EFFECT OF TRAINING SEQUENCE ON EMERGENT STIMULUS RELATIONS

by

REAGAN ELAINE COX

Bachelor of Science, 2016
Angelo State University
San Angelo, 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

Master of Science

December 2018

EFFECT OF TRAINING SEQUENCE ON EMERGENT STIMULUS RELATIONS

by

Reagan E. Cox

Thesis approved:

Major Professor

For The College of Science and Engineering

ACKNOWLEDGEMENTS

I would first and foremost like to thank my advisor Anna I. Petursdottir, Ph.D. for dedicating so many hours to editing my drafts, shaping up my writing, and pouring her knowledge into me. I have learned so much about behavior analysis and research in the short time I have worked with you and am so excited to continue learning from such an amazing scientist. I would also like to thank my thesis committee members, Kenneth Leising, Ph.D. and Mauricio Papini, Ph.D. for helping shape this project from the beginning and asking the tough questions that allowed me to grow in my knowledge of this field.

The research assistants in our lab, Remington Swensson, Camille Roberts, Alexandra Wilkins, and Angela Gallo played essential roles in data collection and analysis. There is no way I could have finished my thesis without them. Thank you so much for spending countless hours listening to me talk through every single obstacle we came across and entering millions of numbers into excel files without ever complaining.

Lastly, I want to thank my incredibly supportive husband, Wesley Cox, for encouraging me to press on every time I complained and for listening to me talk about this project hundreds of times and still being (or at least pretending to be) interested.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Introduction

Quickly after birth, humans begin to form concepts based on their learning history. Keller and Schoenfeld (1950) specified that “when a group of objects gets the same response, when they form a class the members of which are reacted to similarly, we speak of a concept” (p.154). For example, the sights and sounds of many different types of cats may evoke the response of saying “cat” in a human, exemplifying the concept of a cat. Further, when first exposed to a bobcat, cheetah, or an ocelot, a young child may say “cat” as an instance of stimulus generalization. As humans age, these concepts change based on our learning history.

Humans and nonhumans can form concepts based on perceptual features, relations among features, or acquired associations (Zentall, Galizio, & Critchfield, 2002). A perceptual concept exists when stimuli in a category are related based upon common physical features. For example, every day humans categorize people they see into gender based upon common physical features. As another example, Herrnstein and Loveland (1964) trained pigeons to categorize photographs based on the presence or absence of humans in the photos. In relational concept learning, which is more abstract than perceptual concept learning, stimuli are categorized based on similar relations among features in the group (Zentall et al., 2002). For example, when shown sample XX and given a choice between YZ and BB, the reinforced response would be BB because it is relationally similar to XX; therefore, it is not the relationship between XX and BB that is learned, but rather the similarity between the XX relationship and the BB relationship; this is sometimes referred to as same versus different concept learning. Cook, Cavoto, Katz, and Cavoto (1997) showed relational concept learning in pigeons by teaching them to peck keys in a target area that differed in color or shape from its surroundings. In this study, the researchers found that pigeons were able to discriminate between same and different colors and textures.

Finally, in associative concept learning, stimuli are grouped together based on learned associations. For example, when thinking about a specific number, the number three, one can think about the numeral 3, the quantity three (e.g., three balls), or the spoken word "three". All belong to the concept of three even though there is no perceptual similarity or abstract relational similarity between these stimuli (Zentall et al., 2002). Instead, the concept reflects prior learning about the relations between its members, such as learning that both the numeral 3 and the spoken word "three" represent the quantity three.

The formation of an associative concept does not necessarily require directly experiencing all possible associations between all members. Instead, only a subset of associations may be taught directly while the rest are emergent. For example, after being taught that A is related to B, B is related to C, and C is related to D, humans will perform consistent with A being related to C, B to D, and so forth. Research on stimulus equivalence (Sidman & Tailby, 1982) has focused on how physically dissimilar stimuli may come to be grouped together based on common associations with other stimuli. The present study similarly focuses on how dissimilar visual stimuli may be grouped together based on acquired associations between their names.

Stimulus Equivalence Tasks

Stimulus equivalence is often demonstrated in matching-to-sample (MTS) tasks. In an MTS task, participants are shown one stimulus, called the sample stimulus, followed by multiple stimuli at once, called the comparison stimuli. Participants are then taught through multiple trials which comparison stimulus goes with which sample stimulus. Sidman & Tailby (1982) for example, used an MTS format to teach baseline relations. They used three sets of stimuli, A, B, and C, each with three stimuli, A1, A2, A3, B1, B2, B3, C1, C2, and C3. After being taught via an MTS task to select B and C stimuli conditionally when presented with A stimuli, children could readily match B and C stimuli regardless of not being directly taught these associations (Sidman & Tailby, 1982).

Stimulus equivalence has also been demonstrated in go/no-go tasks, during which participants are shown sets of compound stimuli (two stimuli presented together side by side) that either "go together" or not (e.g., Silva & Debert, 2017). Participants are to respond to some stimulus combinations (e.g., A1B1, A2B2, A1C1, and A2C2) and not respond to other variations (e.g., A1B2, A2B1, A1C2, A2C1). Participants are then tested by presenting novel stimulus combinations (e.g., B1C1) in the same manner. In addition, stimulus equivalence has been demonstrated in MTS tests following Pavlovian type baseline training that simply involves contiguous presentation of stimulus pairs (e.g., Leader, Barnes, & Smeets, 1996).

Humans may also demonstrate novel grouping of visual stimuli based on verbal learning. First, visual stimuli may be grouped together after learning that they share common verbal labels (e.g., Lowe, Horne, Harris, & Randle, 2002; Horne, Lowe, & Randle, 2004). Second, stimuli may be grouped together based on acquired associations between verbal labels already associated with the visual stimuli. In the intraverbal naming task, which will be used in the present study, this has been demonstrated by teaching subjects to verbally label two sets of stimuli; this phase is referred to as tact (Skinner, 1957) training, see panel A of Figure 1. Once mastery criterion has been reached, subjects are taught arbitrary relations between pairs of verbal labels; this phase is referred to as intraverbal training, see panel B of Figure 1. Subjects are then given an MTS test in which one visual stimulus is presented as a sample and the subject is asked to select a comparison, none of which are visually similar to the sample, but one of which is related to the sample via the taught verbal associations, see panel C of Figure 1. Children and adults typically pass these MTS tests if the baseline verbal relations are maintained throughout testing (Jennings & Miguel, 2017; Ma, Miguel, & Jennings, 2016; Petursdottir, Carp, Peterson, & Lepper, 2015; Petursdottir, Cox, Mellor, & McKeon, in preparation; Santos, Ma, & Miguel, 2015). In a more complex task, the first phase involves teaching labels for three sets of stimuli and the second phase involves teaching intraverbal relations between two pairs of labels that link three visual stimuli together. Participants generally perform well on MTS tests of visual relations in these paradigms as well (Jennings & Miguel, 2017; Ma, Miguel, & Jennings, 2016).

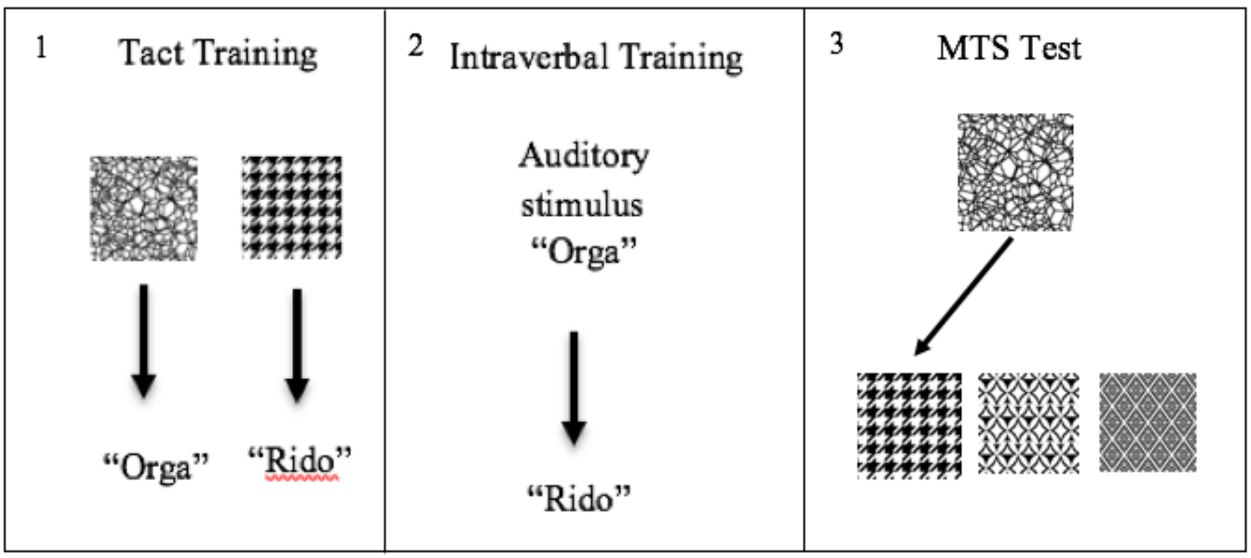


Figure 1. Depiction of tact and intraverbal training and MTS testing in the intraverbal naming task. In tact training, the labels of visual stimuli are taught (A-A'). In intraverbal training, the relations between verbal labels are taught (A'-B'). In the MTS test, subjects are tested for the emergence of A-B relations.

Verbal Mediation Hypothesis

It has been debated whether behavioral mediating events are needed to account for the emergence of novel stimulus relations seen in the previously described tasks. Sidman (2000) argues that equivalence relations result from the reinforcement of overlapping conditional discriminations without any behavioral mediating events. Similarly, relational frame theory (RFT) does not consider mediating behavioral events to be necessary to account for stimulus equivalence or other types of derived relations (Hayes, Barnes-Holmes, & Roche, 2001). RFT suggests that, after an appropriate learning history, higher-order relational operants may be formed, brought under contextual control, and applied to arbitrary stimulus relations (Barnes-Holmes, Barnes-Holmes, Smeets, Cullinan, & Leader, 2004). From this perspective, behavioral mediating events may well be involved in the derivation of novel stimulus relations, but such events are redundant to theory because the end results of training overlapping relations can be predicted without taking mediating events into account. By contrast, mediational theories suggest that performance on a test of derived stimulus relation is a function of whichever sources of stimulus control are present at the time of test, and that in the case of humans, their prior problem-solving histories may engender behaviors that both help them respond correctly during training and allow them to self-construct supplemental sources of stimulus control during testing (Palmer, 2012; Moustakis & Mellon 2018). An example of such a theory is the verbal mediation hypothesis (e.g. Horne & Lowe, 1996) that suggests that after learning a subset of relations between stimuli, humans will utilize verbal strategies to infer the remaining relations. Put differently, a subject will be able to select the correct comparison in a matching-to-sample task even when they have never previously encountered the sample and the comparison together by way of talking themselves through previously learned relations (e.g., “This is sample A, which goes with B, and B goes with C, so C must be correct”). This hypothesis is unlikely to account for all instances of emergent associations between stimuli, as such associations have been demonstrated in nonhuman animals in various types of tasks (e.g., Swisher & Urcuioli, 2013;

Urcuioli & Swisher, 2015; Zentall, Clement, & Weaver, 2003). Nevertheless, the relative ease with which certain types of inferred associations are demonstrated in humans (Lionello-DeNolf, 2009) may suggest that verbal or other behavioral mediating events derived from previous learning histories play a role in their performance. In possible support of this notion, humans perform better in stimulus equivalence tasks with nameable stimuli (e.g., Mandell & Sheen, 1994) and when the names of stimulus class members rhyme (e.g., Randell & Remington, 2006), and when children pass equivalence tests with stimuli they can name, they are also found to be able to state the relations between the stimuli verbally in their absence (Carp & Petursdottir, 2015).

Horne and Lowe (1996) proposed intraverbal naming as one mechanism by which verbal behavior may mediate human performance in stimulus equivalence tasks. They hypothesized that subjects tend to assign names to the stimuli presented in the tasks, followed by the acquisition of intraverbal relations between those names during relational training. These intraverbal relations, in turn, mediate correct test performance, see panel A of Figure 2 for a visual depiction of this process. The previously described intraverbal naming task was explicitly designed to model this history (Petursdottir et al., 2015). That is, instead of teaching relations between pairs of visual stimuli that a human might give names to and go on to relate verbally, the participant is taught simply to name the stimuli and relate them verbally in the absence of other relational training. The emergence of correct matching in this task is sometimes interpreted in support of Horne and Lowe's theory, as talk-aloud procedures and verbal reports of adult participants are often consistent with the notion that they talk themselves through the MTS test (e.g., Ma et al., 2016). But although test performance is clearly a result of learning verbal relations, research with children suggests that they actually do not talk themselves through the test (Petursdottir et al., 2015). The intraverbal naming hypothesis requires that the intraverbal associations learned must be bidirectional (Horne & Lowe, 1996). For example, if a subject is taught that “A-name goes with B-name”, then the subject must also infer that the B-name goes with the A-name in order to perform correctly in trials in which a B stimulus is the sample and the comparisons are A stimuli. However, it has been shown that children can still perform well on the MTS tests even when these intraverbal associations are unidirectional. Specifically, Petursdottir et al. (2015) found that, after assessing the emergence of visual-visual discriminations following training of vocal tact and intraverbal relations, of the five children who passed the MTS test, four did not acquire bidirectional associations. Additionally, reaction time analyses in this study did not suggest that the participants were using a more time-consuming verbal strategy that would require only unidirectional intraverbals.

Although adults may tend to solve the MTS test verbally, these data from children suggest it may be possible to do so without verbal mediation. From the perspective of a mediational account, this leaves the question of whether other behavioral events might mediate performance in those cases such that relevant sources of stimulus control are present at test.

Visual Mediation

Even when verbal behavior does not seem to mediate performance, it is possible that other behavioral events are involved in facilitating it. In the intraverbal naming task, some evidence exists that participants may visualize the visual stimuli together when being taught the verbal associations between their names. In research conducted by Jennings and Miguel (2017), several adult participants (four out of seventeen) reported visualizing the stimuli together when being taught the verbal associations and attributed their MTS performances to having done so. In addition, participants in Petursdottir et al. (2015) were observed to often close their eyes during the instruction of verbal relations, which could suggest that they were attempting to visualize the absent visual stimuli. If participants tend to visualize the stimuli during intraverbal training (i.e., visual representations are activated), this might facilitate their subsequent MTS performance. For example, if a subject is first taught that stimulus A is called “pogo” and stimulus B is called “regi,” then when the subject is taught that “pogo” goes with “regi” he may visualize stimuli A and B and acquire associations between the visual stimuli along with the verbal associations, see panel B of Figure 2 for a visual depiction of this process. Therefore, during the MTS test, the sample stimulus may evoke an image of the correct comparison that controls selection in the absence of verbal mediation.

Visual imagining has previously been shown to enhance various forms of verbal learning; for example, Kisamore, Carr, & LeBlanc (2011) found that young children were unable to emit many responses to questions about category membership (e.g., what are some animals?) after being taught to tact stimuli within four different categories, but when prompted to use visual imagining, the children began emitting a large number of categorization responses. However, visual imagining has not been previously shown to affect the grouping of visual stimuli based on verbal learning.

If visual imagining does play a role in the intraverbal naming task, then it must be crucial to first learn the name corresponding to each stimulus before learning the verbal associations between the names of the stimuli. If this is not the case, then it would be impossible to use visualization because the visual stimuli would have not yet been presented when learning the verbal associations. If participants are using visualization instead of verbal mediation, reaction times on the MTS test might be expected to be faster, as visual mediation provides a short cut alternative to verbal mediation (see Figure 1). Verbal mediation involves going through three steps (naming sample, recalling the associated verbal label, and recalling the image that goes with that label) in order to generate the correct response during the MTS test while individuals who have acquired associations between the visual stimuli via imaginal exposure during intraverbal training may have an immediate source of stimulus control available if the sample stimulus evokes an image of the correct comparison. A previous study (Petursdottir et al., in preparation) was conducted in our lab to test this sequence effect prediction. We found that participants who received tact training first had significantly shorter reaction times on the MTS test when controlling for baseline retention. Although there was no significant difference between the mean MTS accuracy scores in each group when controlling for baseline retention, there was a non-significant tendency for participants who learned to name the stimuli in tact training prior to intraverbal training (standard sequence) to perform more accurately at the beginning of the MTS test compared to when intraverbal training was conducted before tact training. A limitation of this study was that when retention of baseline relations (i.e., the taught tacts and intraverbals) was tested following the MTS test, the majority of participants scored below 88% correct. Because MTS accuracy is likely to be influenced by baseline retention, this lack of retention may have introduced excessive variability in MTS accuracy, leading to an inability to detect an effect of condition on MTS accuracy. An approximately equal number of participants in both groups did retain baseline relations, and of those participants, the ones in the standard group tended to perform better on the MTS test, but the difference between these subgroups was not significant, perhaps because of the small number of participants in them (six in the standard and five in the reverse group).

Present Experiments

The present study consisted of two experiments. The purpose of the first experiment was to replicate the Petursdottir et al. (in preparation) study while attempting to increase baseline retention. It was predicted that with increased baseline retention, less variability in MTS accuracy would be attributable to this variable, making it possible to detect an effect of training sequence on accuracy in addition to reaction times. The following changes were made in an attempt to increase maintenance scores:

1. The mastery criteria were increased to 30 correct trials in a row for the tact instruction from the original 24 trial criterion and increased to 15 correct trials in a row from the original 12 trial criterion for the intraverbal instruction.
2. Participants were only given 3-s to respond after the onset of the visual stimulus (tact instruction) or verbal stimulus (intraverbal instruction) rather than the original 5-s. This change was made due to the observation made in the original study that even though participants may not be confident in their answers and take a few seconds to think about the answer before providing it, if the answer was correct, the trial was still marked as correct. Decreasing the amount of time allotted to provide an answer was intended to increase fluency and, subsequently, require the participant to know the stimuli better.
3. In an attempt to make the tasks easier for participant, we used only 6 visual and verbal stimuli rather than the original 8.
4. The visual stimuli used were different from the original study, due to the observation that the majority of participants had trouble discriminating between two specific stimuli due to their similarity.

The purpose of the second experiment was to assess the generality of the previous finding by examining if the previously observed effect on reaction times (and possibly MTS accuracy) would be obtained in a version of the intraverbal naming task in which training is conducted in the same MTS format as testing. In research on stimulus equivalence, MTS tasks represent the most prevalent methodology. Because of this, it would be of interest to demonstrate an effect of training sequence in an all-MTS task. Additionally, applications in the area of equivalence-based instruction, such as learning new material in a classroom setting, typically use computerized MTS instruction to teach baseline relations (e.g. Fienup & Critchfield, 2011); therefore, it would be of practical interest to know if training sequence matters. The second experiment, thus, changed the training format of experiment 1 to an MTS format in which auditory stimuli and participant vocal responses were replaced with textual stimuli.

Therefore, the present experiments compared the effects of two vocal training sequences in experiment 1 and two MTS training sequences in experiment 2 on matching-to-sample (MTS) performance: The standard groups learned associations between the verbal/text stimuli A1'-B3' and visual stimuli A1-B3 before learning verbal/text associations between the stimuli A1'-B3', whereas the reverse groups first learned the verbal/text associations, and only after that, learned to name the visual stimuli. The standard groups were predicted to have faster reaction times on the MTS test due to being able to use the visual mediation "short cut". Additionally, the standard groups were predicted to obtain superior MTS test scores compared to the reverse groups.

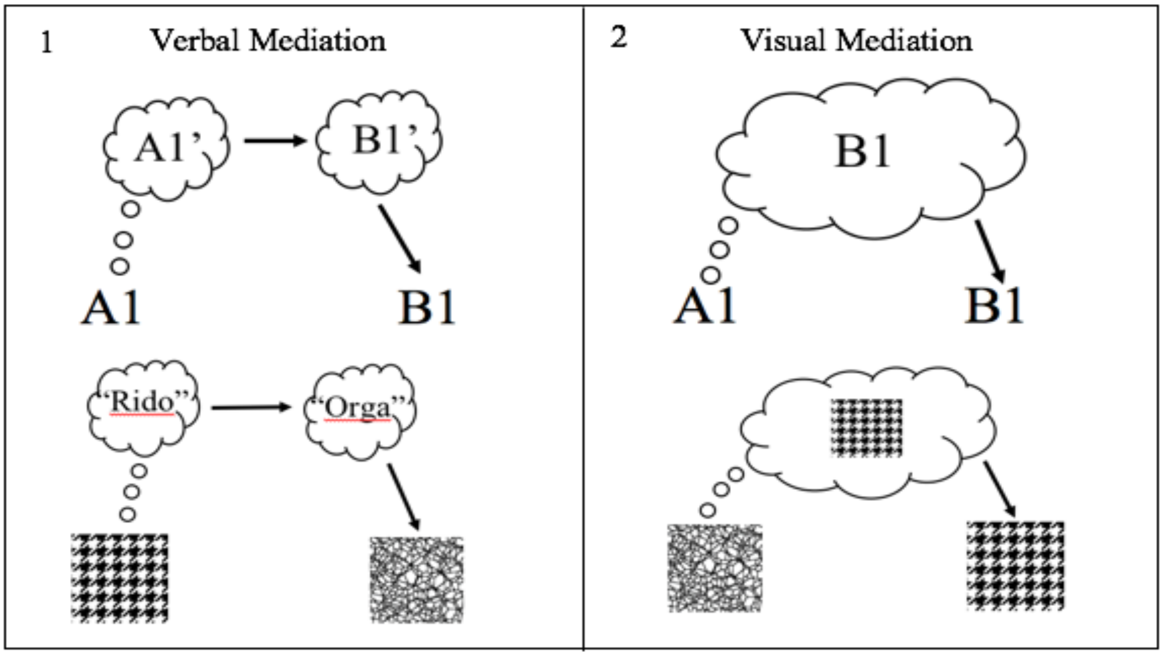


Figure 2. Depiction of individual steps for the verbal mediation and visual mediation processes. As seen here, verbal mediation involves more steps (A1-A1'-B1'-B1) than visual mediation (A1-B1).

Experiment 1

Method

Participants and setting. Following approval by the Texas Christian University Institutional Review Board, 32 adults (27 females and 5 males, age range: 18 – 43 years, M = 20.16 years) were recruited from the Psychology department's human subjects pool via the research participation software *Sona-Systems*. Experimental sessions were conducted one participant at a time in a small, quiet room with only the investigator(s) present. During the session, the participant was seated at a desk with a laptop and a computer mouse. The experimenter sat to the side of the participant and a second observer, when present, sat behind the participant. A screen recording software, *FlashBack Express,* was running in the background of the computer to record the computer screen and the participant's vocal responses. Each participant completed one session that was expected to last about 45 minutes on the average, and no more than 60 minutes. The *SuperLab 5.0* experiment software was used to program and run the experiment.

Stimuli. Visual stimuli consisted of 6 black and white vector patterns from *wallpaper-mobile* (http://wallpaper-mobile.mdvwi.cricket/files/images/black-and-white-patterns-vector.html). Three stimuli were designated as A stimuli and labeled A1, A2, and A3 and the other three were designated as B stimuli and labeled B1, B2, and B3. The participants were taught the names “orga,” “huzo,” and “luti,” for the A stimuli and the names “rido,” “fodi,” and “poga,” for the B stimuli (see Figure 3). These names will be referred to as A1’ through B3’ when presented as auditory stimuli, vocalized by participants, or presented as text stimuli.

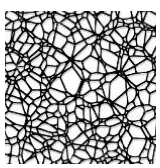
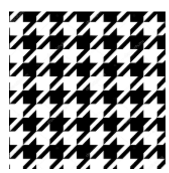
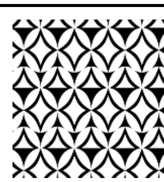
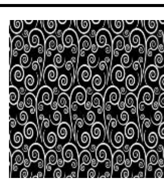
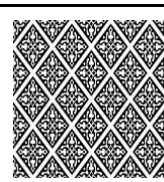
Experimental Stimuli				
	A	A’	B	B’
1		Orga		Rido
2		Huzo		Fodi
3		Luti		Poga

Figure 3. Visual stimulus classes and their designated names.

Verbal visual stimuli will consist of the 6 typed words Rido, Orga, Huzo, Fodi, Luti, and Poga.

Stimulus similarity survey. 10 participants recruited from personal contacts were given surveys that consisted of 24 MTS trials where each visual stimulus served as a sample 4 times while the other 5 stimuli served as comparisons. Participants were instructed to circle the picture they felt “went with” the sample. Any stimuli that were consistently paired together were assigned to be either both A or both B stimuli to control for any predisposition to pair stimuli together.

Procedure. Participants were randomly assigned to one of 2 groups: standard and reverse. There were 12 females and 4 males in the standard group and 15 females and 1 male in the reverse group; the mean age was 20.75 in the standard group and 19.56 in the reverse group. The standard group received tact instruction and then intraverbal instruction while the reverse group received the opposite. All procedures following training were identical across groups.

Overview. After completing the necessary consent and demographic forms, participants were prompted to read the instructions on the screen that read, “When you are ready to begin, press the SPACE bar.” Participants were then taught to either vocally tact visual stimuli A1 through B3 or match text stimuli A1’ through B3’ to the corresponding visual stimuli A1 through B3. Next, the participants were exposed to A’B’ intraverbal training, during which the participant either learned to respond to the verbal stimuli A1’, A2’, and A3’ by vocalizing B1’, B2’, and B3’ respectively, or learned to match A’ text stimuli to B’ text stimuli. Once the A’B’ relations were learned, the participants were given a visual-visual MTS test for AB and BA conditional discriminations. After 30 MTS trials, participants were given a reverse intraverbal test in which the participants heard the B' stimuli and should have vocalized the corresponding A' stimuli. Participants were then given six intraverbal maintenance test trials and six tact maintenance test trials followed by a listener test during which participants heard stimuli A1' through B3' and responded by selecting a visual stimulus. See Table 1 for a table depicting the methodological overview.

Table 1

Design Overview

Study	Group	Baseline Training I	Baseline Training II	Post- test 1	Post-test 2	Post-test 3	Post-test 4
	Standard (n = 16)	Vocal tact training (A-A')	Vocal intraverbal training (A'-B')	MTS test	Reverse Intraverbal (B'-A')	Baseline Retention (A-A', A'-B')	Listener test (A'-A, B'-B)
Study 1	Reverse (n = 16)	Vocal intraverbal training (A'-B')	Vocal tact training (A-A')	MTS test	Reverse Intraverbal (B'-A')	Baseline Retention (A-A', A'-B')	Listener test (A'-A, B'-B)

Vocal tact training (AA'/BB'). During tact training, participants were taught to vocalize the verbal stimuli A1' through B3' when presented with visual stimuli A1 through B3, respectively. Trials were presented in blocks of six, with one presentation of each stimulus per trial block. The order of stimulus presentation was varied by the experiment software within each trial block. Tact training began with an instruction on the computer screen that will read, "You will see six images paired with words. Please attend carefully and try to remember what you see and hear." To begin, participants were exposed to each visual stimulus in the middle of the screen for 2-s while the corresponding auditory stimulus was played. After being exposed to all stimuli, a second set of instructions appeared on the screen that read, "You will now see the images again, one at a time. Please say the word that goes with each image. The experimenter will let you know if you are right or wrong and help you if you don't remember." A trial began with a 2-s presentation of one of the visual stimuli; participants then had 3-s to respond with the correct verbal stimulus. If the participant did not vocalize the correct response within the 3-s interval, the experimenter vocalized the correct response. If the participant vocalized the correct response within the allotted time, the experimenter provided feedback by saying "correct." Training continued in this manner until participants responded with 100% accuracy across 30 trials.

Vocal intraverbal training (A'B'). During intraverbal training, participants were taught to vocalize the verbal responses B1' through B3' when presented with auditory stimuli A1' through A3', respectively. Trials were presented in blocks of three, with one presentation of each stimulus per trial block. The order of stimulus presentation varied across trial blocks. Intraverbal training began with an instruction on the computer screen that read, "You will hear three pairs of words. Please attend carefully and try to remember what you hear." To begin, participants were exposed to each auditory stimulus A1 through A3 immediately followed by the auditory stimulus B1 through B3, respectively. Each pair of auditory stimuli were followed by the instructions "press the SPACE bar to continue". After being exposed to all stimuli, a second set of instructions appeared on the screen that reads, "You will now hear one of the words in each pair. Please say the other word in the pair. The experimenter will let you know if you are right or wrong and help you if you don't remember." A trial began with the presentation of one of the A' auditory stimuli; participants had 3-s to respond with the correct B' verbal stimulus. If the participant did not vocalize the correct response within the 3 second interval, the experimenter vocalized the correct response. If the participant vocalized the correct response within the allotted time, the experimenter reinforced the correct response by saying "correct." Training continued in this manner until participants responded with 100% accuracy across 15 trials.

MTS test (AB and BA). The MTS test consisted of 30 trials ordered randomly within blocks of 6. Each visual stimulus (A1 through B3) was presented as a sample five times. B stimuli served as comparisons for A samples while A stimuli served as comparisons for B samples. The participants first saw a set of instructions that read, "You will see an image in the center of the screen. When you click on the image, three images will appear across the bottom of the screen. Click on the image that goes with the image in the center. You will not be told if your responses are correct or incorrect, but please do your best to respond correctly." One of the six visual stimuli appeared in the center of the screen serving as the sample stimulus for that trial. The participants had as much time as they wanted to look at the sample and then click on it. When the sample was clicked on, it disappeared, and the three comparison stimuli were presented across the bottom of the screen. Participants had 5-s to choose the comparison stimulus that corresponded to the sample stimulus previously presented. After clicking on a comparison stimulus, participants were prompted to press the spacebar to begin the next trial. Comparisons were randomly assigned to locations on the screen in each trial.

Reverse vocal intraverbal test (B'A'). The reverse vocal intraverbal test occurred immediately following the MTS test for the standard and reverse groups. One block of 6 trials was presented in the same manner as intraverbal training with one exception: The B' stimuli were presented instead of the A' stimuli, and participants had to respond with the corresponding A' names.

Tact and intraverbal maintenance. Tact maintenance occurred immediately following the reverse intraverbal test. One block of six trials was presented in the same manner as tact training; however, no error correction was provided for incorrect responses, and no reinforcement was provided for correct responses. Intraverbal maintenance occurred immediately following tact maintenance. One block of six trials was presented in the same manner as intraverbal training; however, no error correction was provided for incorrect responses, and no reinforcement was provided for correct responses.

Listener test. The listener tests occurred immediately following the maintenance tests. The test consisted of one block of six trials. For the standard and reverse groups, first, auditory stimuli A' or B' were presented. Then, all six comparison visual stimuli were presented across the top and bottom of the screen. Participants were expected to choose stimulus A1 when auditory stimulus A1' was presented and so on.

Results and Discussion

Baseline training and retention. An independent samples t-test was conducted total trials to criterion, and retention test scores. There were no significant differences between groups on total trials to criterion or retention test scores ps $\geq .507$ (see Table 2 for means and standard deviations; see Figures 4 and 5 for retention test and trials to criterion data). Each group did, however, receive more training trials than the other group in whichever type of training they received first (e.g. the standard group received more overall tact training trials than the reverse group); ps $\leq .011$. One of the major goals of this experiment was to increase baseline retention in comparison to the previous study (Petursdottir et al., in preparation). The proportion of participants who met the passing criterion for baseline retention was increased from 11 out of 32 participants in the pilot to 19 out of 32 participants in the current study. More specifically, 9 participants in the standard group and 10 participants in the reverse group passed both retention tests (at least 5 out of 6 correct trials). Additionally, mean MTS scores increased (standard M =

76%, reverse M = 63%) from Petursdottir et al. (in preparation) (standard *M* = 59%, reverse M = 47%). Also interesting, is that intraverbal maintenance in the reverse group (which was relatively low in Petursdottir et al. (in preparation) with a mean of 60%) was increased to 80%.

Table 2

Means and Standard Deviations for reverse intraverbal, listener, and retention tests, and trials to criterion for Experiment 1

	Standard Group		Reverse Group	
	Mean	Standard Deviation	Mean	Standard Deviation
Tact Trials to Criterion	322.00	174.75	182.44	107.43
Intraverbal Trials to Criterion	73.63	49.92	165.75	93.89
Total Trials to Criterion	393.75	200.35	348.19	183.00
Reverse Intraverbal Test	63.54%	29.95	58.33%	33.33
Tact Maintenance Test	87.50%	17.74	89.58%	13.44
Intraverbal Maintenance Test	79.17%	26.87	80.21%	31.75
Listener Test	82.29%	33.59	93.75%	10.32

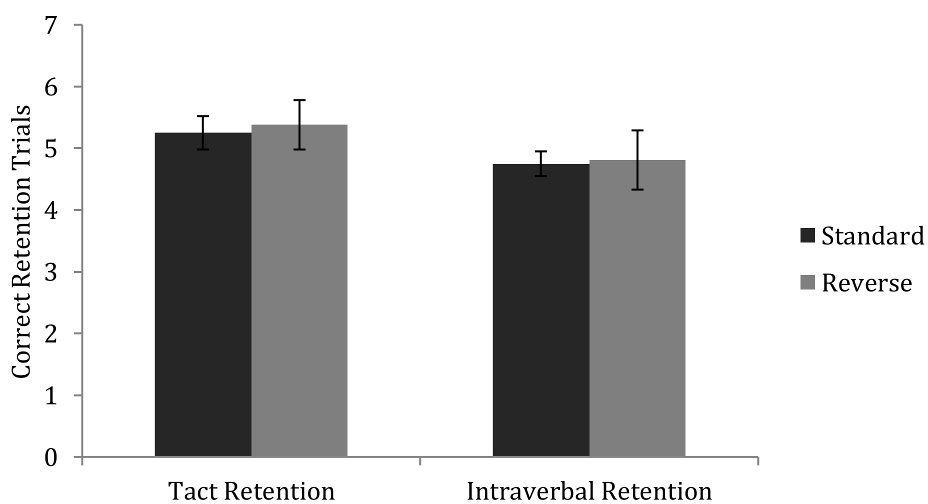


Figure 4. Mean number of correct trials on the retention tests. Error bars represent the standard error for the mean in each group.

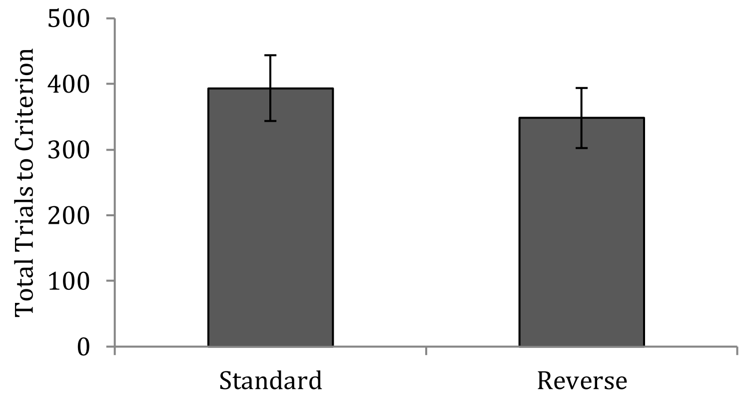


Figure 5. Mean total number of training trials to reach mastery criterion. Error bars represent the standard error for the mean in each group.

MTS and other emergent relations. MTS accuracy was analyzed by a 5 (trial block) X 2 (condition: standard, reverse) mixed-model analysis of variance, with repeated measures on the trial block factor, see Figure 6. Due to a violation of sphericity according to Mauchly's test of sphericity, the Greenhouse-Geisser correction was used. Although mean accuracy was higher in the standard (M = 76.44%, SD = 20.29) than in the reverse group (M = 62.56%, SD = 31.49), there was no significant main effect of condition, $F(1, 30) = 1.48$, $p = .233$. There was a significant main effect of trial block, $F(2.71, 120) = 3.91$, $p = .014$; as the blocks progressed, MTS accuracy increased. There was, however, no significant interaction between condition and trial block, $p = .252$.

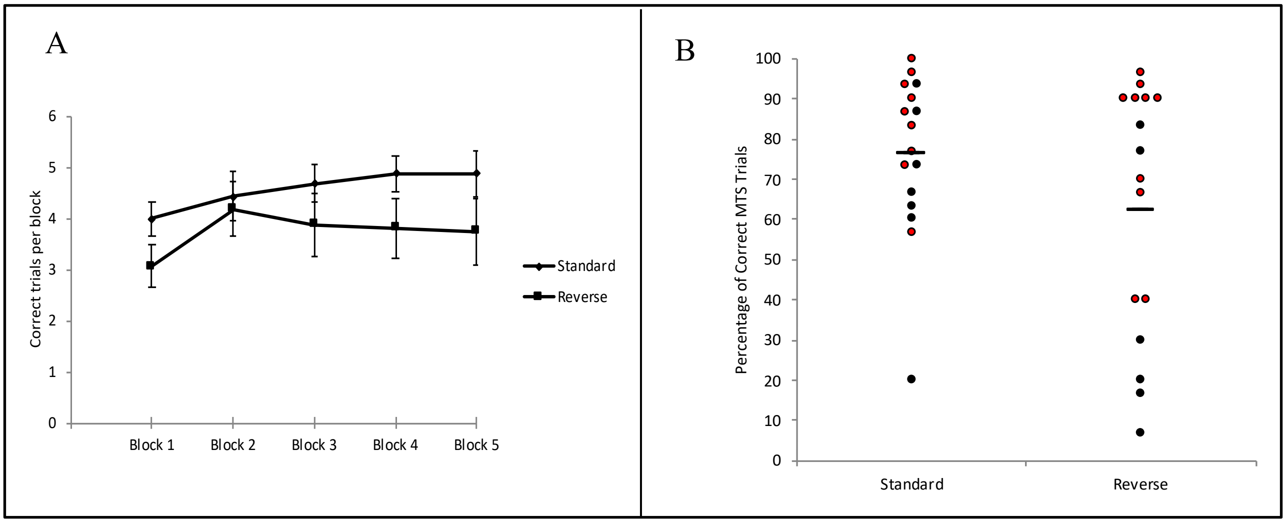


Figure 6. Panel A shows the number of correct trials per block during the MTS test. Error bars represent standard error. Panel B shows the individual MTS averages. Red circles indicate participants who passed both retention tests (at least 5/6 correct trials). The vertical lines represent the group means.

Individual distributions for the MTS reaction times (i.e. time between comparison onset and comparison selection) were positively skewed; because of this, median reaction times were used for analyses rather than the means (see Figure 7). Trials in which participants did not choose a comparison within the time allotted were scored as no response, and these trials were not included in the reaction time analyses. The reaction time results were also analyzed by a 5 (trial block) X 2 (condition: standard, reverse) mixed-model analysis of variance, with repeated measures on the trial block factor (see Figure 8). There was no main effect of condition, $F(1, 30) = 1.008$, $p = .323$; however, the differences between groups is in the expected direction of the standard group ($M = 1606.44$, $SD = 437.36$) reacting faster than the reverse group ($M = 1750.00$, $SD = 509.41$). There was a significant main effect of trial block, $F(4, 120) = 8.082$, $p < .001$; as the blocks progressed, reaction times overall decreased. There was, however, no significant interaction between condition and trial block, $p = .560$.

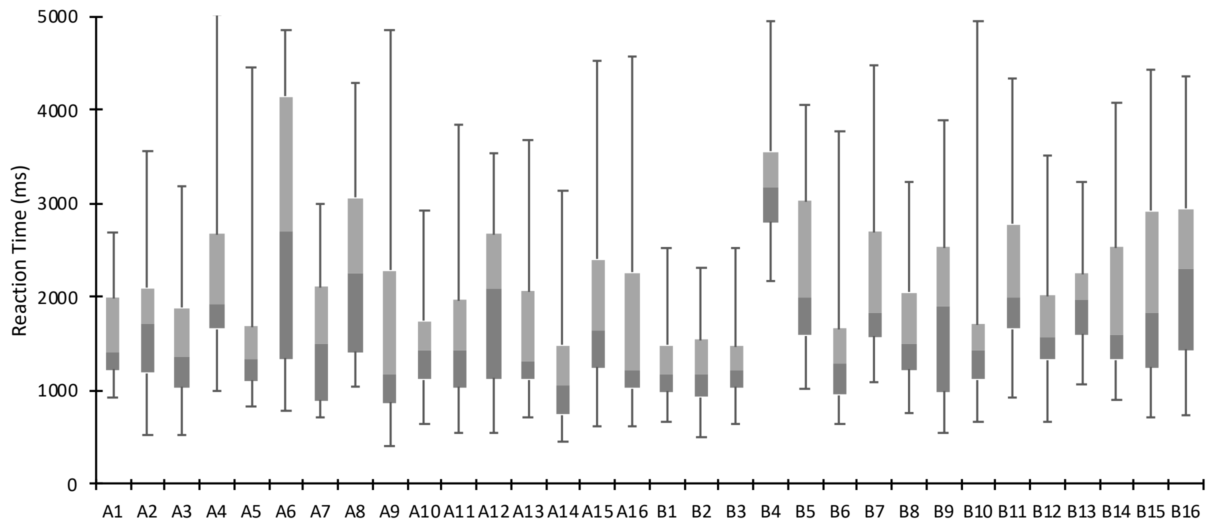


Figure 7. Box and whiskers show the individual reaction time distribution for the MTS test.

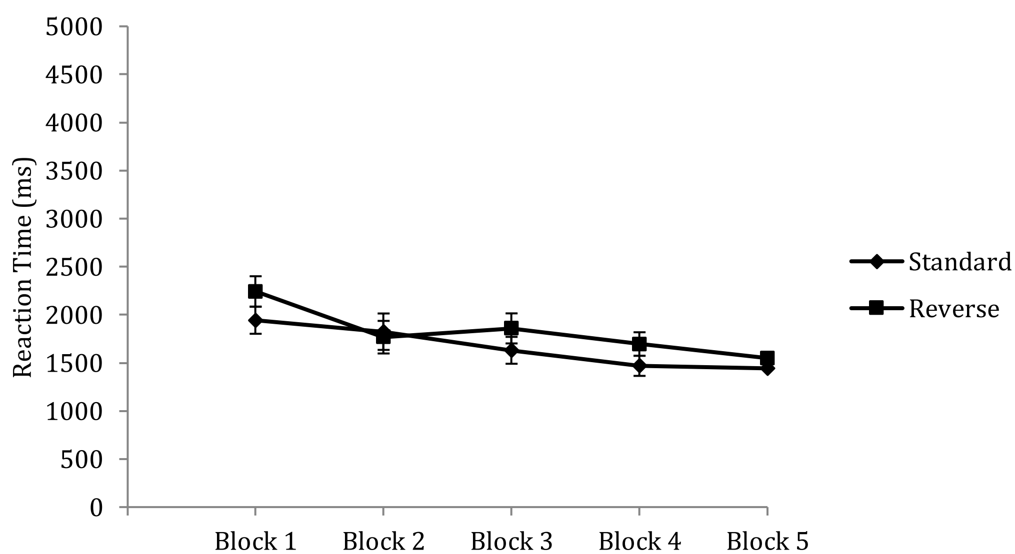


Figure 8. Mean of individual median reaction times per block during the MTS test. Error bars represent the standard error.

It is possible that there was no difference in reaction times because mediating events occurred during sample viewing before the participant brought up the comparisons, so sample viewing time was analyzed. The sample viewing time (i.e. time between sample onset and participant clicking on the sample to make the comparisons appear) results were also analyzed by a 5 (trial block) X 2 (condition: standard, reverse) mixed-model analysis of variance, with repeated measures on the trial block factor (see Figure 9). Due to a violation of sphericity according to Mauchly's test of sphericity, the Greenhouse-Geisser correction was used. There was significant main effect of trial block, $F(4, 120) = 26.145$, $p < .001$; as the blocks progressed, sample viewing times overall decreased. There was, however, no significant interaction between condition and trial block, $p = .096$.

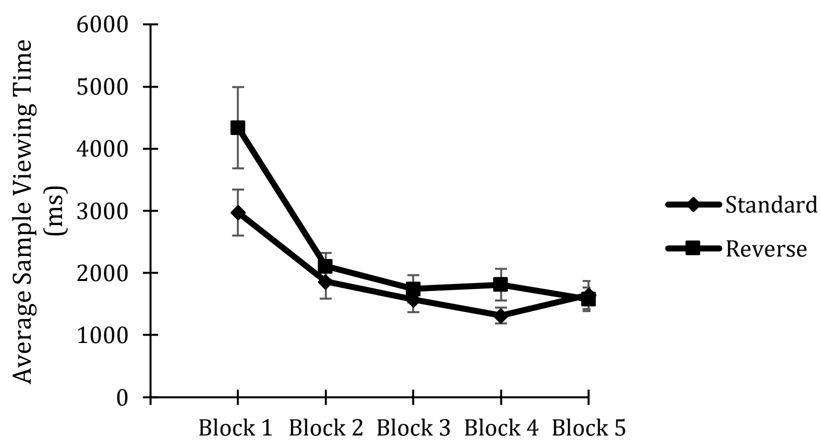


Figure 9. Mean of individual median sample viewing times per block during the MTS test. Error bars represent the standard error.

Independent samples t-tests were done on reverse intraverbal test scores, and listener test scores (see Figures 10 and 11). No significant differences were found between groups for either test, *p*s ≥ .202. A bivariate correlation revealed that reverse intraverbal test results were highly correlated with MTS accuracy for the reverse group, $r = .802$, $p < .001$, but not for the standard group, $r = .310$, $p = .242$ (see general discussion for the implications of these results). Additionally, bivariate correlations found intraverbal maintenance to be highly correlated with MTS accuracy in the standard group, $r = .652$, $p = .006$, and the reverse group, $r = .744$, $p = .001$; however, tact maintenance was not highly correlated with MTS accuracy for either group, $r\text{s} \leq .493$, $p\text{s} \geq .052$.

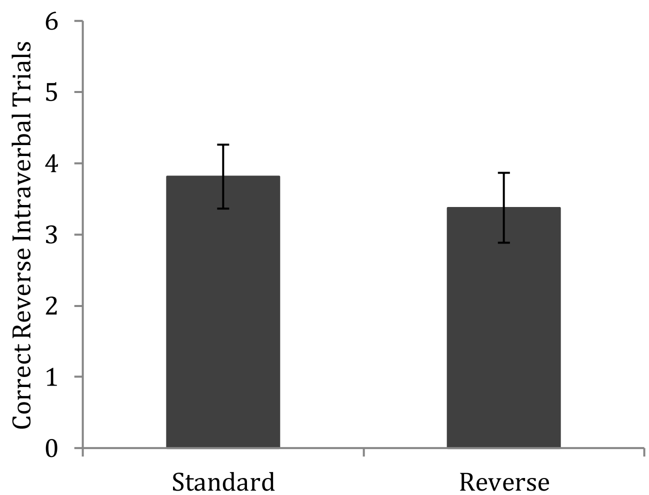


Figure 10. Average number of correct trials on the reverse intraverbal test. Error bars represent the standard error.

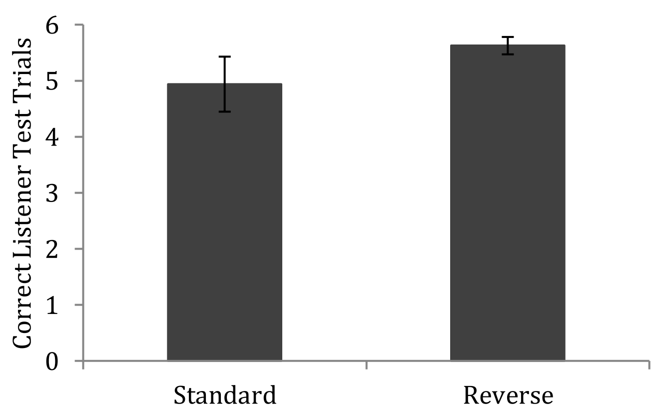


Figure 11. Average number of correct trials on the listener test.

Because many participants did not pass the retention tests, additional analyses were conducted excluding participants who did not get at least 5 out of the 6 trials correct on both retention tests (9 participants in the standard group and 10 participants in the reverse group met the criterion to be included in the analyses). An independent samples t-test was conducted on the MTS scores for participants meeting this criterion. There was no significant difference between standard (M = 25.22, SD = 4.06) and reverse groups (M = 23.00, SD = 6.50) for these participants, $t(17) = .882$, $p = .390$. An independent samples t-test was also conducted on the median MTS reaction times and sample viewing times for participants meeting the retention criterion. There were no significant differences between groups in either analysis ps ≥ .336, so even after removing potential variability due to poor baseline retention, the differences remained non-significant.

In summary, while group differences were in the expected direction, there were no significant differences between groups on MTS accuracy even though baseline retention was increased. Moreover, the difference in reaction times found in the previous study (Petursdottir et al, in preparation) was not replicated. This could be due to a floor effect considering reaction times were much faster in this study than in Petursdottir et al. (in preparation). It is possible that reaction times were faster because baseline tact and intraverbal relations were better rehearsed, so they occurred with greater speed at test, resulting in lower reaction times particularly for the reverse group that had to rely on their use. It is important to note that even though baseline retention was increased, scores on the MTS test were still low; only 10 participants performed with at least 90% accuracy. By contrast, adult participants in previous studies using the intraverbal naming task (Jennings & Miguel, 2017; Ma et al., 2016; Santos et al., 2015) have generally passed the MTS test, possibly because these studies included additional cues that may have encouraged responding in accordance with previously trained baseline relations at test. If low MTS accuracy is in some cases due to participants guessing or matching based on physical features instead of baseline relations, the experimental manipulation would not be expected to affect the reaction times of those participants; as a result, the effects of the manipulation might be more pronounced if more participants were responding in accordance with baseline relations. One way to encourage such responding might be to conduct training and testing in the same format (i.e., MTS). As a result, Experiment 2 was conducted as planned in spite of no effect being found in Experiment 1.

Experiment 2

Method

Participants and setting. Following approval by the Texas Christian University Institutional Review Board, 32 adults (24 females and 8 males, age range: 18 – 47 years, *M* = 20.81 years) were recruited from the Psychology department's human subjects pool via the research participation software *Sona-Systems*. Experimental sessions were conducted in the same room as experiment 1 with the same materials and investigators present.

Stimuli. The same stimuli from experiment 1 were used in experiment 2 with the exception of text stimuli being used as stimuli A1'-B3' instead of auditory stimuli and vocal responses.

Procedure. All procedures and design were the same as in experiment 1 with the exceptions of text stimuli being used in place of auditory stimuli and vocal responses and all training was conducted in MTS format, see Table 3 for a methodological overview.

Participants were randomly assigned to one of 2 groups: standard and reverse. There were 10 females and 6 males in the standard group and 14 females and 2 males in the reverse group; the mean age was 19.63 in the standard group and 22.00 in the reverse group. The standard group received tact instruction and then intraverbal instruction while the reverse group received the opposite. All procedures following training were identical across groups.

Table 3

Design Overview

Study	Group	Baseline Training I	Baseline Training II	Post- test 1	Post-test 2	Post-test 3	Post-test 4
	Standard (n = 16)	MTS tact training (A-A')	MTS intraverbal training (A'-B')	MTS test	MTS Reverse Intraverbal (B'-A')	MTS Baseline Retention (A-A', A'-B')	Listener test (A'-A, B'-B)
Study 2	Reverse (n = 16)	MTS intraverbal training (A'-B')	MTS tact training (A-A')	MTS test	MTS Reverse Intraverbal (B'-A')	MTS Baseline Retention (A-A', A'-B')	Listener test (A'-A, B'-B)

MTS tact training (AA'/BB'). The two MTS groups received MTS tact training instead of the vocal tact training. During MTS tact training, participants were taught to choose the correct text stimuli A1' through B3' when presented with the visual stimuli A1 through B3. Trials were presented in blocks of six, with one presentation of each stimulus per trial block. The order of stimulus presentation was varied by the experiment software within each trial block across trial blocks. Tact training began with an instruction on the computer screen that read, "You will see six images paired with words. Please attend carefully and try to remember what you see." To begin, participants were exposed to each visual stimulus followed immediately by the corresponding text stimulus. After being exposed to all stimuli, a second set of instructions appeared on the screen that read, "You will now see the pictures again, one at a time. Please click on the word that goes with each picture. If you choose the incorrect word, the correct one will be presented before the start of the next trial." A trial began with the presentation of one of the visual stimuli; participants were then presented with three comparison text stimuli (A text stimuli were comparisons for A visual stimuli, and B text stimuli were comparisons for visual stimuli) and had 3-s to click the correct visual stimulus. Comparisons were randomly assigned to locations on the screen in each trial. If the participant did not click on the correct comparison stimulus within the 3-s interval, the computer displayed a screen showing the correct stimulus. If the participant chose the correct response within the allotted time, the computer automatically moved to the next trial. Training continued in this manner until participants responded with 100% accuracy across 30 trials.

MTS intraverbal training (A'B'). During intraverbal training, participants were taught to select the word (text stimuli B1' through B3') when presented with the A1' through A3' words (text stimuli), respectively. Trials were presented in blocks of three, with one presentation of each stimulus per trial block. The order of stimulus presentation varied across trial blocks. Intraverbal training began with an instruction on the computer screen that reads, "You will see three pairs of words. Please attend carefully and try to remember what you see." To begin, participants were exposed to each text stimulus (stimuli A1' through A3') immediately followed by the corresponding text stimulus (stimulus B1' through B3'), respectively. After being exposed to all stimuli, a second set of instructions appeared on the screen that read, "You will now see one of the words in each pair. Please click on the other word in the pair. If you choose the incorrect word, the correct one will be presented before the start of the next trial." A trial began with the presentation of one of the A words; then three comparison words appeared, and participants had 3-s to click on the correct B word. Comparisons were randomly assigned to locations on the screen in each trial. If the participant did not choose the correct response within the 3 second interval, the computer displayed a screen showing the correct stimulus. If the participant chose the correct response within the allotted time, the computer continued to the next trial. Training continued in this manner until participants responded with 100% accuracy across 15 trials.

MTS test. The same MTS test from experiment 1 was used in experiment 2.

Reverse MTS intraverbal test (B'A'). The reverse MTS intraverbal test occurred immediately following the MTS test for the standard and reverse groups. One block of 6 trials was presented in the same manner as MTS intraverbal training with one exception: The B' text stimuli were presented instead of the A' text stimuli, and participants had to respond by clicking on the corresponding A' names.

Tact and intraverbal maintenance. Tact maintenance occurred immediately following the reverse intraverbal test. One block of six trials was presented in the same manner as tact training; however, no error correction was provided for incorrect responses, and no reinforcement was provided for correct responses. Intraverbal maintenance occurred immediately following tact maintenance. One block of six trials was presented in the same manner as intraverbal training; however, no error correction was provided for incorrect responses, and no reinforcement was provided for correct responses.

Listener test. The listener test used in experiment 2 was identical to the one used in experiment 1 with the exception of using text stimuli in place of auditory stimuli.

Results and Discussion

Baseline training and retention. An independent samples t-test was conducted total trials to criterion, and retention test scores. There were no significant differences between groups on total trials to criterion or retention test scores $ps \geq .542$ (see Table 4 for means and standard deviations; see Figures 12 and 13 for retention test and trials to criterion data). Each group did, however, receive more training trials than the other group in whichever type of training they received first (e.g. the standard group received more overall tact training trials than the reverse group); $ps \leq .034$. The proportion of participants who met the passing criterion for baseline retention was increased from 19 out of 32 participants in experiment 1 to 25 out of 32 participants in the current study. More specifically, 14 participants in the standard group and 11 participants in the reverse group passed both retention tests (at least 5 out of 6 correct trials). Additionally, mean MTS scores stayed consistent (standard *M* = 73%, reverse M = 59%) as compared to experiment 1 (standard *M* = 76%, reverse M = 63%).

Table 4

Means and Standard Deviations for reverse intraverbal, listener, and retention tests, and trials to criterion for experiment 2

	Standard Group		Reverse Group	
	Mean	Standard Deviation	Mean	Standard Deviation
Tact Trials to Criterion	148.25	71.39	94.81	64.92
Intraverbal Trials to Criterion	81.38	53.82	148.94	72.75
Total Trials to Criterion	229.63	91.44	243.75	82.25
Reverse Intraverbal Test	91.67%	14.91	87.50%	21.52
Tact Maintenance Test	92.71%	25.07	89.58%	15.96
Intraverbal Maintenance Test	90.63%	13.57	85.42%	30.96
Listener Test	79.17%	9.62	78.13%	14.55

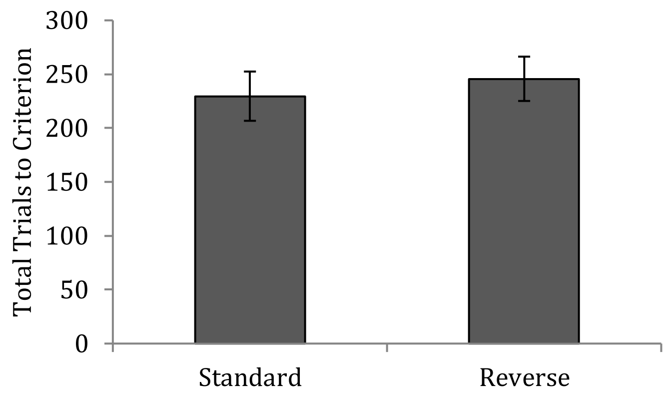


Figure 12. Average number of total training trials needed to reach mastery criterion. Error bars represent the standard error.

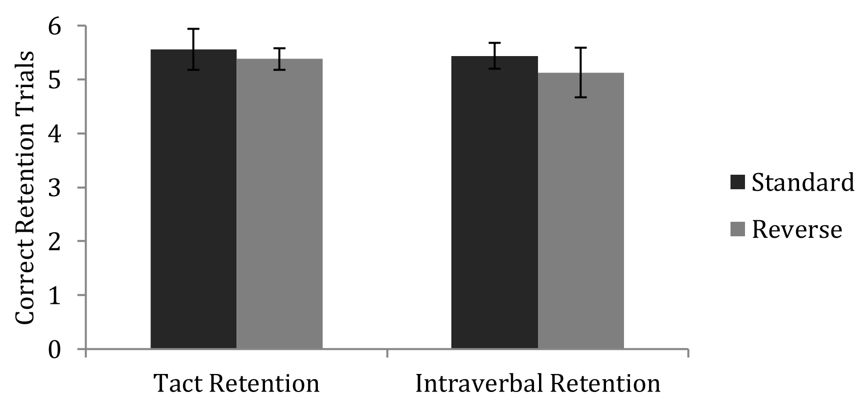


Figure 13. Average number of correct trials on the retention tests. Error bars represent the standard error.

MTS and other emergent relations. MTS accuracy was analyzed by a 5 (trial block) X 2 (condition: standard, reverse) mixed-model analysis of variance, with repeated measures on the trial block factor, see Figure 14. Due to a violation of sphericity according to Mauchly's test of sphericity, the Greenhouse-Geisser correction was used. Although mean accuracy was higher in the standard (M = 72.50%, SD = 31.40) than in the reverse group (M = 58.96%, SD = 30.28), there was no significant main effect of condition, $F(1, 30) = 1.59$, $p = .218$. There was a significant main effect of trial block, $F(2.70, 120) = 3.62$, $p = .020$; as the blocks progressed, MTS accuracy increased. There was, however, no significant interaction between condition and trial block, $p = .224$.

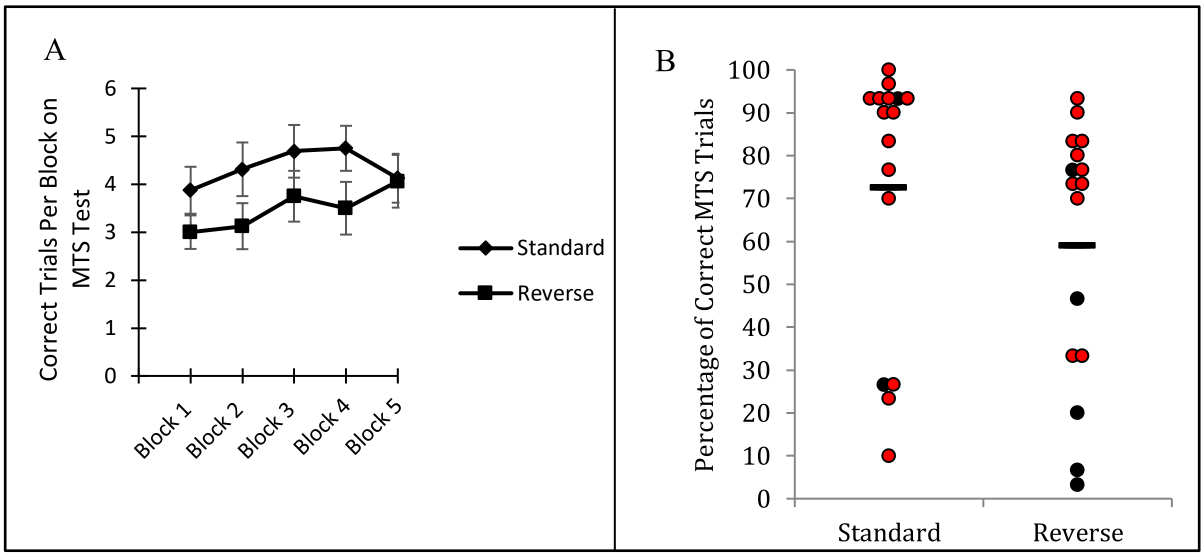


Figure 14. Panel A shows the number of correct trials per block during the MTS test. Error bars represent the standard error. Panel B shows the individual MTS averages. Red circles indicate participants who passed both retention tests (at least 5/6 correct trials).

Individual distributions for the MTS reaction times were positively skewed; because of this, median reaction times were used for analyses rather than the means (see Figure 15). Trials in which participants did not choose a comparison within the time allotted were scored as no response, and these trials were not included in the reaction time analyses. The reaction time results were also analyzed by a 5 (trial block) X 2 (condition: standard, reverse) mixed-model analysis of variance, with repeated measures on the trial block factor (see Figure 16). Due to a violation of sphericity according to Mauchly's test of sphericity, the Greenhouse-Geisser correction was used. There was a main effect of condition, $F(1, 30) = 6.339$, $p = .017$, the standard group ($M = 1453.72$, $SD = 345.30$) reacted faster than the reverse group ($M = 1831.00$, $SD = 403.87$). There was a significant main effect of trial block, $F(2.68, 120) = 9.203$, $p < .001$; as the blocks progressed, reaction times overall decreased. There was, however, no significant interaction between condition and trial block, $p = .333$.

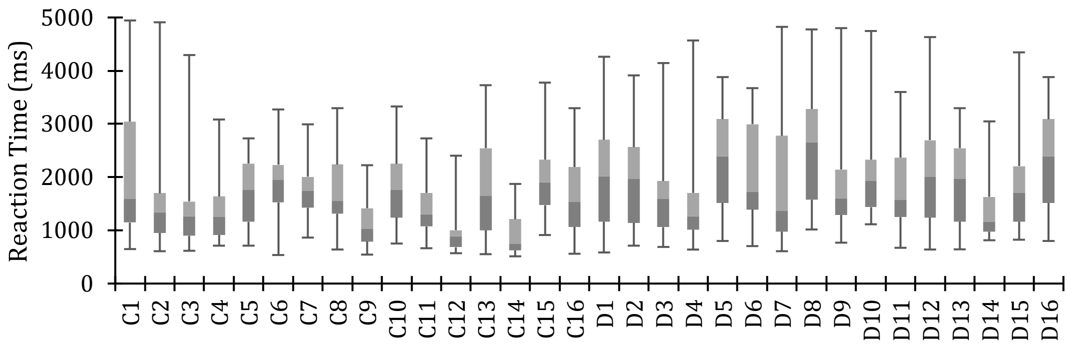


Figure 15. Box and whiskers show the individual reaction time distribution for the MTS test.

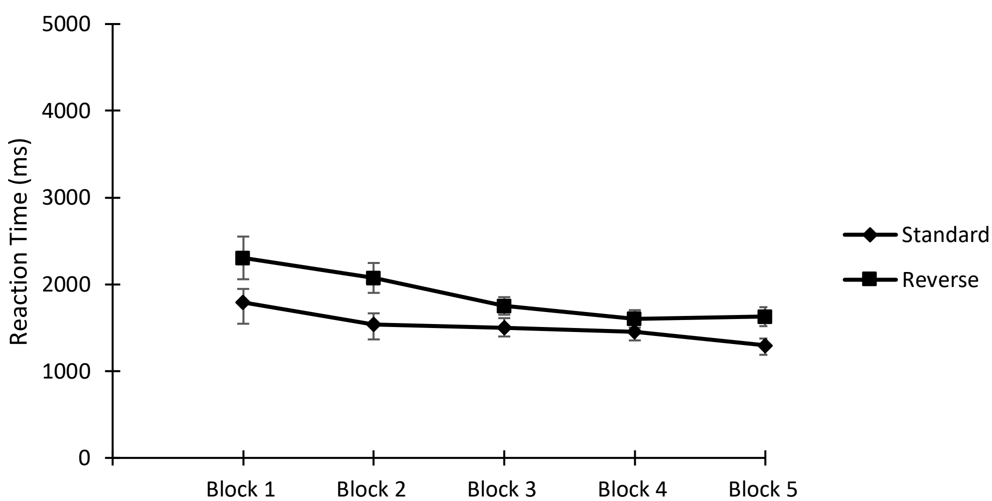


Figure 16. Average median reaction time per block during the MTS test. Error bars represent the standard error.

The sample viewing time results were also analyzed by a 5 (trial block) X 2 (condition: standard, reverse) mixed-model analysis of variance, with repeated measures on the trial block factor (see Figure 17). Due to a violation of sphericity according to Mauchly's test of sphericity, the Greenhouse-Geisser correction was used. There was significant main effect of trial block, $F(1.95, 120) = 15.150$, $p < .001$; as the blocks progressed, sample viewing times overall decreased. There was, however, no significant main effect of condition or interaction between condition and trial block, *p*s ≥ .378.

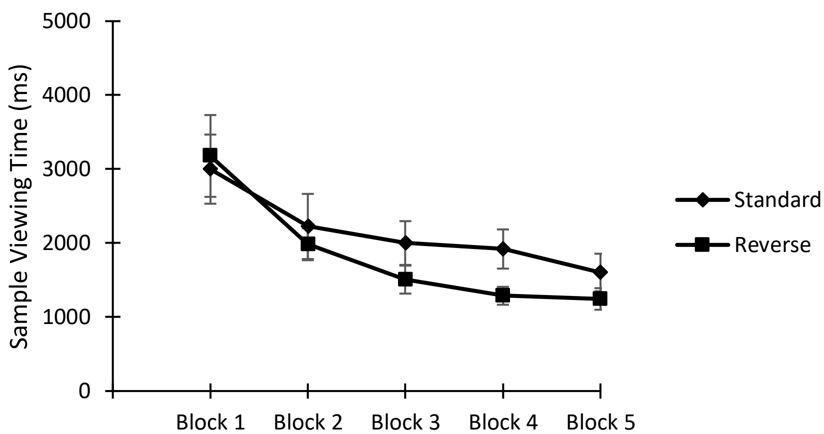


Figure 17. Average median sample viewing time per block during the MTS test. Error bars represent the standard error.

Independent samples t-tests were done on listener test scores and reverse intraverbal test scores (see Figures18 and 19). No significant differences were found between groups for either test, *p*s ≥ .529. A bivariate correlation revealed that reverse intraverbal test results were highly correlated with MTS accuracy for the reverse group, $r = .513$, $p = .042$, but not for the standard group, $r = .129$, $p = .634$ (see general discussion for the implications of these results).

Additionally, bivariate correlations found intraverbal maintenance to be highly correlated with MTS accuracy in the reverse group, $r = .603$, $p = .013$, but not the standard group, $r = .164$, $p = .544$. Additionally, tact maintenance was not highly correlated with MTS accuracy for either group, $r\text{s} \leq .313$, $p\text{s} \geq .114$.

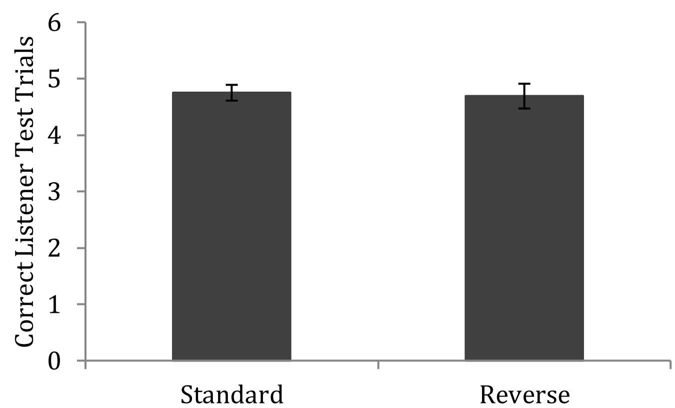


Figure 18. Average number of correct trials on the listener test. Error bars represent the standard error.

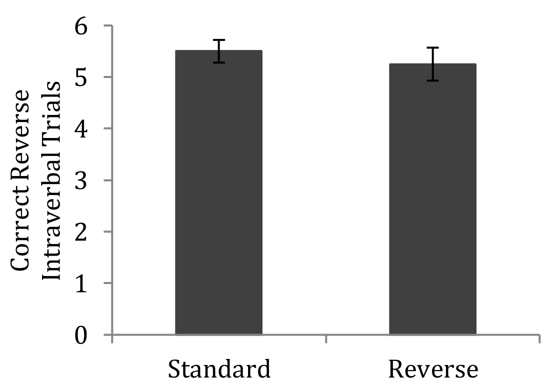


Figure 19. Average number of correct trials on the reverse intraverbal test. Error bars represent the standard error.

Because a few participants did not pass the retention tests, additional analyses were conducted excluding participants who did not get at least 5 out of the 6 trials correct on both retention tests (14 participants in the standard group and 11 participants in the reverse group met the criterion to be included in the analyses). An independent samples t-test was conducted on the MTS scores for participants meeting this criterion. There was no significant difference between standard (M = 74.29, SD = 30.65) and reverse groups (M = 71.82, SD = 20.30) for these participants, $t(23) = .230$, $p = .820$. An independent samples t-test was also conducted on the median MTS reaction times. Even when excluding participants who did not pass both retention tests, there was still a reaction time effect, $t(23) = 2.55$, $p = .018$.

In an exit interview, before being debriefed, participants were asked if there was anything they did to help them remember the word pairs. Only one of the participants in the standard group reported visualizing the visual stimuli. When asked the more leading question "When you were memorizing which words went together, did you at any point try to visualize the images together?", five more participants (six total) said yes. An independent samples t-test found that participants who reported using visualization performed significantly better on the MTS test than participants who reported not using visualization $t(14) = 2.67$, $p = .025$ (see Figure 20).

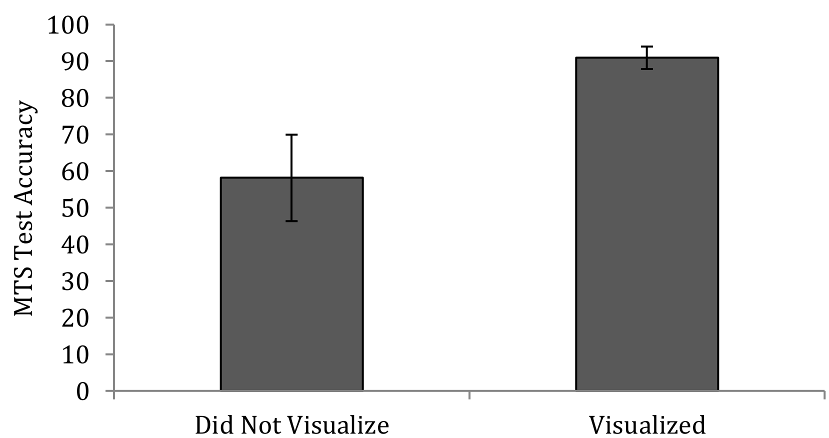


Figure 20. MTS accuracy for participants in the standard group who reported using or not using visualization during intraverbal training. Error bars represent the standard error.

In summary, while there was no effect of training sequence on MTS accuracy, there was a significant difference between groups on reaction times such that the standard group had faster reaction times than the reverse group, replicating previous findings (Petursdottir et al., in preparation). This is in line with the study hypothesis and suggests that participants may be using the "short cut" of visual mediation during the MTS test. The lack of effect on MTS score may be due to participants still having 5000 ms per trial in the test, which seems to be long enough for participants to verbally mediate and thus, also come to the correct answer just in a longer amount of time. Interestingly, changing the format of training to match testing did not increase scores on the MTS test.

General Discussion

The goal of Experiment 1 was to replicate the results of Petursdottir et al. (in preparation), in which training sequence affected reaction times in the intraverbal naming task, and to examine if increased baseline retention would be accompanied by an effect on MTS accuracy as well. The goal of increasing baseline retention was met, as 19 participants maintained baseline relations compared to 11 in Petursdottir et al. (in preparation). However, the results of the MTS test and reaction time analyses were not significant, even though group differences were in the expected direction when including all participants as well as when including only those who passed the retention tests.

In Experiment 2, however, the results of Petursdottir et al. (in preparation) were replicated using an MTS training task in place of the vocal tact and intraverbal instruction typically included in the intraverbal naming task. In this experiment, as in the previous study, the standard group did have a significantly faster reaction times than the reverse group, but there was not a significant effect on MTS accuracy.

It is not clear why the effect on reaction times was replicated in Experiment 2 but not in Experiment 1. Floor effects due to the task being easier and better practiced that the task in Petursdottir et al. (in preparation) are an unlikely explanation, as reaction times in Experiment 2 were just as fast, and faster in some blocks, as in experiment 1. An alternative explanation could have been the increase in baseline retention from Experiment 1 to Experiment 2, but baseline retention was even lower in Petursdottir et al. (in preparation) where the effect was still present. The most plausible explanation it seems is that the number of participants used in these studies (n = 32) is not sufficient to reliably detect this effect.

Because the study by Petursdottir et al. (in preparation) was close to finding a significant effect of training sequence on MTS accuracy, it was initially hypothesized that increasing baseline retention would result in the detection of a significant effect, but neither experiment found a significant effect of training sequence on MTS accuracy. The absence of an effect on MTS accuracy could be because, regardless of which group the participant is in, it is still possible to solve the task verbally, and the 5000 ms window given to respond may have been sufficient for the reverse group to perform as accurately as baseline retention permitted.

Although replicated in only one of the two experiments, the effect of training sequence on MTS reaction times supports the hypothesis that availability of visual representation during intraverbal training facilitates MTS performance. An analysis of the participants that reported using visualization during intraverbal training in the standard group of Experiment 2 found that participants who reported using visualization performed significantly better on the MTS test than participants who reported not using visualization. Future research in which participants are instructed to use visualization when learning intraverbal relations may shed some more light on this effect.

An alternative explanation, however, could be related to the number of relations trained in each phase. During tact training, six relations are taught while in intraverbal training, only three relations are taught. From a stimulus equivalence perspective, moving from tact to intraverbal training could prompt stimulus class merger (i.e., if tact training establishes six equivalence relations between images and verbal labels, intraverbal training might serve to merge these six two-member stimulus classes into three four-member classes that include two images and two labels); whereas moving from intraverbal to tact training would entail class expansion (i.e., expanding the previously learned relations from 2 to 4. If merger occurs more readily than expansion, then this could be an explanation for the finding in study 2 that participants in the reverse group have slower reaction times. However, there is no literature indicating that class merger should occur more easily than expansion, nor is it clear that this prediction would be derived from the major theory of stimulus equivalence (Sidman, 2000).

An interesting finding in both studies was that reverse intraverbal test accuracy was highly correlated with MTS test accuracy for the reverse groups but not for the standard groups. Since the reverse intraverbal test examines the emergence of bidirectional relations between the verbal (or text) relations, these results indicate that the emergence of bidirectional relations could be essential to MTS accuracy for reverse participants. This finding is consistent with Horne and Lowe's naming hypothesis (1996). They suggest that in a task such as those used in these studies, participants must use verbal mediation in order to accurately perform on the test, and that verbal mediation cannot occur without the understanding of the bidirectional intraverbal relations. For example, when presented with visual stimulus B3 in the MTS test, according to the verbal mediation hypothesis, the subject must be able to recall the word associated with the visual stimulus (B3'), then remember the other word associated with B3' (A3'), and finally remember the visual stimulus associated with A3’ (A3). If the bidirectional intraverbal relations do not emerge, then the participant would not be able to relate B3’ to A3’, and therefore, not be able to come to the correct answer, see Figure 21. Our finding of the strong correlation between reverse intraverbal test accuracy and MTS test accuracy in the reverse groups is in line with this prediction. More interesting however, is our finding that for the standard groups, there was not a significant correlation between these tests. This suggests that participants do not have to use verbal mediation in order to perform accurately on the MTS test. While this may still be a possibility for participants for which bidirectional intraverbal relations emerge, it is clearly not necessary. This could mean that participants in the standard group are using visual mediation while participants in the reverse group are using verbal mediation, as is consistent with our hypothesis.

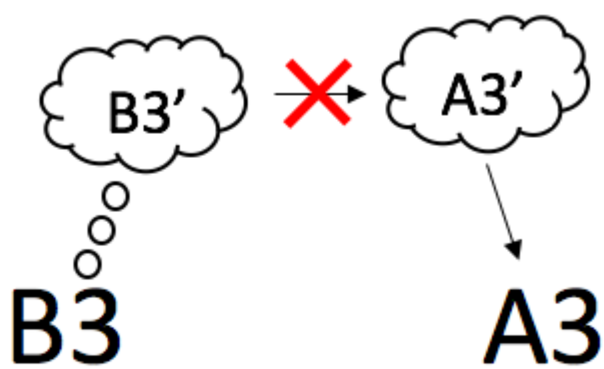


Figure 21. This figure depicts the verbal mediation chain and shows the broken link that occurs when bidirectional intraverbal relations do not emerge.

Another interesting finding is the overall similar MTS scores across studies. In study 1 the overall average MTS score was 69.4% compared to 67.1% in study 2. These scores are very low compared to other studies that use similar tasks with adult participants (Jennings & Miguel, 2017; Ma et al., 2016; Santos et al., 2015). These lower MTS scores could be due to the use of arbitrary stimuli rather than stimuli that participants may be more familiar with increasing the difficulty of the task; however, Santos et al. (2015) used arbitrary stimuli as well and still found high MTS test accuracy. Another possible reason is that participants in prior studies reviewed baseline relations before testing which could have increased baseline retention; baseline review was omitted in the present study because it would have compromised the integrity of the independent variable. Finally, in the present study, intraverbal training consisted of simply training participants to respond to one verbal label by providing (Experiment 1) or selecting (Experiment 2) another, whereas in previous studies, intraverbal training consisted of filling in statements that contained additional cues (e.g., "Isa goes with . . ." in Santos et al., 2015, and "The state for cardinal is . . . in Ma et al., 2016) that may have cued participants to later match the visual stimuli accordingly. Although it was hypothesized that training and testing in the same format in Experiment 2 would have a similar effect, it apparently did not.

Additionally, these data suggest that there is no difference in testing regardless of using vocal or MTS training. There are, however, a couple of benefits to the MTS training format. First, training took overall less time (about 15-20 minutes) as compared to about 30 minutes for vocal training. Additionally, participants seemed more comfortable during MTS training because they were not getting feedback from the experimenter. During vocal training, many participants seemed mildly embarrassed when receiving corrective feedback. Using MTS training rendered this type of feedback from an experimenter unnecessary. Because performance was similar across experiments, it may be hypothesized that selecting a pronounceable textual stimulus is functionally similar to having a participant vocally name an item. This means it should be safe to substitute computerized MTS training for the actual teaching of vocal responses in the intraverbal naming task.

One limitation of this study is the use of arbitrary stimuli. While this is ideal for experimental control, it is more difficult to speculate the extent to which this effect would be present in real world situations, such as if stimuli had been things participants experience in the real word (e.g. names of state birds and flowers). Future research should try to replicate these results with "real world" stimuli in order to make broader predictions about how this effect exists in more practical situations.

Conclusion. Although the results were inconsistent across experiments, the results from Experiment 2 support the notion that emergent relations between visual stimuli in associative concept formation is sensitive to the sequence in which baseline relations are learned. In turn, this suggests that different processes may underlie concept formation depending on the details of the history that produces it. The effect seen in Experiment 2 and a previous study (Petursdottir et al., in preparation) was hypothesized to reflect that correct matching in the reverse group depended on verbal problem solving strategies that were unnecessary for the standard group because it had a visual representation of the correct response immediately available without verbal problem solving. Although other explanations are possible, findings consistent with this hypothesis include the correlation between MTS accuracy and emergence of reverse intraverbal relations only in the group hypothesized to rely on verbal problem-solving, as well as the superior performance of participants who reported awareness of visualization during intraverbal training. The results may have implications for theories of stimulus equivalence and related phenomena, as well as implications for building repertoires of concept formation and inference-making.

References

Barnes-Holmes, D., Barnes-Holmes, Y., Smeets, P. M., Cullinan, V., & Leader, G. (2004). Relational frame theory and stimulus equivalence: Conceptual and procedural issues. *International Journal of Psychology and Psychological Therapy, 4*, 181-214.

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

Cook, R. G., Cavoto, B. R., Katz, J. S., & Cavoto, K. K., (1997). Pigeon perception and discrimination of rapidly changing texture stimuli. *Journal of Experimental Psychology: Animal Behavior Processes, 23,* 390-400.

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

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

Herrnstein, R. J., & Loveland, D. H. (1964). Complex visual concept in the pigeon. *Science*, *146*, 549-551. doi: 10.1126/science.146.3643.549

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. doi:10.1901/jeab.1996.65-185

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. http://doi.org/10.1901/jeab.2004.81-267

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

Keller, F. S., & Schoenfeld, W. N. (1950) *Principles of psychology.* New York: Appleton-Century-Crofts.

Kisamore, A. N., Carr, J. E., & LeBlanc, L. A. (2011). Training preschool children to use visual imagining as a problem-solving strategy for complex categorization tasks. *Journal of Applied Behavior Analysis, 44*, 255-278. doi:10.1901/jaba.2011.44-255

Leader, G., Barnes, D., & Smeets, P. (1996). Establishing equivalence relations using a respondent-type training procedure. *The Psychological Record, 46*, 685-706. https://doi.org/10.1007/BF03395192

Lionello-DeNolf, K. M. (2009). The search for symmetry: 25 years in review. *Learning & Behavior, 37*, 188-203. https://doi-org.ezproxy.tcu.edu/10.3758/LB.37.2.188

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*, 527–549. http://doi.org/10.1901/jeab.2002.78-527

Ma, M. L., Miguel, C. F., & Jennings, A. M. (2016). Training intraverbal naming to establish equivalence lass performances. *Journal of the Experimental Analysis of Behavior*, *105*, 409-426. doi: 10.1002/jeab.203

Mandell, C., & Sheen, V. (1994). Equivalence class formation as a function of the pronounceability of the sample stimulus. *Behavioural Processes, 32*, 29-46. http://dx.doi.org/10.1016/0376-6357(94)90025-6

Moustakis, I. S. & Mellon, R. C. (2018). Transitivity as skinnerian problem solving controlled by self-constructed relational stimuli. *Journal of the Experimental Analysis of Behavior,* 1-23. https://doi.org/10.1002/jeab.473

Palmer, D. C. (2012). The role of atomic repertoires in complex behavior. *The Behavior Analyst, 35,* 59-73.

Petursdottir, A. I., Carp, C. L., Peterson, S. P. & Lepper, T. L. (2015), Emergence of visual–visual conditional discriminations. *Journal of the Experimental Analysis of Behavior, 103*, 332–348. doi:10.1002/jeab.136

Petursdottir, A. I., Cox, R. E., Mellor, J. R., & McKeon, C. (in preparation). Effects of instructional sequence on emergent conditional discriminations.

Randell, T., & Remington, B. (2006). Equivalence relations, contextual control, and naming. *Journal of the Experimental Analysis of Behavior, 86*, 337–354. http://doi.org/10.1901/jeab.2006.82-05

Santos, P. M., Ma, M. L., & Miguel, C. F. (2015). Training intraverbal naming to establish matching-to-sample performances. *The Analysis of Verbal Behavior, 32*, 162-182. https://doi-org.ezproxy.tcu.edu/10.1007/s40616-015-0040-4

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

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

Silva, R. A., & Debert, P. (2017). Go/no-go procedure with compound stimuli with children with autism. *Journal of Applied Behavior Analysis, 50*, 750-755. https://doi.org/10.1002/jaba.421

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

Swisher, M., & Urcuioli, P. J. (2013). Symmetry in the pigeon with sample and comparison stimuli in different locations. *Journal of the Experimental Analysis of Behavior, 100,* 49-60. doi: 10.1002/jeab.31

Urcuioli, P. J., & Swisher, M. (2015). Transitive and anti-transitive emergent relations in pigeons: Support for a theory of stimulus-class formation. *Behavioural Processes, 112,* 49-60. doi: 10.1016/j.beproc.2014.07.006

Zentall, T. R., Clement, T. S., & Weaver, J. E. (2003) Symmetry training in pigeons can produce functional equivalences. *Psychonomic Bulletin and Review, 10,* 387-391. https://doi.org/10.3758/BF03196496

Zentall, T. R., Galizio, M., & Critchfield, T. S. (2002). Categorization, concept learning, and behavior analysis: An introduction. *Journal of the Experimental Analysis of Behavior. 78*, 237-248. doi:10.1901/jeab.2002.78-237

VITA

Reagan E. Cox

EDUCATION

2014-2016	Angelo State University, San Angelo, TX B.S. Psychology
2017-Present	Texas Christian University, Fort Worth, TX Graduate Student, Ph.D. program in Experimental Psychology

CURRENT RESEARCH INTERESTS

- Experimental Analysis of Behavior
- Language acquisition of children with Autism

PUBLICATIONS

McGee, R. (2016). Birth order and anxiety levels in young adults with siblings with special needs. *Angelo State University Social Sciences Research Journal, 1*(4).

MANUSCRIPTS IN PREPARATION

Petursdottir, A. I., **Cox**, R. E., Mellor, J. R., & McKeon, C. (2018) Effects of instructional sequence on emergent conditional discriminations.

CONFERENCE PRESENTATIONS

Cox, R. E., Petursdottir, A. I., Mellor, J. R., & McKeon, C. (2018, May). *Effects of training sequence on stimulus class formation via intraverbal naming.* Association for Behavior Analysis International Annual Convention, San Diego, CA.

Molina-Lomeli, C., **Cox**, R. E., & Petursdottir, A. I., (2018, April). *Effects of an echoic response requirement on object naming.* Michael and Sally McCracken Student Research Symposium, Texas Christian University, Fort Worth, TX.

Cox, R. E. (2018, March) *Effects of instructional sequence on emergent conditional discriminations.* Texas Association for Behavior Analysis, Houston, TX.

Petursdottir, A. I., **Cox**, R. E., Mellor, J. R., & McKeon, C. (2018, February). *A training sequence effect on emergent conditional discriminations*. Winter Conference on Animal Learning and Behavior, Logan, UT.

McGee, R. (2017). Birth Order and Anxiety Levels in Young Adults with Siblings with Special Needs. *Angelo State University Social Sciences Research Journal.* Annual Convention of the Southwestern Psychological Association, San Antonio, TX.

TEACHING EXPERIENCE

2017-Present	**Principles of Behavior I**, Lab Instructor *Texas Christian University*
2018	**Applied Behavior Analysis,** Teaching Assistant (including guest lectures)

ABSTRACT

EFFECT OF TRAINING SEQUENCE ON EMERGENT STIMULUS RELATIONS

by Reagan E. Cox, B.S., 2016
Department of Psychology
Texas Christian University
Thesis Advisor: Anna I. Petursdottir, Associate Professor of Psychology

The purpose of these experiments was to compare two instructional sequences on emergent relations between visual stimuli. Participants were college students between the ages of 18 and 47 years. In each of two experiments, 16 participants were assigned to a standard group that received training to relate visual stimuli to verbal labels prior to training to relate pairs of labels, and 16 participants to a reverse group that received the opposite training sequence. Emergent relations between visual stimuli were assessed in match-to-sample (MTS) format. In line with previous findings, we predicted that the standard group would perform with greater speed on the MTS test. Experiment 1 attempted to replicate the results of a prior study in which baseline relations were trained as vocal tacts and intraverbals, while taking steps to increase baseline retention from the prior study. Baseline retention increased but although between-group differences in reaction times were in the predicted direction they were not statistically significant. In Experiment 2, vocal tact and intraverbal training were replaced with analogous MTS tasks involving visual stimuli and print labels, and the standard group had significantly faster reaction times during testing than the reverse group.