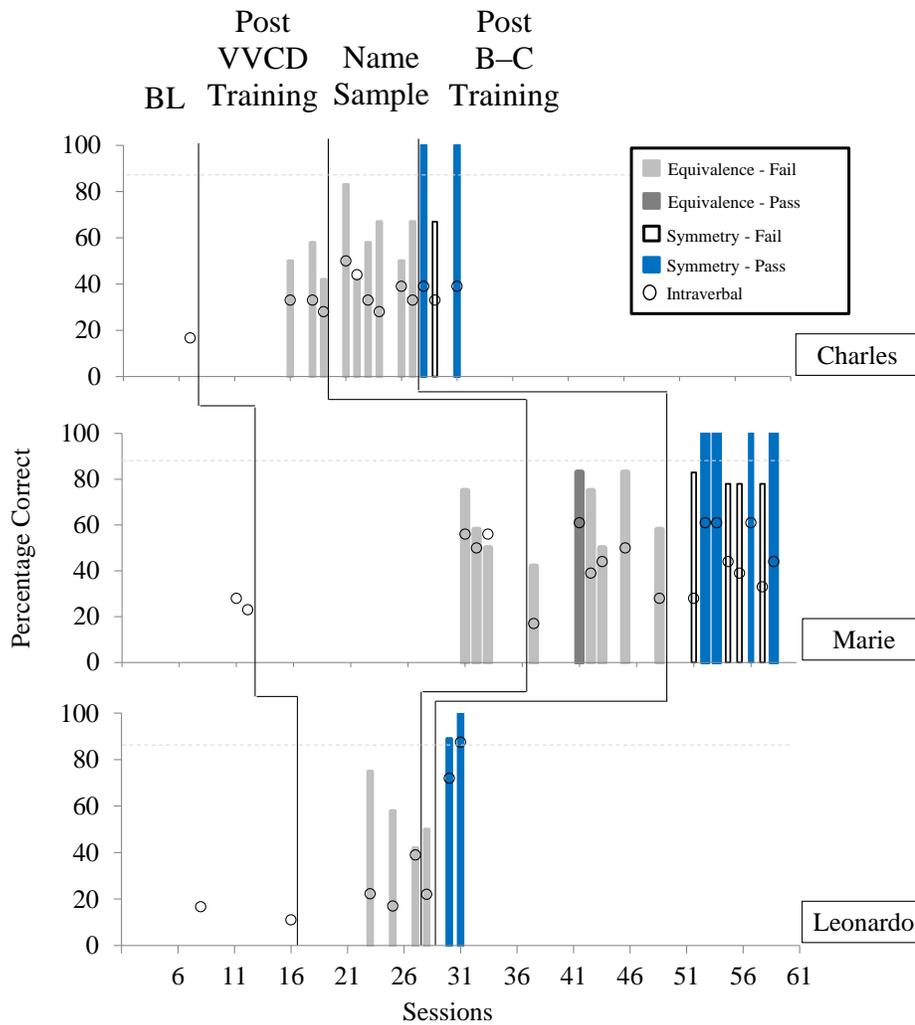
Figure 6. This figure shows the percentage of correct responses during test sessions for Charles, Marie, and Leonardo. "BL" indicates baseline intraverbal pre-testing data, "Post VVCD Training" indicates equivalence and intraverbal test data following VVCD training, "Name Sample" indicates equivalence and intraverbal test data when participants were required to name the sample stimulus, and "Post B-C Training" indicates equivalence and intraverbal testing data following B-C training. The first black line indicates the start of VVCD training, the second black line represents when participants were required to name the sample stimulus during equivalence tests, and the third black line represents the start of B-C training. The initial "BL" testing session was preceded by tact training to criterion, and all subsequent "BL" testing sessions were preceded by tact maintenance blocks. The initial "Post VVCD Training" testing session was preceded by VVCD training to criterion. Open circles represent performance on the intraverbal test, light gray bars represent failing performance on the equivalence test, dark gray bars represent passing performance on the equivalence test, open bars represent failing performance on the symmetry test, blue bars represent passing performance on the symmetry test, and light gray horizontal dashed lines represent criterion performance for intraverbal tests. With the exception of the pre-test, data are shown only for tests in which criterion was met for the baseline relations during equivalence testing.

intraverbal pre-test and Marie and Leonardo were exposed to two intraverbal pre-tests.

Following VVCD training, none of the participants met criterion for either the equivalence or

intraverbal test following three additional test sessions. Furthermore, performance on

equivalence tests for Marie and Leonardo gradually decreased over testing sessions. A slight

improvement in equivalence test performance was demonstrated during the second round of

testing for Charles; however, in the third testing session, performance decreased to levels

below the initial testing session. Intraverbal performance for the first three testing sessions

remained stable for Charles and Marie at around 33% and 56%, respectively. Intraverbal

performance for Leonardo decreased from 22% to 17% in the second test session following

VVCD training, and then increased to 39% in the third test session. Because criterion was not

met for either tests and performance on either test was not gradually increasing, all three

participants were required to name the sample stimulus in additional rounds of testing.

Charles and Marie were exposed to six additional test sessions in which naming the sample

stimulus was required on equivalence tests. For Charles, performance on both tests increased

for the first test session in this phase, however, the increase was not maintained and

performance decreased back to levels similar to testing sessions following VVCD training

(M=37.8% correct across the six intraverbal tests). Marie passed the equivalence test in the

second test session in this phase, but performance was not maintained. For Marie,

performance levels for both tests remained similar to levels in test sessions following VVCD

training (M=39.8 % correct across the six intraverbal tests). Due to time constraints,

Leonardo was only exposed to one test session requiring naming of the sample stimulus.

Performance levels for both tests remained similar to levels in test sessions following VVCD

training. As no participants met criterion in either test after being required to name the

sample stimulus and no gradual increasing trend was demonstrated for any of the participants, all participants were exposed to training on the B-C relation. Following B-C training, all participants were exposed to symmetry tests, which all participants consistently passed. For Charles and Marie, performance levels on intraverbal tests remained consistent with previous phases and were 39% correct (M= 37%) and 44% correct (M= 46.4%) respectively at the end of the evaluation. For Leonardo, performance on intraverbal tests increased from 72% to 87.5% in which criterion was met during that second test session following B-C training. For all three participants, only testing data in which criterion was met for the trained relations during equivalence testing was included in the graph; criterion was not met for the trained relations during two equivalence tests for Charles and one equivalence test for Leonardo.

Figure 7 shows data from intraverbal tests for Niels, Rita, and Emilie, broken down into the type of relation tested. For Niels, performance on baseline-like and symmetry-like relations remained above 83% for all five intraverbal tests. Performance on the transitivity-like relations was at 33% correct for the first intraverbal test, then increased to 50% correct for the following two intraverbal tests. On the fourth intraverbal test, performance on the transitivity-like relations increased to 100% correct and remained at that level during the follow-up test. For Rita, performances for all relations were at 100% correct on the first intraverbal test; however, on the follow-up test, performance for the baseline-like and transitivity-like relations dropped to 67% correct and performance for the symmetry-like relations dropped to 50% correct. For Emilie, performance on the baseline-like relations was at 100% correct for test one, followed by a decrease to 17% correct on test three. Following test three, performance increased and tests 5 and 6 were at or above 83% correct, followed by
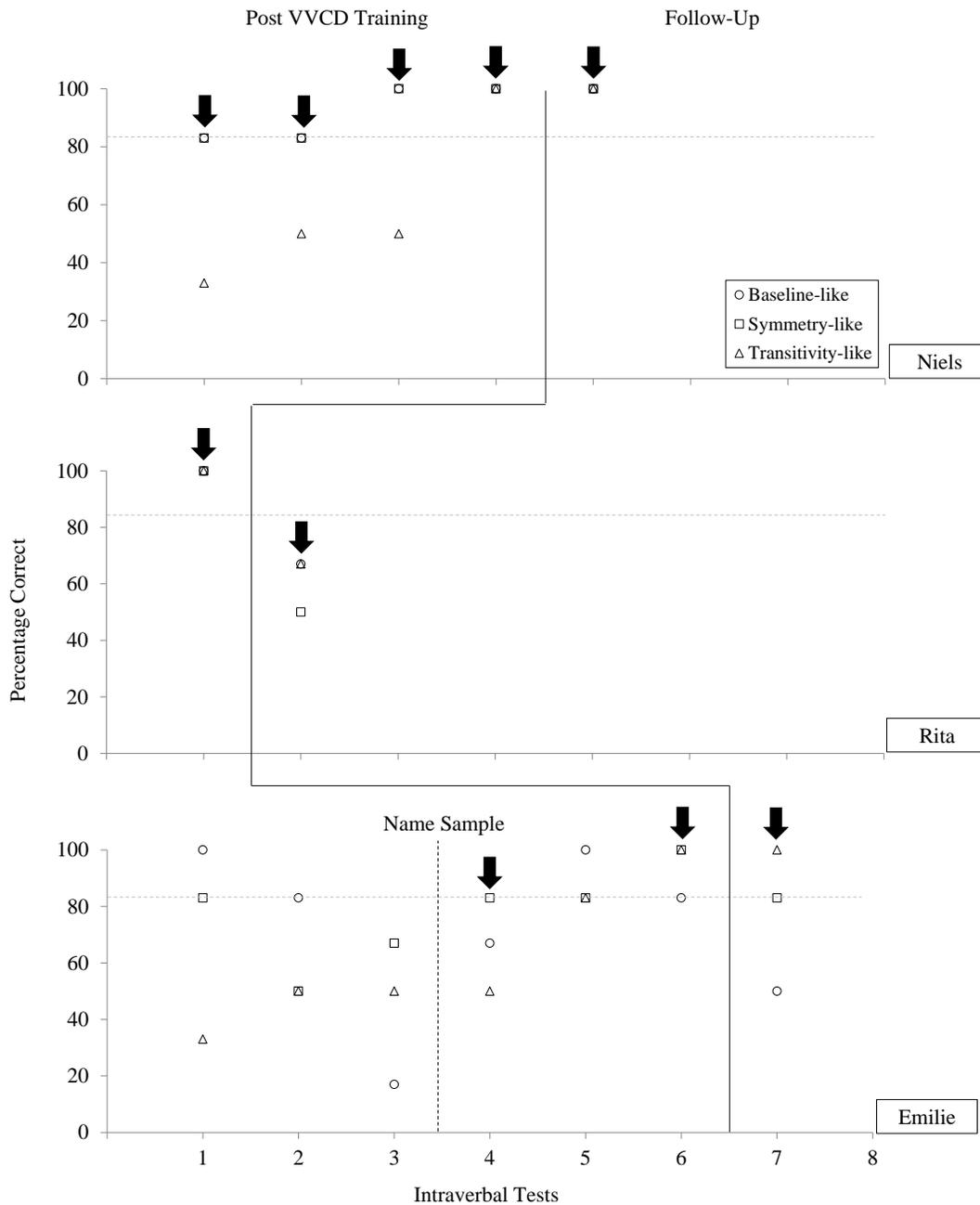
Figure 7. This figure shows the percentage of correct responses during intraverbal tests for Niels, Rita, and Emilie. Circles represent baseline-like relations, squares represent symmetry-like relations, and triangles represent transitivity-like relations as stated in the legend. The dashed line indicates when the naming requirement was added to equivalence tests. Black arrows indicate intraverbal test that followed a passing performance on the prior equivalence test. Light gray horizontal dashed lines represent criterion performance for intraverbal tests. Baseline intraverbal tests are not presented, and data are shown only for tests in which criterion was met for the trained relations during equivalence testing.

50

a decrease to 50% on the follow-up test. Figure 7 also shows that except in the case of Rita's follow-up test, passing performance on the equivalence test was always followed by passing performance on either baseline-like or symmetry-like relations on the intraverbal test; in most cases, passing performance was observed on both types of relations.

Figure 8 shows data from intraverbal tests following VVCD and B-C training for Charles, Marie, and Leonardo by the type of relation tested. For Charles, performance on all relations was highly variable. The highest performance level was demonstrated with the symmetry-like relations ranging from 33–83% correct. The next highest performance was demonstrated with the baseline-like relations ranging from 17–67% correct. Performance on transitivity-like relations demonstrated the lowest levels ranging from 0–33% correct. For Marie, performance on baseline-like relations were highly variable starting at 50% correct for test one, decreasing to 17% correct on test four, increasing to 83% correct for test five, followed by a decrease in performance back to 17% correct on the first test following B-C training (i.e., test 10), increasing back up to 83% correct for test 15, followed by another decrease back to 17% correct for test 17. Performance on symmetry-like relations remained fairly constant at 83% correct until test five. Following test five, performance gradually decreased to 33% correct on the first test following B-C training (i.e., test 10), then increased back to 83% correct on test 13, then decreased and remained stable around 50% correct until the end of the evaluation. Performance on transitivity-like relations remained low throughout the evaluation consistently ranging from about 0% - 33% correct. Performance levels on transitivity-like relations peaked above 33% correct on tests 1, 6, 9, 13, and 15. For Leonardo, performance on baseline-like and symmetry-like relations remained constant around 17% and 33% correct until tests following B-C training. Following B-C training,
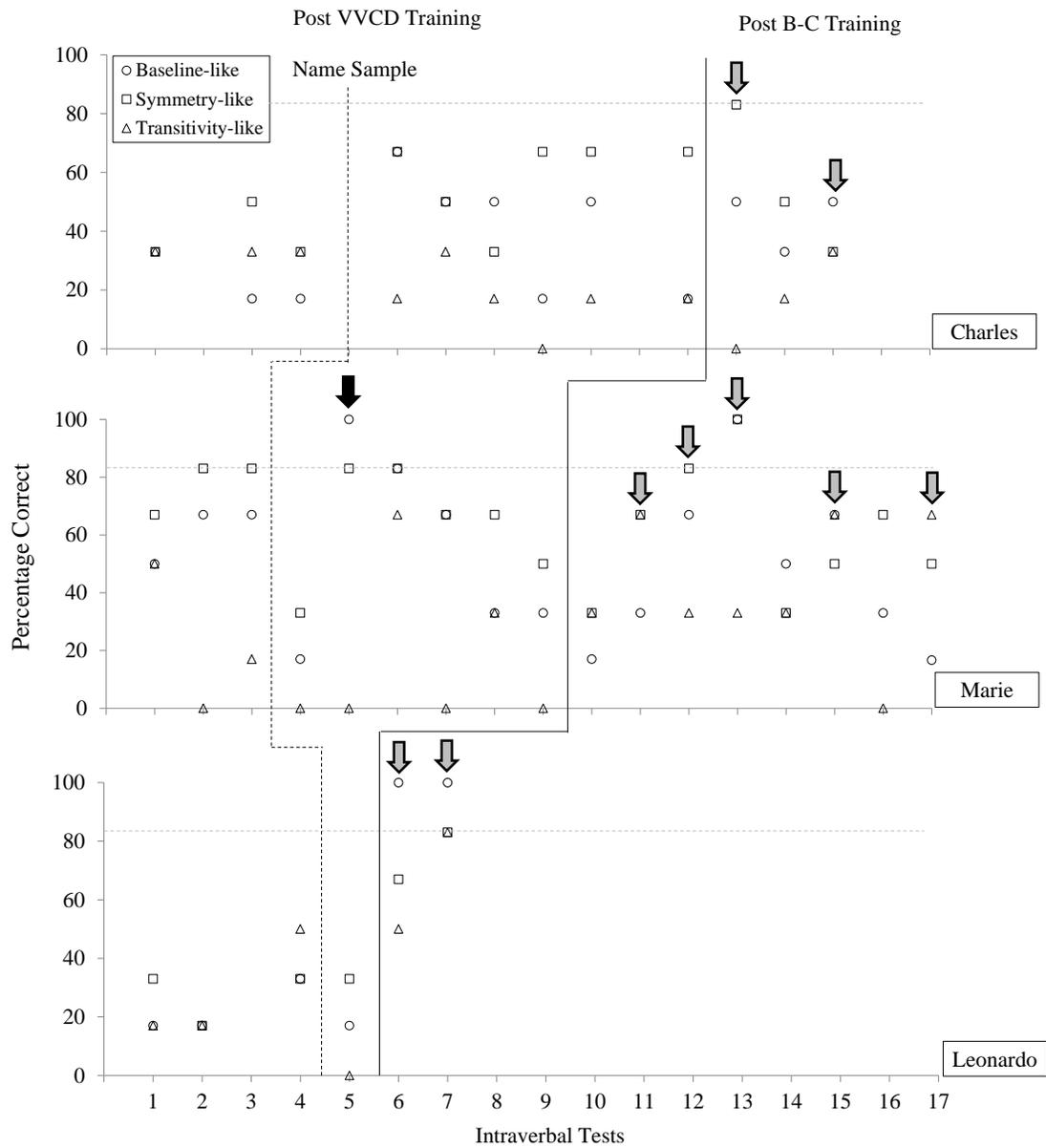
51

Figure 8. This figure shows the percentage of correct responses during intraverbal tests for Charles, Marie, and Leonardo. Circles represent baseline-like relations, squares represent symmetry-like relations, and triangles represent transitivity-like relations as stated in the legend. The dashed line indicates when the naming requirement was added to equivalence tests. Black arrows indicate intraverbal test that followed a passing performance on the prior equivalence test, and gray arrows indicate intraverbal test that followed a passing performance on the prior symmetry test following B-C training. Light gray horizontal dashed lines represent criterion performance for intraverbal tests. Baseline intraverbal tests are not presented, and data are shown only for tests in which criterion was met for the trained relations during equivalence and symmetry testing.

performance levels for baseline-like relations increased to 100% correct and remained at that level throughout the end of the evaluation, and performance on symmetry-like relations increased to 83% correct. Performance on transitivity-like relations began low at 17% correct on test one, then gradually increased to 83% correct following B-C training. Overall, Figure 8 shows that following VVCD training, the participants who consistently failed the equivalence test also showed failing performance on all three types of intraverbal relations as well, with the exception of a few of Marie's intraverbal tests in which she met the criterion for the emergence of symmetry-like intraverbals. In the test that followed, her only passed equivalence test, Marie met the criterion for the emergence of both baseline-like and symmetry-like intraverbals. Following B-C training, when participants passed the symmetry test, they typically did not meet the criterion for the emergence of both baseline-like and symmetry-like intraverbals, and in some cases met neither criterion.

Figure 9 shows data for symmetry and transitivity relations on failed equivalence tests prior to B-C training for Charles, Marie, and Leonardo. For all three participants, the number of correct responses was consistently higher for the symmetry relation than the transitivity relation. For Charles and Leonardo, the mean number of responses for the symmetry relation was 5 and the mean number of responses for the transitivity relation was 2. For Marie, the mean number of responses for the symmetry relation was 5 and the mean number of responses for the transitivity relation was 3. All three participants frequently met the criterion for the emergence of symmetry relations, but rarely for the emergence of transitivity relations. Thus, their failing performance on the intraverbal test was mostly due to below-criterion responding in transitivity trials.
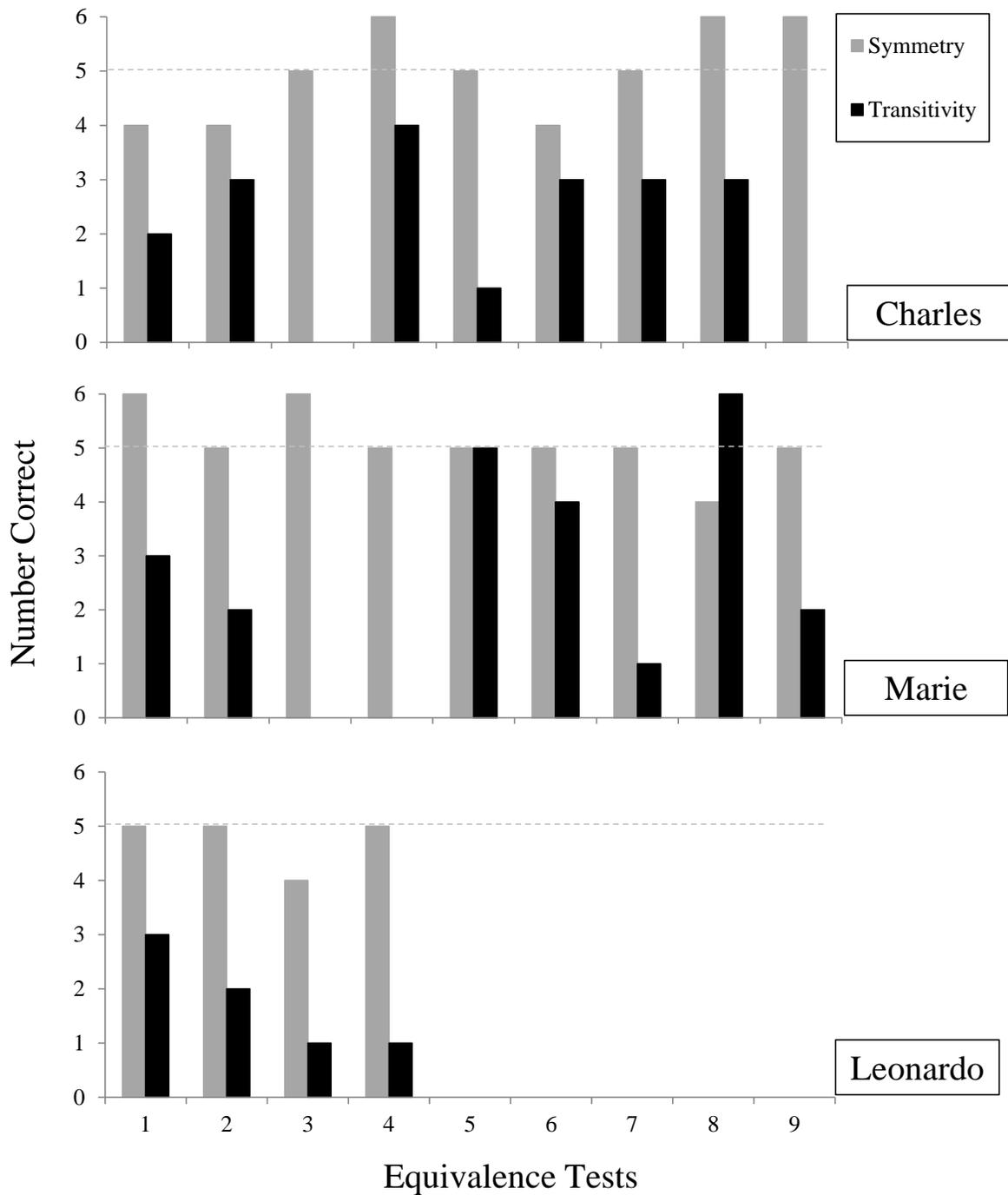
Figure 9. This figure shows data for symmetry and transitivity relations on failed equivalence tests prior to B-C training for Charles, Marie, and Leonardo. Light gray bars represent the symmetry relation and black bars represent the transitivity relation. Light gray horizontal dashed lines represent criterion performance for each relation on equivalence tests.
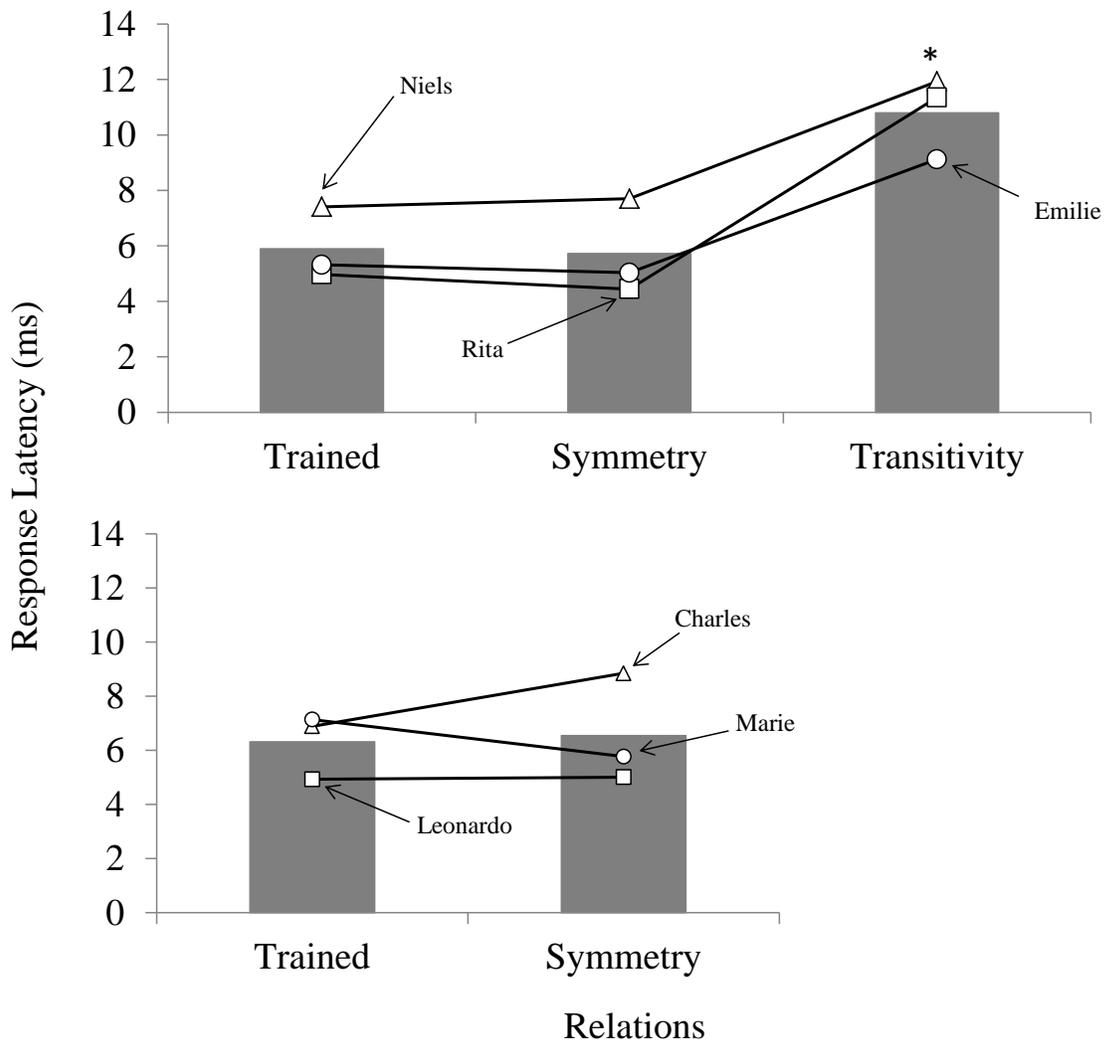
Figure 10. This figure shows response latency data (ms) in the upper panel on equivalence tests for Niels, Rita, and Emilie for each of the three relations, and in the lower panel on symmetry tests for Charles, Marie, and Leonardo for the symmetry and trained relations. Gray bars represent the overall mean response time for all participants. Open circles represent the mean response time for each participant. "*" indicate significance of $p < 0.05$. Data used in the analysis consisted of only tests in which passing performance was demonstrated on equivalence tests.

Figure 10 shows response latency data for (a) equivalence tests for participants who

passed the equivalence test (i.e., Niels, Rita, and Emilie) in the upper panel and (a) symmetry

tests for participants who passed the symmetry test and failed the intraverbal test (i.e.,

Charles

and Marie) in the lower panel. The computer started recording response latency data when

the participant clicked on the sample stimulus (which produced the comparison stimuli), and

ended when the participant selected a comparison stimulus. In the upper panel, a repeated-

measures one-way analysis of variance (ANOVA) was conducted to test for differences

between the amount of time it took for participants to select a comparison stimulus on

trained, symmetry, and transitivity trials during equivalence testing. The one-way ANOVA

revealed significant differences between the trial types, $F(2, 4) = 33.39$, $p < .05$. A Least

Significant Difference (LSD) pairwise comparison revealed that response latencies on

transitivity trials ($M = 10.80$ s) were significantly longer ($p < .05$) than response latencies on

trained ($M = 5.90$ s) and symmetry ($M = 5.72$ s) trials, whereas there was no significant

difference ($p = 0.54$) between response latency in trained and symmetry trials. Due to the fact

that the symmetry test consisted of only two relations (i.e., trained and symmetry), a paired

samples t-test on the means of the response latencies on trained ($M = 6.32$ s) and symmetry

($M = 6.54$ s) trials for Charles, Marie, and Leonardo during the symmetry tests were

conducted and failed to reveal any significant differences.

Figure 11 shows data from the listener tests for all six participants. All participants

met criterion and made 100% correct responses on the listener test. Figure 12 shows data

from the post-tests for participants who never passed the intraverbal tests (i.e., Charles and

Marie). Neither participant met criterion for the tact-intraverbal test, suggesting that the

substitution of a visual stimulus for a verbal stimulus did not facilitate correct verbal

Figure 11. This figure shows data from the listener tests for all six participants. Light gray horizontal dashed line represents criterion performance for listener tests.
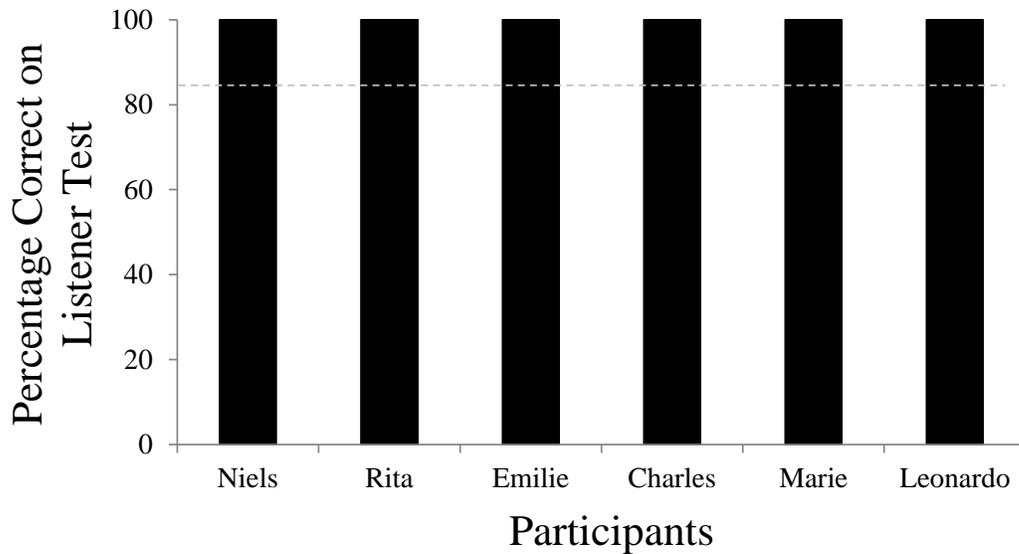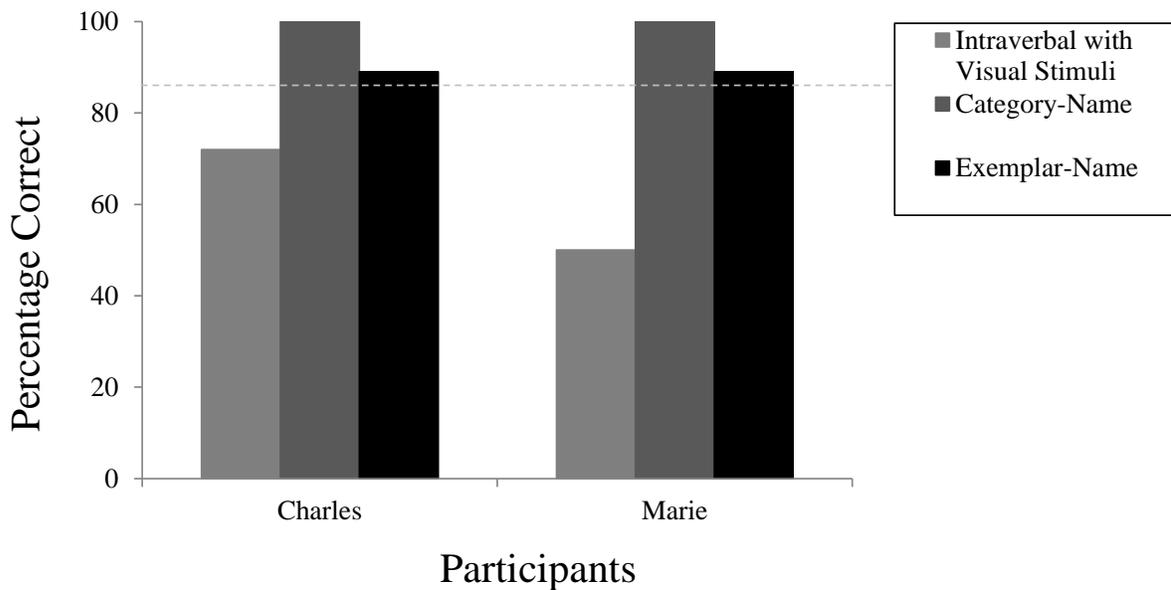

Figure 12. This figure shows data from the post-tests for participants whom never passed the intraverbal tests (i.e., Charles and Marie). Light gray bars represent the tact-intraverbal test,

dark gray bars represent the category-name test, and black bars represent the exemplar-name test as stated in the legend. The light gray horizontal dashed line represents criterion performance for all tests.

responses. Charles made correct responses on 72% of the trials and Marie made correct responses on 50% of the trials. Both participants met criterion on the category-name and exemplar-name tests with correct responses on 100% of the trials for the category-name test and 89% of the trials for the exemplar-name test, suggesting that they did not have trouble responding to the verbal stimuli "state". "bird", and "flower" that were presented during intraverbal test trials.

Figure 13 shows the percentage of correct responses during the last intraverbal pre-test and the first intraverbal test following VVCD training for Niels, Rita, and Emilie. For all three participants, performance on the first intraverbal test following VVCD training substantially increased compared to the last intraverbal pre-test. Performance levels increased from 17% to 67% for Niels, from 22% to 100% for Rita, and from 11% to 72% for Emilie. Figure 14 shows the percentage of correct responses during the last intraverbal pre-test and the first intraverbal test following VVCD training for Charles, Marie, and Leonardo. For all three participants, performance on the first intraverbal test following VVCD training only slightly increased compared to the last intraverbal pre-test. Performance levels increased from 17% to 33% for Charles, from 23% to 56% for Marie, and from 11% to 22% for Leonardo.

Figure 13. This figure shows the percentage of correct responses during the last intraverbal pre-test and the first intraverbal test following VVCD training for Niels, Rita, and Emilie.
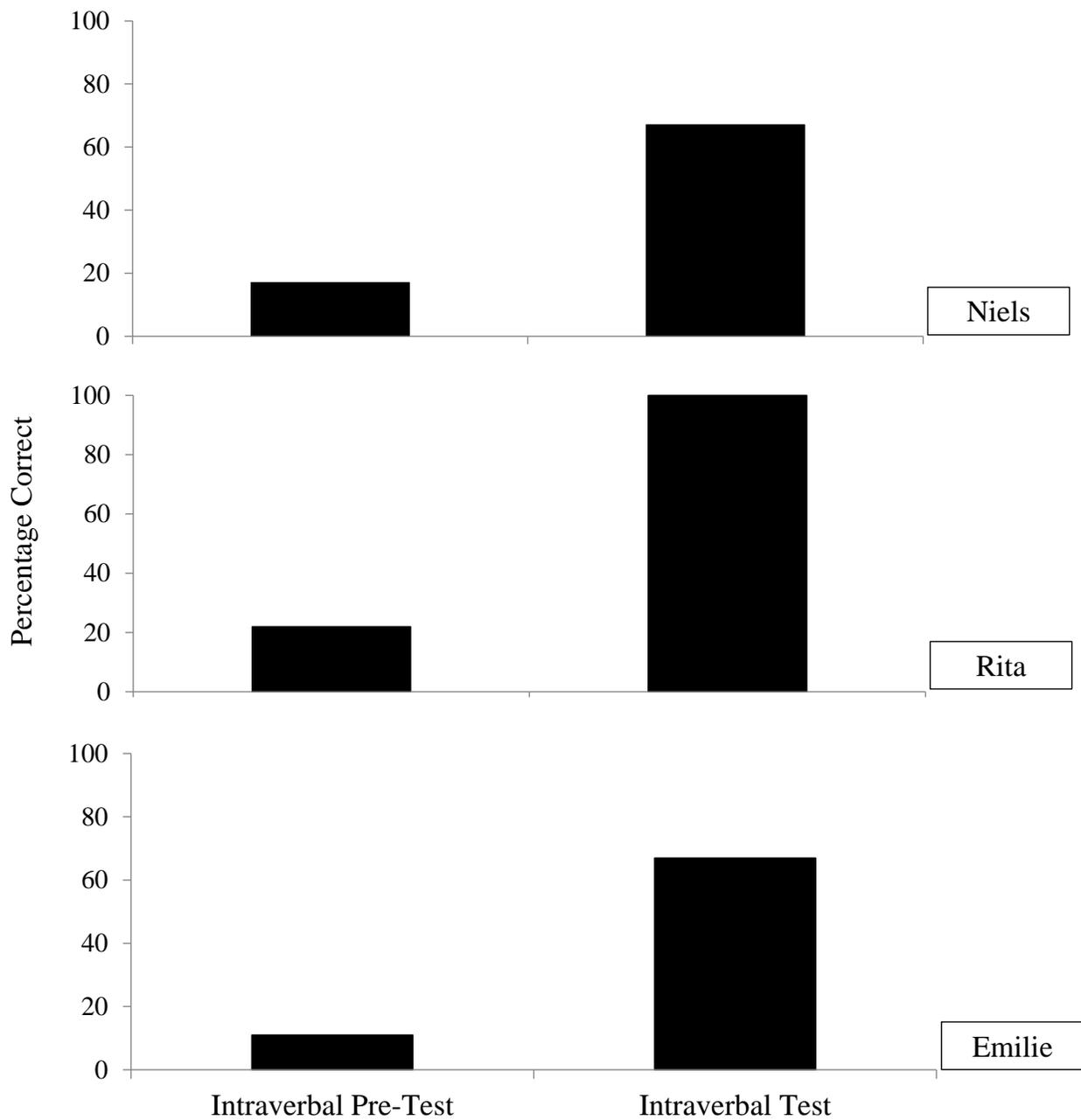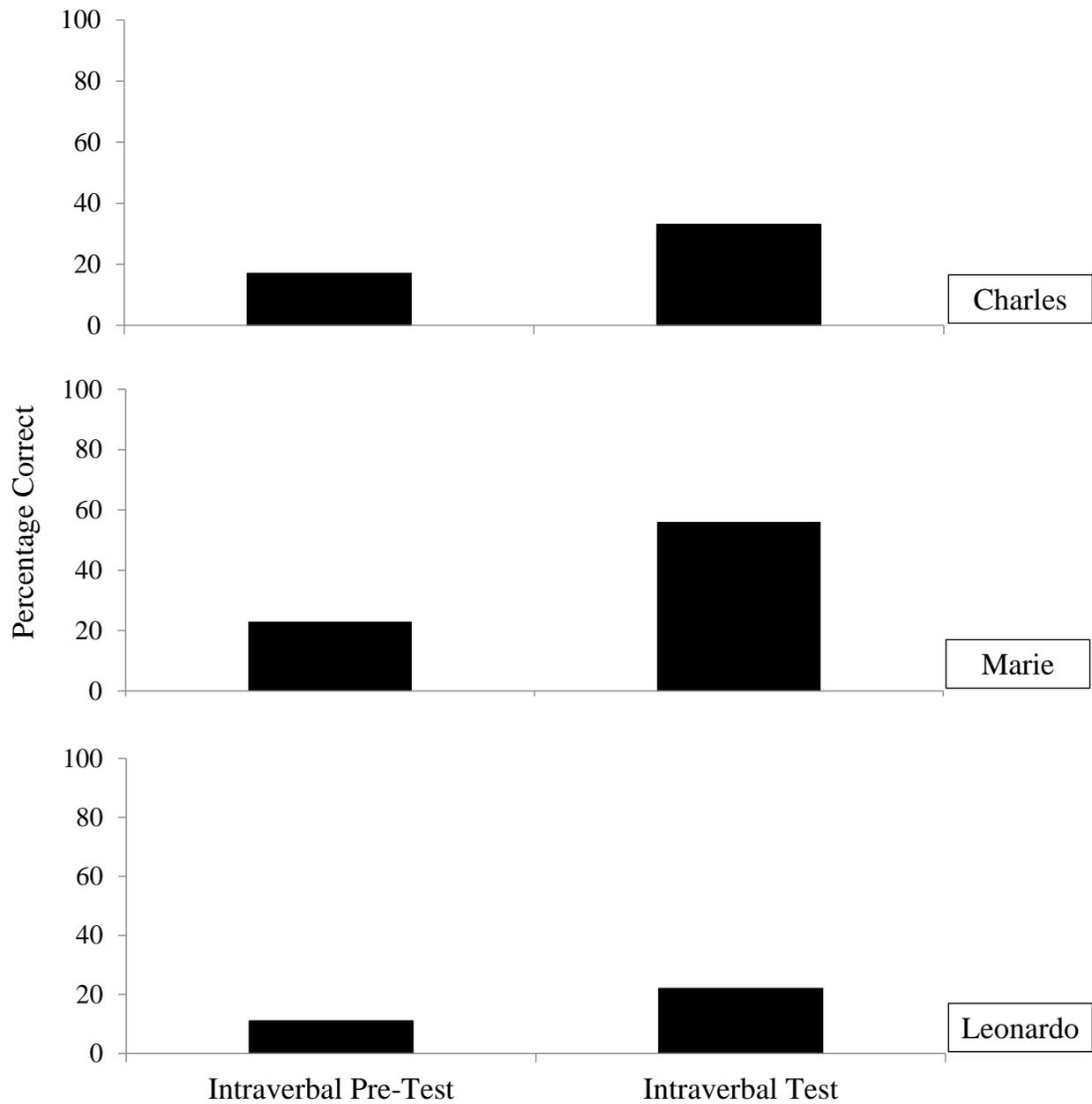
Figure 14. This figure shows the percentage of correct responses during the last intraverbal pre-test and the first intraverbal test following VVCD training for Charles, Marie, and Leonardo.

Summary of Results

        The purpose of the present study was to investigate Horne and Lowe's (1996) intraverbal naming hypothesis by evaluating the effects of visual-visual equivalence class formation on the emergence of novel intraverbals. Results showed two out of the six participants (i.e., Niels and Rita) passed the equivalence test following VVCD training. A third participant (i.e., Emilie) passed the equivalence test, but only after being required to name the sample stimulus during testing. The remaining three participants (i.e., Charles, Marie, and Leonardo) never passed the equivalence test, even after repeated rounds of testing with the original equivalence test and the version of the equivalence test in which they were required to name the sample stimulus. These results seem to be consistent with Horne and Lowe's verbal mediation account regarding the formation of equivalence classes for several reasons. First, all participants who passed the equivalence test passed the subsequent intraverbal test, and all participants who failed the equivalence test also failed the subsequent intraverbal test. Second, Emilie did not demonstrate passing performance on the intraverbal test until after she passed the version of the equivalence test that required her to name the sample stimulus. Finally, according to Horne and Lowe, correct responding on transitivity trials requires longer chains of intraverbals than does responding for either the trained or symmetry relations. The response latency data for the participants who passed equivalence supports this notion by showing that participants took a significantly longer amount of time to select a comparison stimulus on transitivity trials than on trained and symmetry trials.

        Of the three participants who demonstrated equivalence, two (i.e., Niels and Emilie) showed delayed passing performance on the intraverbal compared to passing performance on

the equivalence test. Upon a closer examination of the data, however, passing performance on the equivalence test was typically accompanied by passing performance on baseline-like and symmetry-like intraverbals. These data are consistent with Horne and Lowe's (1996) intraverbal naming hypothesis as, according to Horne and Lowe, the baseline-like and symmetry-like intraverbals are technically the only ones necessary to mediate passing performance on the equivalence test.

Participants who did not pass the equivalence tests were exposed to B-C training followed by a subsequent symmetry test. All three participants passed the symmetry test, but only one of those participants passed the intraverbal test (i.e., Leonardo). For the other two participants (i.e., Charles and Marie), no type of intraverbal relation (baseline-like, symmetry-like, or transitivity-like) was consistently present following passing performance on the symmstery test.

In summary, the data from the equivalence and intraverbal tests are mostly consistent with Horne and Lowe's (1996) intraverbal naming hypothesis. However, because two participants passed the symmetry test seemingly without mediating intraverbals, the data also suggest that verbal behavior may play a greater role in performance on transitivity trials than symmetry trials, and may even prove unnecessary for passing performance in symmetry trials.

Evaluation of Horne and Lowe's (1996) Naming Hypothesis

Horne and Lowe's (1996) intraverbal naming hypothesis predicts that if participants pass a purely visual equivalence test, they have done so because the training contingencies have established the appropriate intraverbal relations needed to mediate their performance (i.e., that intraverbal naming is a mechanism that mediates equivalence). Therefore, if

62

participants pass the equivalence test they should also pass the subsequent intraverbal test. The results of the present study provide no evidence to disprove Horne and Lowe's account; no participants passed the equivalence test without evidence of the intraverbal relations in their repertoires. Instead, participants who passed the equivalence test passed the intraverbal test, and participants who failed the equivalence test also failed the intraverbal test. Although this does not necessarily mean that the intraverbal relations mediated performances on equivalence tests, there are several other aspects of the data that are also consistent with Horne and Lowe's account.

Dymond and Rehfeldt (2001) suggest that response latency data should be collected during equivalence testing because it provides an indirect way to assess whether or not participants are engaging in a mediating strategy (e.g., verbal behavior). According to Horne and Lowe, during trained and symmetry trials on equivalence tests, participants need to only engage in one verbal link. For example, when participants are presented with a trained (A-B) trial as in Figure 12, participants will: tact the sample stimulus (i.e., "California") either covertly or overtly, hear themselves say "California," engage in the verbal link (i.e., "California goes with quail"), hear themselves say "quail," then select the appropriate comparison stimulus as a listener. The same process occurs during a symmetry trial, due to the training contingencies establishing the verbal link as bidirectional. However, on a transitivity trial, participants will need to engage in two possible verbal links. For example, when participants are presented with a transitivity (B-C) trial as in Figure 12, participants will: tact the sample stimulus (i.e., "quail") either covertly or overtly, hear themselves say "quail," engage in the first verbal link (i.e., "quail goes with California"), which then allows

them to engage in the second verbal link (i.e., "California goes with poppy"), they hear themselves say "poppy," and then select the appropriate comparison stimulus as a listener.

Therefore, if participants are engaging in verbal behavior during equivalence test trials, then transitivity trials, in theory, should result in longer response latencies than trained and symmetry trials. This is consistent with the results from the latency data for participants who passed the equivalence test, which demonstrated a significant different in response latency scores for transitivity trials compared to both trained and symmetry trials, but no significant differences between the trained and symmetry trials. Further, these findings are consistent with previous research demonstrating longer response latencies in transitivity trials than symmetry trials (e.g., Arntzen & Holth, 1997; 2000; Arntzen & Lian, 2010; Eilifsen & Arntzen, 2009; Holth & Arntzen, 1998, 2000; Saunders, Chaney, & Marquis, 2005; Spencer & Chase, 1996), and are consistent with the hypothesis that verbal behavior may be mediating performances on equivalence tests.

Niels and Emilie both passed the equivalence test several times before criterion was met on the intraverbal test. As stated previously, Horne and Lowe's (1996) intraverbal naming hypothesis only requires that participants have acquired the trained and symmetry-like intraverbals for passing performance on equivalence tests. For example, when presented with a B-C trial (i.e., a transitivity trial), as in Figure 15 during equivalence testing, one way participants can select the appropriate comparison stimulus is by engaging in the correct transitivity-like intraverbal. Participants can tact the sample stimulus (i.e., "quail") either covertly or overtly, hear themselves say "quail," engage in the transitivity-like verbal link (i.e., "quail goes with poppy"), and then select the appropriate comparison stimulus as a listener. Engaging in transitivity-like intraverbals is not the only way that verbal mediation
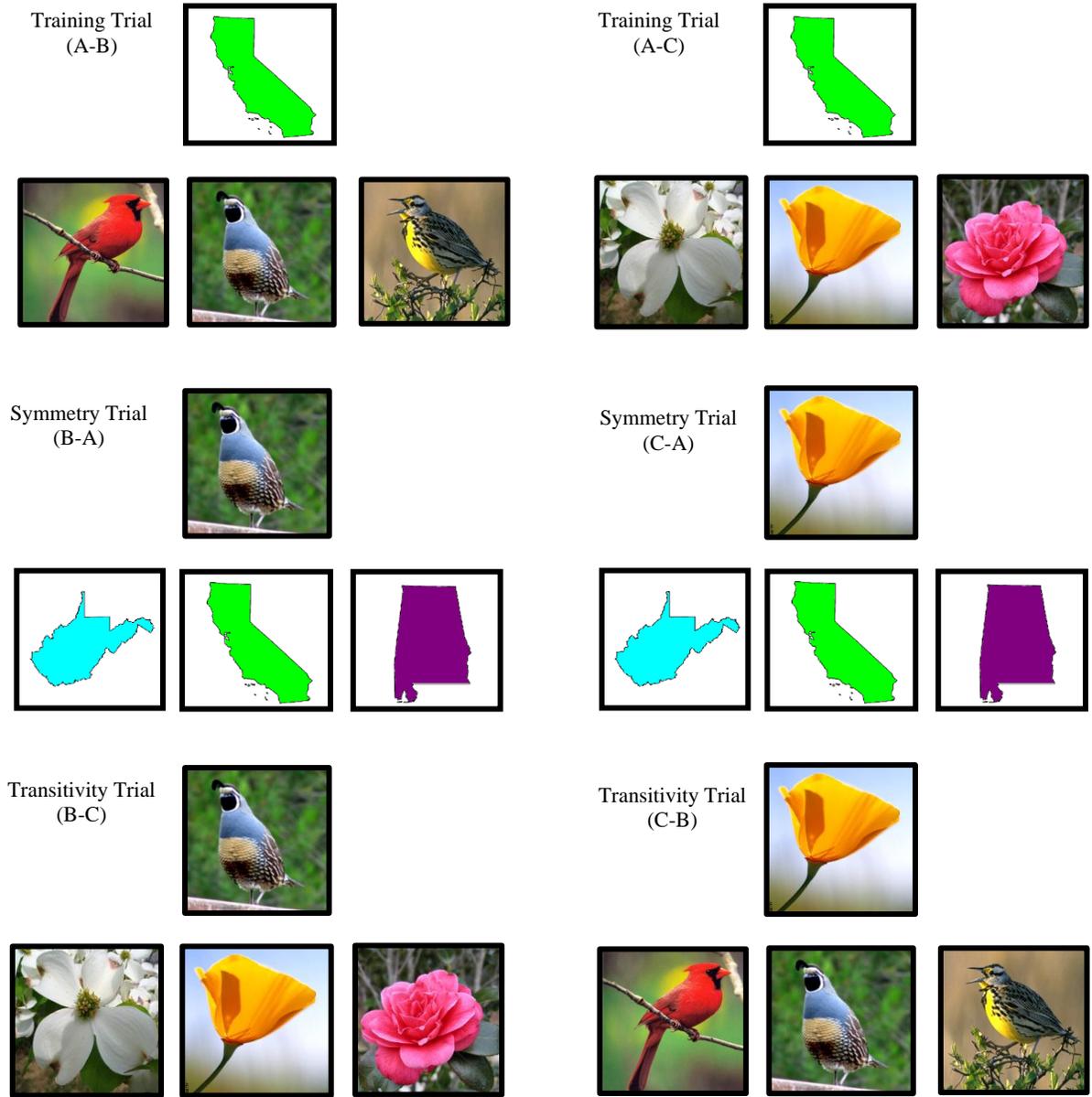
Figure 15. This figure shows an example of a trial for each relation (A-B, A-C, B-A, C-A, B-C, and C-B) presented during equivalence tests.

can result in the selection of the correct comparison stimulus, that is, participants can also select comparison stimulus by engaging in only the baseline-like and symmetry-like intraverbals. Therefore, on that same B-C trial, participants can tact the sample stimulus (i.e., "quail") either covertly or overtly, hear themselves say "quail," engage in a symmetry-like verbal link (i.e., "quail goes with California"), which then allows them to engage in the second baseline-like verbal link (i.e., "California goes with poppy"), they hear themselves say "poppy," and then as a listener select the appropriate comparison stimulus. This suggests that if participants fail the intraverbal test due to the absence of the transitivity-like intraverbals in their repertoires, then verbal mediation could still occur, as Horne and Lowe predict, if performance on the baseline-like and symmetry-like intraverbals is high.

Niels demonstrated three rounds of testing in which passing performance on the equivalence test was followed by a failing performance during the subsequent intraverbal test. A more in-depth analysis of the three intraverbal tests shows that Niels met criterion on the baseline-like and symmetry-like intraverbals during those three intraverbal tests: passing performance for both intraverbals was at 83% for the first two tests and at 100% for the third test (see Figure 7). Therefore, both baseline-like and symmetry-like intraverbals were present when the equivalence tests were passed. Emilie's first passed equivalence test (i.e., test 4 in Figure 7) showed passing performance only on the symmetry-like intraverbals and not the baseline-like intraverbals. This could be inconsistent with Horne and Lowe's (1996) intraverbal naming hypothesis because Horne and Lowe requires that both baseline-like and symmetry-like intraverbals be present to mediate transitivity trials on equivalence tests. However, performance on baseline-like intraverbals for Emilie was well above chance level, and it is possible that other uncontrolled factors resulted in her not reaching passing criterion

for these intraverbals on that test. Furthermore, the data (see Figures 7 and 8) show that participants rarely failed the equivalence test and then met the criterion for both the baseline-like and symmetry-like intraverbals during the subsequent intraverbal test. This happened twice for Emilie in tests 1 and 5, but never occurred for any other participants.

On the other hand, Rita and Emilie passed the follow-up equivalence test without showing passing performance on the baseline-like and symmetry-like intraverbals. This brings up the possibility that verbal mediation may be necessary for the formation of equivalence classes in children, but not their maintenance. Horne and Lowe's (1996) account speaks only to the initial formation of equivalence classes and does not necessarily exclude the possibility that once formed, equivalence classes may be maintained in the absence of verbal mediation.

Four participants did not pass the initial unmodified equivalence test (i.e., Emilie, Charles, Marie, and Leonardo). When this occurred, participants were exposed to additional equivalence tests in which they were required to name the sample stimulus during testing. It has been shown that participants who fail equivalence tests after learning common names for the stimuli in each class will pass the tests when prompted to name sample stimuli (using the common name) during testing (Horne et al., 2006; Horne et al, 2004; Lowe et al., 2002; Lowe et al., 2005). According to Horne and Lowe (1996), this occurs because naming the stimuli may prompt participants to engage in the appropriate verbal behavior. If naming the sample stimulus during testing does, in fact, prompt the appropriate verbal behavior, then we should expect to see an increase in equivalence test performance only for those participants who have the intraverbals in their repertoires. Alternatively, it is possible that naming the sample stimulus during testing simply increases attention to the stimuli. In this case, we

should expect to see an increase in equivalence test performance regardless of whether or not the intraverbals are in the participants' repertoires; therefore, we may see an increase to passing performance in the equivalence test, but no increase in performance for the intraverbal test.

Consistent with the former possibility, Charles, Marie, and Leonardo, who never passed the intraverbal test, did not benefit from the addition of the naming requirement to the equivalence test. On the other hand, Emilie passed the first equivalence test when required to name the sample stimulus, even though she did not pass the subsequent intraverbal test. Figure 7, however, shows that she met passing criterion for baseline-like and symmetry-like intraverbals during the first intraverbal test. It is possible that the intraverbal relations were in her repertoire, but were fragile and not fully established.

Finally, two participants (i.e., Charles and Marie) passed the symmetry test, but still continued to fail the subsequent intraverbal test. According to Horne and Lowe's (1996) naming hypothesis, the VVCD training contingencies should establish bidirectional intraverbal relations. If participants passed symmetry, then they would be expected to also demonstrate passing performance on the baseline-like and symmetry-like intraverbals. Horne and Lowe do not explicitly address what role verbal mediation plays in the emergence of symmetry alone, and thereby only address the formation of equivalence classes. Regardless, Figure 8 details a number of instances in which the symmetry test was passed, but criterion was not met on either the baseline-like or symmetry-like intraverbals, suggesting that the symmetry relation was not verbally mediated at all. Charles (i.e., the last intraverbal test) demonstrated this once and Marie (i.e., tests 11, 15, and 17) demonstrated this three times. Furthermore, Figure 9 evaluates the data on equivalence tests for the symmetry and

transitivity relations prior to B-C training for the participants who did not pass equivalence (i.e., Charles, Marie, and Leonardo). As these participants did not pass any of the intraverbal tests and never even demonstrated passing performance on both the baseline-like and symmetry-like intraverbals following equivalence testing, it is highly unlikely that verbal behavior was mediating performances on any of the relations during equivalence testing. Figure 9 demonstrates that in the absence of verbal behavior, all three participants made substantially more correct responses in symmetry trials than transitivity trials. The high number of correct responses in symmetry trials and low number in transitivity trials brings up the possibility that, for children, different mechanisms may mediate performance on symmetry and transitivity trials. Therefore, verbal mediation may be crucial or have a great facilitative effect on the transitivity relation and the formation of equivalence classes, but may be unnecessary for passing performance on symmetry trials.

Verbal Behavior and Symmetry

It is possible that performance on symmetry tests was in fact intraverbally mediated and the intraverbal test that was conducted in the present study simply failed to capture them. As Charles and Marie never passed the intraverbal tests following symmetry testing, a number of additional tests were conducted with these participants to determine whether any evidence exists that verbal behavior mediated performances on the symmetry tests.

First, during tact training, participants may have been emitting a specific vocal response in the presence of a specific visual stimulus (i.e., tacting; Skinner, 1957) and not engaging in the full *naming relation*. As stated earlier, the *naming relation* interlocks the tact, echoic, and listener behaviors. The full *naming relation* is necessary for the tact to have *meaning* (i.e., for the environment to establish listener, speaker, and echoic repertoires for a

stimulus) and it is what separates verbal behavior from other types of operant responses (Horne & Lowe, 1996). Typically, if a child can tact and respond as a listener to a stimulus, then naming has been established. As tact training was conducted at the beginning of the study, a listener test was given to all participants at the conclusion of the study to ensure that tact training had established *naming* and not just tacting. All participants scored 100% correct responding on the listener test (see Figure 11), ensuring that the *naming relation* for all stimuli had been established.

Second, it is possible that these participants assigned their own names (e.g., common names or their own specific names) to the stimuli, and thus, used their own names rather than the experimenter-provided names within the intraverbal relation. To assess this, an alternate tact test was conducted in which the experimenter would hold up a stimulus (e.g., California) and say, "What is this?" Both participants responded with the appropriate experimenter-provided names for all stimuli. Following the initial vocal response from the participant, the experimenter would then say, "Do you call it anything else?" The names that both participants most consistently responded with were the category names (e.g., "a state," "a bird," "a flower"). It is possible that Charles and Marie formed equivalence classes consisting of states, birds, and flowers via common naming (i.e., using the category names). However, these equivalence classes would be different from the ones promoted by the training contingencies established consisting of one state, one bird, and one flower, thereby, impeding performance on both the equivalence tests and the tests for symmetry alone. As both Charles and Marie ultimately passed the symmetry test, it is unlikely that their behavior was under control of the category names rather than the experimenter-provided names.

Third, it is possible that Charles and Marie had the intraverbal relations between the state, bird, and flower names, but they still failed the equivalence test because the category names "state," "bird," and "flower" failed to control the intraverbal responses. Therefore, Charles and Marie were exposed to a category-name and exemplar-name intraverbal test. In the category-name test, questions containing the exemplar names (e.g., California, quail, poppy, etc.) were presented to the child and the child's response consisted of the category names (i.e., state, bird, or flower). The experimenter would say for example, "What is California?" and the correct response from the participant would be the category name "a state." In the exemplar-name test, the experimenter's question contained the category names and the child's response consisted of an exemplar. For example, the experimenter would say "Name some states," and the correct response from the child would be one of the states (e.g., "California"). Both participants met criterion and passed both category- and exemplar-name tests. This establishes that failure on the intraverbal test was not due to the category names being unable to control the appropriate intraverbal responses.

Finally, it is possible that participants could emit the correct intraverbal responses upon being presented with a visual sample stimulus. Therefore, participants could emit the intraverbals during the symmetry test when the sample stimulus was present, but not during the intraverbal test when no sample stimulus was present. To investigate this possibility, Charles and Marie were exposed to an intraverbal test with visual stimuli. The intraverbal test with visual stimuli consisted of the experimenter holding up a stimulus (e.g., the picture of California) and saying, "When you see this, what do you pick?" Neither participant met criterion; however, Charles did demonstrate a higher level of responding on the intraverbal test with visual stimuli compared to the original intraverbal test suggesting that some verbal

behavior that was not captured in the original intraverbal test could have mediated performances on symmetry trials for him. Marie's level of responding was similar to the original intraverbal test. The results of the follow-up tests do not provide any consistent evidence that the intraverbal test in the following study could not have captured the appropriate mediating intraverbals and/or that verbal behavior in general mediated Charles and Marie's performances on symmetry tests.

The relationship between verbal behavior and symmetry is interesting because it is very challenging to find symmetry in animals, but it is easily demonstrated in humans. For example, symmetry has been demonstrated in children as young as two years of age (Boelens et al., 2000). This is interesting when taking into consideration the data from the present study in which all participants easily demonstrated symmetry, but it is clear that verbal behavior did not mediate that performance (see data from Charles and Marie in Figure 8). If verbal behavior does not mediate symmetry, then what does? Sidman (1994; 2000) suggested that symmetry is simply a natural outcome of the operant reinforcement contingency, thus rendering a search for mediating factors unnecessary. Another possibility stems from the fact that the stimuli presented on symmetry trials have already been presented (i.e., paired) together on the screen during training. Therefore, symmetry could be, in part, a product of Pavlovian processes.

If nonverbal processes are responsible for symmetry, then why is symmetry so difficult to demonstrate in non-humans? Lionello-DeNolf (2009) reviewed the animal literature on symmetry in MTS training and found 24 studies. Of those 24 studies, 11 failed to find any evidence of symmetry, 10 found some evidence of symmetry, and three found convincing evidence of symmetry. These studies were conducted across several different

species such as monkeys, chimpanzees, pigeons, baboons, rats, and sea lions. Frank and Wasserman (2005) have demonstrated the most convincing evidence for symmetry thus far with pigeons using clip-art stimuli. Pigeons were first trained on MTS tasks that included interspersed identity relations. For example, A-B training did not only include A-B trials, but also A-A and B-B trials, controlling for temporal variables (i.e., that a given stimulus could appear at either the beginning of a trial as a sample or in the middle of the trial as a comparison). During MTS training, the sample stimulus was presented for 10 s followed by the presentation of a comparison stimulus for 10 s, and peck rates during stimulus presentation were recorded. Half of the training trials were reinforced (i.e., the correct comparison stimulus would follow the sample stimulus followed by access to grain) and half were not (i.e., an incorrect comparison stimulus would follow the sample stimulus and there was no access to grain). During symmetry testing, unreinforced B-A probes were inserted into the training procedure. The two pigeons exposed to this procedure demonstrated more pecks to the comparison stimulus that was consistent with symmetry than to the comparison stimuli that were inconsistent with symmetry.

Lionelle-DeNolf (2009) concluded that the emergence of symmetry in animals might be a function of the training procedures, in that the choice of training type could discourage or encourage symmetry relations. In animal studies, symmetry is more likely to occur in tandem with procedures that intermix identity-matching relations into the training and include specific training that diminishes stimulus control by location variables, such as training with multiple locations (e.g., Tomonaga, Matsuzawa, Fujita, & Yamamoto, 1991) or using only one location (e.g., Frank & Wasserman, 2005). Interspersing identity-matching trials into the training procedures ensures that the animal can make successive and

simultaneous discriminations with all the stimuli involved (i.e., be able to discriminate the stimuli between trials as well as with in the trial) and controls for temporal variables (i.e., the animals learn that stimuli can appear first as a sample stimulus or second as a comparison stimulus). Therefore, the main difference between animals and humans (e.g., the children in the present study) may lie in their differing abilities to respond to the same stimulus regardless of its location in space. For animals, switching the role of the sample and comparison stimuli disrupts their performance. This does not seem to happen in humans, in that a favorite toy car is the same stimulus regardless of whether it is above eyesight, turned around, upside down, etc. Humans and animals have different prior experiences that may explain these differences, and procedures in animals that control and train these prior experiences seem to set the occasion for symmetry to be more likely to emerge.

Alternative Theoretical Explanations for Equivalence

If it is the case that symmetry is not mediated by verbal behavior but transitivity is, it is possible that verbal behavior may not mediate equivalence in exactly the same way as Horne and Lowe's (1996) Naming Hypothesis predicts. Although the present study overall does not demonstrate any major inconsistencies with Horne and Lowe's account, several minor inconsistencies emerge when examining Emilie's and Rita's data. As stated earlier, however, it cannot be ruled out that these inconsistencies are due to uncontrolled factors or that the intraverbals were in the participant's repertoire in a fragile way. Regardless, Emilie did pass two equivalence tests (i.e., tests 4 and 7 in Figure 7) in the absence of passing performance on both the baseline-like and symmetry-like intraverbals, and Rita passed the follow-up equivalence test in the absence of passing performance on all three types of intraverbals.

Another verbal mediation theory that has received less attention than Horne and Lowe's is Lowenkron's (1998; 2004) joint control account. According to Lowenkron (1998; 2004), joint control is established when a currently rehearsed topography of a verbal operant is simultaneously evoked by another stimulus, setting the occasion for an appropriate response. It is based on the effect of two discriminative stimuli acting together to control a common response topography. For example, if a person is looking for a specific telephone number in a phone book, that person may vocally rehearse the number. When a phone number is encountered that allows the rehearsed numbers to be entirely emitted as a series of tacts for the numbers in that phone number, the searching comes to an end because at that moment the person is saying the phone number under joint control. That is, saying the phone number is under control of both the rehearsal (i.e., self-echoic) and also the printed numbers on the page.

According to Lowenkron (1998; 2004), when participants in the present study were presented with a trained A-B trial as in Figure 12, they tacted the sample stimulus (e.g., "California") and then go through one of two routes to select a comparison stimulus. If the participant had the A-B intraverbal, then following the tact of the sample stimulus, they would engage in the verbal link ("California [goes with] quail") and then tact each comparison stimulus until one of them evoked the tact "quail." At this point they would make a selection response to indicate that they found a response under joint control (i.e., intraverbal and tact control). If the participant did not have the A-B intraverbal but had the B-A intraverbal, then following the tact of the sample stimulus, the participants would first tact the sample stimulus (e.g., "California") and then tact each comparison stimulus until one of the tacts evoked "California" as an intraverbal response (e.g., "quail [goes with] California").

Furthermore, B-C and C-B equivalence test trials only require the participants to have the symmetry-like intraverbals. In these trials, the participant could tact the sample stimulus "quail," emit the intraverbal "California," then tact the comparisons until one of the tacts also evokes the intraverbal response "California." This suggests that participants could pass the entire equivalence test by only having the symmetry-like intraverbals in their repertoires. Therefore, Lowenkren (1998; 2004) does not require that participants have *both* baseline-like and symmetry-like intraverbals, only that one is sufficient to mediate performance.

For all participants who passed equivalence, every equivalence test was followed by passing performance on either or both of the baseline-like or symmetry-like intraverbals. For the two testing sessions in which passing performance on equivalence tests was not followed by passing performance on *both* baseline-like and symmetry-like intraverbals (e.g., test four for Emilie in Figure 7), the equivalence test was followed by passing performance on the symmetry-like intraverbals. According to Lowenkren, this would still provide sufficient evidence that verbal behavior mediated performance.

The only possible inconsistency with Lowenkren's (1998; 2004) joint control account is found in Rita's follow-up tests demonstrating that equivalence was passed in the absence of passing performance for any of the three types of intraverbals. As this was only one test, we cannot rule out the possibility that her low performance on this test may be due to uncontrolled factors, as responding was still above chance level for all three intraverbal relations.

Another possibility is that equivalence classes emerge simply from the reinforcement contingencies (Sidman, 1994, 2000), and that verbal behavior may facilitate their formation but are not *necessary* for them to be established. This perspective by Sidman would account

for all of the inconsistencies during intraverbal tests following passing performances on equivalence tests, in that the emergence of intraverbal relations would be considered a byproduct of the equivalence class formation, not functionally related to it. Results from the present study suggest (a) that verbal behavior may be necessary for the transitivity relation but not the symmetry relations and (b) that verbal behavior may be necessary for the formation of equivalence classes but not the maintenance of the classes. It may be the case that both are necessary in different stages for the formation of equivalence classes for children. For example, both verbal behavior and the reinforcement contingency may be necessary to help bring selection responses on transitivity trials under appropriate stimulus control of the reinforcement contingencies. Then, once the participant's responses are under appropriate stimulus control, verbal behavior is no longer necessary, and the reinforcement contingencies are sufficient to maintain performance on equivalence tests.

Practical Implications

Early and intensive behavioral intervention (EIBI) treatment programs for children with developmental disabilities (e.g., children diagnosed with autism) allocate a large portion of treatment to language interventions that target both receptive and expressive skills. In a receptive protocol, a child responds non-vocally to a teacher's spoken instruction, whereas the goal of expressive protocols is to teach the child to emit spoken responses (Petursdottir & Carr, 2011). Many EIBI manuals suggests teaching receptive repertoires before expressive repertoires (e.g., Lovaas, 2003; Maurice, Green, & Luce, 1996; Leaf & McEachin, 1999); however, the empirical literature suggest that the reversed sequence may be more efficient (for a review see Petursdottir & Carr, 2011). For example, tact training (e.g., holding up a visual stimulus of a cat and saying "What is this?") often results in the emergence of listener

behavior (e.g., an auditory-visual conditional discrimination in which children are presented with an array of stimuli and instructed to "Point to the cat"), whereas listener training does not often result in the emergence of a tact repertoire. Thus, conducting expressive training first may reduce or eliminate the need for receptive training, and conducting receptive training first may increase the overall number of trials to establish both (i.e., receptive and expressive) repertoires. With regard to teaching categorization skills, EIBI curricula often suggests having the child perform visual-visual MTS tasks with the visual stimuli prior to teaching categorizational intraverbals (e.g., Leaf & McEachin, 1999; Maurice et al., 1996). For example, when aiming to teach a child vocal responses to categorizational intraverbals such as "Name some fruit," these manuals suggest having the child match visual stimuli of fruit together in a MTS task prior to intraverbal training. Phase 1 in teaching categories for the Leaf and McEachin (1999) manual is matching. The teacher puts out an array of pictures, one picture for each category, and labels each of the categories (e.g., a picture of a fruit, a picture of an animal, and a picture of a vehicle). Next, the student is presented with additional pictures (i.e.,, one additional picture at a time) and the teacher tells the child to "put with [fruit]." At the end of Phase 1, this visual-visual matching task is turned into a sorting task in which the child is presented with a pile of mixed up pictures and required to sort them into categories. No research has been conducted thus far that empirically demonstrates that having a prior history of visual-visual conditional discrimination training with visual stimuli may facilitate intraverbal training. On the surface this seems like a good suggestion, though the empirical evidence with regard to the teaching sequence of receptive and expressive protocols has typically not supported teaching receptive before expressive repertoires.

Although the present study did not directly evaluate the effects of conditional discrimination training (i.e., MTS tasks) with visual stimuli on the rate of acquisition of intraverbal training, it provides a starting point to evaluate whether novel intraverbals emerged following conditional discrimination training with the corresponding visual stimuli. Results from the present study demonstrate that several participants showed an increase in performance on intraverbal test following VVCD training compared to their performance on the intraverbal pre-test (see Figures 10 and 11); however, this was a small increase for participants who just passed the symmetry test. This is interesting in that the participants whose performances increased the most, were the participants who demonstrated equivalence. As the participants that demonstrated equivalence were older; therefore, it is possible that this effect is due to the participants' age. However, it is also possible that the best way to facilitate intraverbal training is not just to have the children perform MTS tasks with the visual stimuli, but for equivalence classes to form with the visual stimuli. Therefore, based on the results from the present study, VVCD training should not be expected to generate intraverbals unless it establishes equivalence classes.

Limitations

The main limitation of the present study is that half of the participants (i.e., three out of six) did not demonstrate visual-visual equivalence, which, in turn, did not allow a thorough examination of the effects of the naming requirement during equivalence testing. Other studies with similar participants have demonstrated similar failures in passing visual-visual equivalence tests (e.g., Smeets & Barnes-Holmes, 2005), though this is not what is typically demonstrated in the literature. Indeed, most previous studies utilizing children of similar ages or younger have readily demonstrated visual-visual equivalence (e.g., Barnes,

Browne, Smeets, & Roche, 1995; Carpentier, Smeets, & Barnes-Holmes, 2002; Goyos, 2000; Lazar et al., 1984; Sidman et al., 1986; Smeets & Barnes-Holmes, 2003). This discrepancy could be related to several issues in the present study.

First, the present study did not fully assess the participants' prerequisite skills prior to their participation in the study. The only pre-assessment conducted was the PPVT, which assessed participant's receptive skills. Although PPVT scores were comparable to each other, expressive skills were not assessed. As intraverbal behavior is an expressive skill, and as Horne and Lowe propose intraverbal behavior as one possible mechanism that mediates VVCD training, it would have been interesting to observe whether there were any differences between skill levels of expressive skills between the older and younger participants. A more thorough assessment of participants' skills could have resulted in the identification of certain prerequisite skills affecting the formation of equivalence classes. This could have applied significance in that clinicians would know what skills to train, so that the stimulus equivalence paradigm could be a more useful tool in language interventions.

Second, the stimuli that were used in the present study were nonarbitrary; therefore, participants could have had prior histories with the states, birds, and/or flowers that interfered with the training contingencies of the experiment. Several studies have shown that certain stimuli with which participants have a learning history could have an inhibitory effect on equivalence class formation (e.g., Holth & Arntzen, 1998b; Leslie, Tierney, Robinson, Keenan & Watt, 1993). For example, maybe a grandparent lives in California or maybe a favorite baseball team is the Saint Louis Cardinals. Horne and Lowe (1996) acknowledge that naming may not solve all difficulties persons may have on equivalence tests, they state that "intraverbal naming can work for or against success on tests of stimulus equivalence

depending on whether or not the intraverbal sequences that are formed before such testing are congruent with experimenter-defined classes" (p. 226). Therefore, whether or not names participate in classes requires a relevant prior and current history with the contingencies of reinforcement. Thus, a naming intervention may fail to facilitate class formation as often as it succeeds, due to additional factors such as the training and testing contingencies.

Third, the procedures in the present study incorporated a number of features in both training and testing. VVCD training was conducted in a one-to-many training structure in which one set of sample stimuli (i.e., the A stimuli) are presented with several sets of comparison stimuli (i.e., the B and C stimuli). There are several other training structures, however, and studies have shown that the formation of equivalence classes varies as a function of the training structure (e.g., Arntzen, 2004; Hove, 2003; Saunders et al., 2005; Saunders & McEntee, 2004). VVCD training also consisted of three comparison stimuli. Previous research has shown that the number of comparison stimuli (i.e., typically two, three, or four) affects which training procedure is more effective in establishing equivalence classes (e.g., Saunders et al., 1993, 1999; Sidman, 1994). It is also possible that the way the equivalence test was conducted affected the formation of equivalence classes. In the present study, the trained trials (i.e., A-B and A-C trials) were interspersed between non-trained trials and reinforced during the equivalence test. Previous research has shown that sometimes when reinforced trials are interspersed with non-reinforced trials equivalence is not demonstrated; however, when the same non-reinforced trials are then interspersed with different non-reinforced trained trials then equivalence is demonstrated (e.g., Lazar et al., 1984; Sidman et al., 1985). Furthermore, the fact that all relations (trained, symmetry, and transitivity) were presented during the equivalence test as opposed to testing for symmetry

and transitivity separately may have affected the emergence of equivalence (e.g., Smeets & Barnes-Holmes, 2005). One of these procedures or a combination of them could have rendered the formation of equivalence classes less easily established, as in some of the previous studies with children.

Finally, there was no VVCD pre-test conducted with the stimuli used in the present experiment before VVCD training. Because of this, it cannot be completely ruled out that the some of the VVCD relations trained in the present experiment had not already been established in the participants' repertoires. It is very unlikely, however, due the ages of the participants that these relations had already been acquired prior to their involvement in the study. In addition, all participants required extensive VVCD training before meeting the acquisition criterion for the baseline relations.

Conclusions and Future Directions

Horne and Lowe's (1996) naming hypothesis suggests two mechanisms by which verbal behavior can mediate the formation of equivalence classes: common naming and intraverbal naming. To date, most of the literature has focused on common naming, with only one study investigating intraverbal naming via an oral naming test following passing performance on an equivalence test (i.e., Smeets & Barnes-Holmes, 2005). The present study provides the literature with an experimental analysis of the intraverbal naming mechanism, and suggests that in children, verbal mediation may be a possible mechanism for equivalence, but not necessarily for symmetry. Therefore, transitivity may be more dependent on verbal behavior than symmetry. Additionally, the issue of verbal mediation in equivalence class formation vs. maintenance may need further investigation.

Future research might evaluate the effects of equivalence class formation via different training structures on the emergence of intraverbals. A training structure is used to refer to the sequence of conditional discrimination training and how the stimuli are linked during baseline training (Saunders & Green, 1999). The stimulus equivalence literature has developed three main training structures to establish equivalence classes: One-to-Many (OTM), Many-to-One (MTO), and Linear Series (LS). In the OTM training structure, the same sample is presented with different comparison stimuli (e.g., A-B and A-C training). In the MTO training structure, different sample stimuli are presented with the same comparison stimuli (e.g., B-A and C-A training). In the LS training structure, the comparison stimuli presented during the first training phase are presented as sample stimuli during the second training phase (e.g., A-B and B-C training). The present study implemented an OTM training structure (i.e., A-B and A-C training), and because of this specific training structure, the distinction between transitivity and equivalence (i.e., the B-C and C-B relations) cannot be determined. In the LS (i.e., A-B and B-C training), however, it is possible to make the distinction between a pure transitivity relation (A-C) and a pure equivalence relation (C-A). In stimulus equivalence studies with animals a LS training structure is typically used, and it seems that transitivity (in this case A-C) is easier to demonstrate than symmetry (e.g., Bunsey & Elchenbaum, 1996; D'Amato et al., 1985). In the present study, the transitivity relation (i.e., B-C and C-B) is more dependent on verbal behavior than the symmetry relation. Use of a LS training structure instead of an OTM training structure may allow a more precise investigation into which part of the formation of equivalence classes is verbal behavior necessary. The mechanism underlying a pure transitivity relation (i.e., A-C) may prove more similar to the mechanism underlying symmetry, a hypothesis consistent with

prior findings suggesting that the pure transitivity relation is more easily demonstrated in animals than symmetry (Donahue & Palmer, 2004). This would then leave verbal behavior as a possible mechanism behind the pure equivalence relation (i.e., C-A), would help close the gap in research findings in stimulus equivalence between animals and humans, and would also provide additional information regarding the relation between verbal behavior and the formation of equivalence classes.

REFERENCES

Arntzen, E. (2004). Probability of equivalence formation: Familiar stimuli and training

    sequence. *The Psychological Record, 54*, 275–291.

Arntzen, E., & Holth, P. (1997). Probability of stimulus equivalence as a function of training

    design. *The Psychological Record, 47*, 309–320.

Arntzen, E., & Holth, P. (2000). Equivalence outcome in single subjects as a function of

    training structure. *The Psychological Record, 50*, 603–628.

Arntzen, E., & Lian, T. (2010). Trained and derived relations with pictures versus abstract

    stimuli as nodes. *The Psychological Record, 60*, 659–678.

Barnes, D. (1994). Stimulus equivalence and relational frame theory. *The Psychological

    Record, 44*, 91–124.

Barnes, D., Browne, M., Smeets, P. M., & Roche, B. (1995). A transfer of functions and a

    conditional transfer of functions through equivalence relations in three- to six-year-

    old children. *The Psychological Record, 45*, 405–430.

Barnes, D., McCullagh, P. D., & Keenan, M. (1990). Equivalence class formation in non-

    hearing impaired children and hearing impaired children, *The Analysis of Verbal

    Behavior, 8*, 19–30.

Barros, R. S., Galvo, O. F., & MvIlvane, W. J. (2002). Generalized identity matching-to-

    sample in *Cebus apella. Psychological Record, 52*, 441–460.

Boelens, H., Van der Broek, M., & Van Klarenbosch, T. (2000). Symmetric matching to

    sample in 2-year old children. *The Psychological Record, 50*, 293–304.

Brady, N. C., & McLean, L. K. (2000). Emergent symbolic relations in speakers and

nonspeakers. *Research in Developmental Disabilities, 21*, 197–214.

Bunsey, M., & Elchenbaum, H. (1996). Conservation of hippocampal memory function in

rats and humans. *Letters to Nature, 379*, 255–257.

Catania, A. C. (2007). *Learning, fourth edition.* NY: Sloan.

Carpentier, F., Smeets, P. M., & Barnes-Holmes, D. (2002). Matching functionally same

relations: Implications for equivalence-equivalence as a model for analogical

reasoning. *The Psychological Record, 52*, 351–370.

Carr, D., Wilkinson, K. M., Blackman, D., & McIlvane, W. J. (2000). Equivalence classes in

individuals with minimal verbal repertoires. *Journal of the Experimental Analysis of

Behavior, 74*, 101–114.

D'Amato, M. R., Salmon, D. P., Loukas, E., & Tomie, A. (1985). Symmetry and transivity of

conditional relations in monkeys (*Cebus apella*) and pigeons (*Columba livia*).

Journal of the Experimental Analysis of Behavior, 44, 35–47.

Devany, J. M., Hayes, S. C., & Nelson, R. O. (1986). Equivalence class formation in

language-able and language-disabled children. *Journal of the Experimental Analysis

of Behavior, 46*, 243–257.

Dugdale, N., & Lowe, C. F. (1990). *Naming and stimulus equivalence.* In D. E. Blackman &

H. Lejeune (Eds.), Behavior analysis in theory and practice: Contributions and

controversies (pp. 115–138). Hove: Lawrence Erlbaum Associates.

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

Publishing.

Dymond, S., & Rehfeldt, R. A. (2001). Understanding complex behavior: The transformation of stimulus functions. *The Behavior Analyst, 23*, 239–254.

Eikeseth, S., & Smith, T. (1992). The development of fucnctional and equivalence classes in high-functioning autistic children: The role of naming. *Journal of the Experimental Analysis of Behavior, 58*, 123–133.

Eilifsen, C., & Arntzen, E. (2009). On the role of trial types in tests for stimulus equivalence. *Eurpoean Journal of Behavior Analysis, 10*, 187–202.

Elias, N. C., Goyos, C., Saunders, M., & Saunders, R. (2008). Teaching manual signs to adults with mental retardation using matching-to-sample procedures and stimulus equivalence. *The Analysis of Verbal Behavior, 24*, 1–13.

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

Fienup, D. M., Covey, D. P., & Critchfield, T. S. (2010). Teaching brain-behavior relations economically with stimulus equivalence technology. *Journal of Applied Behavior Analysis, 43*, 19–33.

Feinup, D. M., & Critchfield, T. S. (2010). Efficiently establishing concepts of inferential statistics and hypothesis decision making through contextually controlled equivalence classes. *Journal of Applied Behavior Analysis, 43*, 437–462.

Frank, A. J., & Wasserman, E. A. (2005). Associate symmetry in the pigeon after successive matching-to-sample training. *Journal of the Experimental Analysis of Behavior, 84*, 147–165.

Goyos, C. (2000). Equivalence class formation via common reinforcers among preschool

children. *The Psychological Record, 50*, 629–654.

Green, G. (1990). Differences in development of visual-visual and auditory-visual

   equivalence relations. *American Journal on Mental Retardation, 95*, 260–270.

Greer, R. D., & Ross, D. E. (2008). *Verbal behavior analysis: Inducing and expanding new*

   verbal capabilities in children with language delays. Boston: Allyn & Bacon.

Hayes, S. C., & Hayes, L. J. (1986). Verbal relations and the evolution of behavior analysis.

   American Psychologist, 47, 1383–1395.

Hayes, S. C., Barnes-Holmes, D., & Roche, B. (Eds.) (2001). *Relational frame theory: A*

   post-Skinnerian account of human language and cognition. New York: Plenum.

Herman, L. M., & Gordon, J. A. (1974). Auditory delayed matching in the bottlenose

   dolphin. *Journal of Experimental Analysis of Behavior, 21*, 19–26.

Herman, L. M., Hovancik, J. R., Gory, J. D., & Bradshaw, G. L. (1989). Generalization of

   visual matching by a bottlenose dolphin (*Tursiops truncates*): Evidence for invariance

   of cognitive performance with visual and auditory materials. *Journal of Experimental*

   Psychology: Animal Behavior Processes, 15, 124–136.

Holth, P., & Arntzen, E. (1998). Stimulus familiarity and the delayed emergence of stimulus

   equivalence or consistent nonequivalence. *The Psychological Record, 48*, 81–110.

Holth, P., & Arntzen, E. (1998b). Symmetry versus sequentially related to prior training,

   sequential dependency of stimuli, and verbal labeling. *The Psychological Record, 48*,

   293–315.

Holth, P., & Arntzen, E. (2000). Reaction times and the emergence of class consistent

   responding: A case for precurrent responding? *The Psychological Record, 50*, 305–

   337.

Horne, P. J., Hughes, J. C., & Lowe, C. F. (2006). Naming and categorization in young

    Children: IV. Listener behavior training and transfer of function. *Journal of the*

    Experimental Analysis of Behavior, 85, 247–273.

Horne, P. J., & Lowe, C. F. (1996). On the origins of naming and other symbolic behavior.

    Journal of the Experimental Analysis of Behavior, 65, 185–241.

Horne, P. J., Lowe, C. F., & Randle, V. R. L. (2004). Naming and categorization in young

    children: II. Listener behavior training. *Journal of the Experimental Analysis of*

    Behavior, 81, 267–288.

Hove, O. (2003). Differential probability of equivalence class formation following a one-to-

    many versus a many-to-one training structure. *The Psychological Record, 53*, 617–

    634.

Kastak, C. R., & Schusterman, R. J. (2002). Sea lions and equivalence: Expanding classes by

    exclusion. *Journal of the Experimental Analysis of Behavior, 78*, 449–465.

Kastak, C. R., Schusterman, R. J., & Kastak, D. (2001). Equivalence classification by

    California sea lions using class-specific reinforcers. *Journal of the Experimental*

    Analysis of Behavior, 76, 131–158.

Kuno, H., Kitadate, T., & Iwamoto, T. (1994). Formation of transitivity in conditional

    matching to sample by pigeons. *Journal of the Experimental Analysis of Behavior,*

    62, 399–408.

Lazar, R. M., Davis-Lang, D., & Sanchez, L. (1984). The formation if visual stimulus

    equivalences in children. *Journal of the Experimental Analysis of Behavior, 41*, 251–

    266.

Leaf, R. & McEachin, J. (1999). *A work in progress. Behavior management strategies and a*

cirriculum for intensive behavioral treatment of autism. New York: DRL

Leslie, J. C., Tierney, K. J., Robinson, C. P., Keenan, M., & Watt, A. (1993). Differences between clinically anxious and non-anxious subjects in a stimulus equivalence training task involving threat works. *The Psychological Record, 43*, 153–161.

Lionello-DeNolf, K. M. (2009). The search of symmetry: 25 years in review. *Learning & Behavior, 37*, 188–203.

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

Lowenkron, B. (1998). Some logical functions of joint control. *Journal of the Experimental Analysis of Behavior, 69*, 327–354.

Lowenkron, B. (2004). Meaning: A verbal behavior account. *The Analysis of Verbal Behavior, 20*, 77–97.

Lowe, C. F., Horne, P. J., Harris, F. D. A., & Randle, V. R. L. (2002). Naming and categorization in young children: Vocal tact training. *Journal of the Experimental Analysis of Behavior, 78*, 527–249.

Lowe, C. F., Horne, P. J., & Hughes, J. C. (2005). Naming and categorization in young children: III. Vocal tact training and transfer of function. *Journal of the Experimental Analysis of Behavior, 83*, 47–65.

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

Ninness, C., Dixon, M., Barnes-Holmes, D., Rehfeldt, R.A., Rumph, R., McCuller, G.,

Holland, J., Smith, R., Ninness, S. K., & McGinty, J. (2009). Constructing and

deriving reciprocal trigonometric relations: A functional analytic approach. *Journal of*

Applied Behavior Analysis, *42*, 191–208.

O'Donnell, J., & Saunders, K. J. (2003). Equivalence relations in individuals with language

limitations and mental retardation, *Journal of the Experimental Analysis of Behavior,*

80, 131–157.

Oden, D. L., Thompson, R. K., & Premack, D. (1988). Spontaneous transfer of matching by

infant chimpanzees (*Pan troglodytes*). *Journal of Experimental Psychology: Animal*

Behavior Processes, 14, 140–145.

Pena, T., Pitts, R. C., & Galizio, O. (2006). Identity matching-to-sample with olfactory

stimuli in rats. *Journal of the Experimental Analysis of Behavior, 85*, 203–221.

Pérez-González, L. A., & Williams, G. (2002). Multicomponent procedure to teach

conditional discriminations to children with autism. *American Journal on Mental*

Retardation, 107, 293–301.

Petursdottir, A. I. & Carr, J. E. (2011). A review of recommendations for sequencing

receptive and expressive language instruction. *Journal of Applied Behavior Analysis,*

44, 859–876.

Pilgrim, C., Jackson, J., & Galizio, M. (2000). Acquisition of arbitrary conditional

discriminations by young normally developing children. *Journal of the Experimental*

Analysis of Behavior, 73, 177–193.

Premack, D. (1976). *Intelligence in ape and man.* Hillsdale, NJ: Erlbaum.

Randell, T., & Remington, B. (1999). Equivalence relations between visual stimuli: The

functional role of naming. *Journal of the Experimental Analysis of Behavior, 71*, 395–415.

Randell, T., & Remington, B. (2006). Equivalence relations, contextual control, and naming. *Journal of the Experimental Analysis of Behavior, 86*, 337–354.

Rehfeldt, R. A., & Barnes-Holmes, Y. (2009). *Derived relational responding: Applications for learners with autism and other developmental disabilities*. Oakland, CA: New Harbinger.

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

Saunders, R. R., & McEntee, J. E. (2004). Increasing the probability of stimulus equivalence with adults with mental retardation. *The Psychological Record, 54*, 423–435.

Saunders, R. R., & Green, G. (1996). Naming is not (necessary for) stimulus equivalence. *Journal of the Experimental Analysis of Behavior, 65*, 341–353.

Saunders, R. R., & Green, G. (1999). A discrimination analysis of training-structure effects on stimulus equivalence outcomes. *Journal of the Experimental Analysis of Behavior, 72*, 117–137.

Saunders, K. J., & Spradlin, J. E. (1989). Conditional discrimination in mentally retarded adults: The effect of training the component simple discriminations. *Journal of the Experimental Analysis of Behavior, 52*, 1–12.

Saunders, K. J., & Spradlin, J. E. (1990). Conditional discrimination in mentally retarded adults: The development of generalized skills. *Journal of the Experimental Analysis of Behavior, 54*, 239–250.

Saunders, K. J., & Spradlin, J. E. (1993). Conditional discrimination in mentally retarded

    subjects: Programming acquisition and learning set. *Journal of the Experimental*

    Analysis of Behavior, 60, 571–585.

Saunders, K. J., Saunders, R. R., Williams, D. C., & Spradlin, J. E. (1993). An interaction of

    instructions and training design on stimulus class formation: Extending the analysis

    of equivalence. *The Psychological Record, 43*, 725–744.

Saunders, R., Drake, K. M., & Spradlin, J. E. (1999). Equivalence class establishment,

    expansion, and modification in preschool children. *Journal of the Experimental*

    Analysis of Behavior, 71, 195–214.

Schusterman, R. J., & Kastak, D. (1993). A California sea lion (*Zalo-phus californianus*) is

    capable of forming equivalence relations. *Psychological Record, 43*, 823–839.

Sidman, M. (2000). Equivalence relations and the reinforcement contingency. *Journal of*

    Experimental Analysis of Behavior, 74, 127–146.

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

    Cooperative.

Sidman, M., Cresson, O., & Willson-Morris, M. (1974). Acquisition of matching to sample

    via mediated transfer. *Journal of the Experimental Analysis of Behavior*, *22,* 261–

    273.

Sidman, M., Rauzin, R., Lazar, R., Cunningham, S., Tailby, W., & Carrigan, P. (1982). A

    search for symmetry in the conditional discriminations of rhesus monkeys, baboons,

    and children. *Journal of the Experimental Analysis of Behavior, 37*, 23–44.

Sidman, M. & Tailby, W. (1982). Conditional discrimination vs. matching to sample: An

expansion of the testing paradigm. *Journal of Experimental Analysis of Behavior, 37*, 5–22.

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

Sidman, M., Willson-Morris, M., & Kirk, B. (1986). Matching-to-sample procedures and the development of equivalence relations: The role of naming. *Analysis and Intervention in Developmental Disabilities, 6*, 1–19.

Smeets, P. M., & Barnes-Holmes, D. (2005). Auditory-visual and visual-visual equivalence relations in children. *The Psychological Record, 55*, 483–503.

Smeets, P. M., & Barnes-Holmes, D. (2003). Children's emergent preferences for soft drinks: Stimulus equivalence and transfer. *Journal of Economic Psychology, 24*, 603–618.

Skinner, B. F. (1957). *Verbal behavior*. Action, MA: Copley.

Spencer, T. J., & Chase, P. N. (1996). Speed analysis of stimulus equivalence. *Journal of the Experimental Analysis of Behavior, 65*, 643–659.

Spradlin, J. E., Cotter, V. W., & Baxley, N. (1973). Establishing a conditional discrimination without direct training: A study of transfer with retarded adolescents. *American Journal of Mental Deficiency, 77*, 556–566.

Stromer, R., MacKay, H. A., & Remington, B. (1996). Naming, the formation of stimulus classes, and applied behavior analysis. *Journal of Applied Behavior Analysis, 29*, 409–431.

Tomonaga, M., Matsuzawa, T., Fujita, K., & Yamamoto, J. (1991). Emergence of symmetry

in a visual conditional discrimination by chimpanzees (*Pan troglodytes*).

Psychological Reports, 68, 51–60.

Walker, B. D., Rehfeldt, R. A., & Ninness, C. (2010). Using the stimulus equivalence

paradigm to teach course material in an undergraduate rehabilitation course. *Journal

of Applied Behavior Analysis, 43*, 615–633.

Wulfert, E., Dougher, M. J., & Greenway, D. E. (1991). Protocol analysis of the

correspondence of verbal behavior and equivalence class formation, *Journal of the

Experimental Analysis of Behavior, 56*, 489–504.

VITA

CHARLOTTE LYNN CARP, Katy, Texas

Education

Ph. D in Experimental Psychology, May 2012
Texas Christian University, Fort Worth, Texas
Advisor: Anna Ingeborg Petursdottir, Ph. D

M.S. in Experimental Psychology, May 2010
Texas Christian University, Fort Worth, Texas

B.S. in Psychology, December 2004
Minor: Mathematics
Texas Christian University, Fort Worth, Texas

Publications

Carp, C. L., Peterson, S. P., Arkel, A., Petursdottir, A. I., & Ingvarsson, E. T.
(accepted pending revisions). Effects of picture prompts on the acquisition of
auditory-visual conditional discriminations and tact emergence. *Journal of
Applied Behavior Analysis*.

Carp, C. L., & Petursdottir, A. I. (in press). Effects of two training conditions on the
emergence of novel intraverbals: An extension of Pérez-González et al. *The
Psychological Record*.

Petursdottir, A. I., **Carp, C. L.**, Matthies, D. W., & Esch, B. E. (2011). Analyzing
stimulus-stimulus pairing effects on preferences for speech sounds. *The
Analysis of Verbal Behavior, 27*, 45–60.

Hubbard, T. L., Kumar, A. M., & **Carp, C. L.** (2009). Effects of spatial cue timing and
relevance on representational momentum. *Journal of Experimental
Psychology: Learning, Memory, and Cognition, 35*(3), 666–677.

Editorial Experience

Editorial Assistant
The Analysis of Verbal Behavior, Present

Ad-Hoc Reviewer
Journal of Applied Behavior Analysis
The Analysis of Verbal Behavior

ABSTRACT

MEDIATION BY INTRAVERBAL NAMING IN

CHILDREN'S EQUIVALENCE TEST PERFORMANCE

by Charlotte Lynn Carp, Ph.D., 2012
Department of Psychology
Texas Christian University

Dissertation Advisor: Dr. Anna Ingeborg Petursdottir, Assistant Professor of Psychology

The mechanisms underlying stimulus equivalence are of considerable debate in the literature, especially regarding verbal behavior. The purpose of the present study was to evaluate Horne and Lowe's (1996) intraverbal naming hypothesis by investigating the effects of equivalence class formation on the emergence of novel intraverbals in six typically developing kindergarteners ages 4–6. Participants were first taught vocal tacts for pictures of states, birds, and flowers. Following tact training, they were exposed to MTS training in which they were first taught an A–B (i.e., state–bird) relation and an A–C (i.e., state–flower) relation and then exposed to an equivalence test probing 12 B–A, C–A, B–C, and C–B emergent intraverbal relations. Immediately following equivalence testing, an intraverbal test was conducted that probed intraverbal relations between the pictures (e.g., "Florida goes with which bird?"). Horne and Lowe predict that under those circumstances, participants who pass the equivalence test should also show a tendency to emit the relevant intraverbals. All three participants who passed the equivalence test passed the intraverbal test; however, two of those participants required additional rounds of both tests. Other participants passed a

symmetry only test after requiring additional B–C training; only one of those participants passed the intraverbal test. Results suggest that verbal behavior may be necessary or have a substantial facilitative effect for performance on transitive relations, but unnecessary for performance on symmetrical relations. Implications for Horne and Lowe's (1996) analysis of intraverbal naming are discussed.