FURTHER ANALYSIS OF THE ORDER OF STIMULUS PRESENTATION IN RECEPTIVE LANGUAGE INSTRUCTION

by

Kiley A. Hiett

Submitted in partial fulfillment of the

requirements for Departmental Honors in

the Department of Psychology

Texas Christian University

Fort Worth, Texas

May 2, 2016

FURTHER ANALYSIS OF THE ORDER OF STIMULUS PRESENTATION IN RECEPTIVE LANGUAGE INSTRUCTION

Project Approved:

Supervising Professor: Anna I. Petursdottir, Ph.D.

Department of Psychology

Kenneth Leising, Ph.D.

Department of Psychology

Dan Williams, Ph.D.

Department of Honors

ABSTRACT

Receptive language training programs are commonly used in clinical settings to build the vocabularies of both typically developing children and those with autism. Conflicting recommendations exist in the literature regarding the optimal order of stimulus presentations in receptive instructional trials. The present study was a systematic replication of previous research (Petursdottir & Aguilar, 2016) that found an advantage of presenting the auditory sample before the visual comparisons in typically developing children. Unlike the previous study, the present study incorporated an error correction procedure in the form of a prompted trial following incorrect responses. Four typically developing 5- and 6-year-old boys participated. A sample-first condition presented the auditory sample before the visual comparisons. The comparison-first condition presented the comparisons before the sample. Acquisition in the two conditions was compared in an adapted alternating-treatments design with a replication across stimulus sets. Unlike the previous study, the present study did not show a consistent advantage of the sample-first condition. More research is needed to determine the contribution of stimulus presentation, prompting procedures, and other components of instructional trials to acquisition in receptive tasks.

Further Analysis of the Order of Stimulus Presentation in Receptive Language Instruction

Receptive language instructional programs designed for students with autism spectrum disorder (ASD) or developmental delays (e.g., Taylor & McDonough, 1996) commonly address auditory-visual conditional discriminations; a skill that is essential for listeners to respond appropriately to spoken words. In an auditory-visual conditional discrimination, reinforcement for approaching or selecting a certain visual stimulus (e.g., a pen) is conditional on the prior presentation of a certain auditory stimulus (e.g., the spoken word "pen"; Carp, Peterson, Arkel, Petursdottir & Ingvarsson, 2012). Programs designed to teach these listener skills commonly employ a matching-to-sample format (MTS; Green, 2001), in which an auditory sample stimulus (e.g., the word "pen") and multiple visual comparison stimuli (e.g., a pen, a book, and a ruler) are presented in a single trial. The goal is to build the student's comprehension of relationships between spoken words and their referents: a crucial component of language acquisition. As a result, these programs are instrumental for teaching children with language delays skills that commonly appear in typically developing children through social interactions (Hart & Risley, 1975) and have implications for multiple facets of their education (Leaf & McEachin, 1999).

Although there is ample knowledge on best practices in the applications of receptive language instruction (Green 2001; Grow & LeBlanc, 2013), conflicting recommendations exist regarding the most effective order of stimulus presentation within trials in MTS procedures. Early intervention curricula tend to suggest arranging the visual comparison stimuli in front of the participant before the auditory sample stimulus is presented (Leaf & McEachin, 1999; Taylor & McDonough, 1996). By contrast, Green (2001) recommended presenting the auditory sample stimulus prior to introducing the visual comparisons in each trial, which is consistent with what is typically done in laboratory research that uses MTS procedures with humans and nonhumans.

In tentative support of the first recommendation, two cases have been reported in the literature in which students who failed to learn auditory-visual discriminations when the sample was presented first benefitted from switching to comparison-first presentation (Carp et al., 2012; Doughty & Saunders, 2009). By contrast, a recent study by Petursdottir and Aguilar (2016) supports Green's (2001) recommendation to present the sample first. This study, which represented the first systematic comparison of the two presentation formats, examined whether sample-first or comparison-first stimulus presentation methods would produce faster acquisition rates in typically developing children who were learning the names of birds and country flags. The results showed that mastery was consistently met quicker in the sample-first condition across six evaluations with three participants. However, the instructional method used by Petursdottir and Aguilar (2016) was not entirely representative of those typically used in receptive language instruction in applied settings. Specifically, the only teaching procedure used in this study was differential consequences, which consisted of a computer animation and a sound clip lasting 4 s for correct responses and a black screen presented for 4 s for incorrect responses. In other words, it was a trial-and-error procedure and incorporated no type of prompting or remedial feedback, which may limit the generality of the results. Although trial-and-error procedures are common in research settings, receptive language instruction in clinical and educational settings usually includes prompting and prompt-fading or error correction procedures. As a result, it is not known if an advantage of sample-first presentation would be seen when these additional instructional components are included. The current study aimed to extend Petursdottir and Aguilar's (2016) work by investigating whether the sample-first advantage would still be observed when incorrect responses were followed by an effective error correction procedure.

Error correction procedures are frequently employed in instructional protocols as a means to reduce the number of errors that occur during training. An error correction event refers to any event that is presented contingent on error responses during instruction (Rodgers & Iwata, 1991). As such, the black screen presented following errors in Petursdottir and Aguilar (2016) is an example of an error correction event. However, error correction events in clinical settings are often instructional in nature and may involve, for example, providing the student with the correct answer, prompting the student to make a correct response, or having the student practice the correct response repeatedly (e.g., Barbetta, Heward, & Bradley, 1993; Carroll, Joachim, St. Peter, & Robinson, 2015; Leaf, Leaf, Taubman, McEachin, & Delmolino, 2014b; McGhan & Lerman, 2013).

In contrast to methods that utilize positive reinforcement only through praise or other desired items, error correction procedures may function to negatively reinforce correct responses. In a study designed to investigate different mechanisms for error correction in the context of discrimination training using a matching task, Rodgers and Iwata (1991) found that adult participants with developmental delays performed better in the conditions that employed negative reinforcement, where correct responding resulted in the avoidance of the error correction procedures, in comparison to the differential reinforcement condition, which only used positive reinforcement for correct responses. The two error correction conditions, one labeled the "practice" condition which used repetition of the incorrect trial with reinforcement prompt fading, and one labeled "avoidance" where an unrelated matching task was presented after incorrect trials, showed similar levels of success in 5 of the 7 participants. Rodgers and Iwata (1991) hypothesized that negative reinforcement by avoidance of the error correction task played an important role in the effects of error correction procedures on acquisition rates. They

also noted that while their data suggest that error correction procedures can improve performance compared to differential reinforcement alone, individual history may play a part in which method works best for a particular learner.

Although Rodgers and Iwata's (1991) data suggest that an effective error correction procedure does not need to be directly related to the learning task, most commonly used error correction procedures involve re-presenting an error trial with a prompt to respond correctly (e.g., McGhan & Lerman 2013). A prompt is defined as an "instructional technique used to help students make correct responses" until they acquire the ability to perform this correct responding independently (MacDuff, Krantz, & McClannahan, 2001; p. 38) and it may consist, for example, of showing or telling the learner how to perform the task at hand. A number of studies have shown positive effects of error correction procedures using a variety of prompts on acquisition in discrete-trial learning tasks. For example, Carroll et al. (2015) compared the effectiveness of four popular error correction procedures used in discrete trial instruction with five- to eight-year-old children with attention deficit hyperactivity disorder and ASD. By comparing acquisition data in an adapted alternating treatments design, they found that while one or two of the error correction procedures produced faster acquisition for each participant, multiple procedures proved to be effective in some capacity for all individuals. As another example, Kodak, Fuchtman, and Paden (2012) compared the effects of echoic prompting with error correction, tact prompting with error correction, and cues-pause-point (CPP) procedure on increasing question-answering in two participants with autism and found that prompting plus error correction procedures was more effective than the CPP procedure.

In receptive instruction, error correction procedures involving picture prompts may be more effective than error correction procedures involving other types of prompts. Picture

prompts (e.g., Fisher, Kodak, & Moore, 2007) are used when the teacher presents an identical picture to the positive comparison when an incorrect response occurs in receptive language training, names the item, and then repeats the initial instruction. For example, a clinician presents a printed image of a pen in addition to the positive and negative comparison images, says "pen", and then instructs the learner to "choose the one you heard." Fisher and his colleagues compared the effects of presenting picture prompts following all errors to a condition in which errors were followed by pointing prompts, and a no-reinforcement control condition, in children with ASD who had histories of failure in receptive language training. Results from two participants suggested that the picture prompt employed in the identity matching task facilitated acquisition over the least-to-most procedure and the control (Fisher et al., 2007). Carp, Peterson, Arkel, Petursdottir, & Ingarsson (2012) replicated the work of Fisher et al. (2007) using more stringent stimulus presentation methods and a new control condition, and extended them to children with histories of success in receptive language training. Using a sample-first presentation structure, Carp et al. (2012) found that picture prompts improved acquisition in four children with ASD more rapidly than the pointing prompt condition, and more rapidly than a differential reinforcement (i.e., trial-and-error) condition used for two of the participants. Together, these two studies (Carp et al., 2012; Fisher et al., 2007) suggest that picture prompts may be a particularly effective error correction procedure in receptive instruction. As noted earlier, one participant in Carp et al. (2012) did not reach mastery in any of the conditions until stimulus presentation was modified from a sample-first to a comparison-first procedure, at which time he quickly reached mastery in the picture-prompt condition. . However, because additional modifications were also implemented (i.e., prompting the participant to look at each comparison stimulus before the sample was presented), it is not certain that the change was directly related to

the order of stimulus presentation in this case.

The current study aimed to extend the work of Petursdottir and Aguilar (2016) through the addition of a picture prompt error correction procedure to both the sample-first and comparison-first conditions. Four typically developing children of kindergarten age completed receptive identification tasks using the same procedures as Petursdottir and Aguilar (2016) with the addition of a picture prompting procedure contingent on an incorrect response. The format of the prompting procedure was identical to that of the training procedures, by being presented in either the sample-first or comparison-first format. It was predicted that a sample-first advantage would be seen as in the previous study, in spite of the addition of this instructional component.

Method

Participants

The participants were four kindergarten and first-grade age male students enrolled in an after-school care program at a local public elementary school. Matthew (pseudonym) was 5 years, 8 months at the beginning of the study and of Hispanic background. He scored an age equivalence of 4 years, 10 months on the Peabody Picture Vocabulary Test – 4th edition (PPVT-4). Rory (pseudonym) was 5 years, 10 months at the beginning of the study and of Caucasian-American background. He scored an age equivalence of 6 years, 5 months on the PPVT-4. Danny (pseudonym) was 6 years, 0 months at the beginning of the study and of Caucasian-American background. He scored an age equivalence of 7 years, 7 months on the PPVT-4. Peter (pseudonym) was 6 years, 7 months at the beginning of the study and of Hispanic background. He scored an age equivalence of 6 years, 9 months on the PPVT-4. None of them had been diagnosed with any sort of developmental disorder and all reported English as the primary language at home.

Sessions were held either in the library or at a table in the hallway of the school after hours. In both instances, the participant and experimenter sat at a rectangular table side by side during sessions. A group of students in the after-care program and their staff member were sometimes present as they passed through the hallway, or students from the school entered the library occasionally. The experimenter met with each participant 3-5 times a week, and conducted two to four 16-trial sessions with 5-minute breaks in between each.

Stimulus and Apparatus

Visual and auditory stimuli were presented in automated MS PowerPoint slideshows on a HP® EliteBook laptop with a 13" monitor. The presentation and timing of stimuli and corresponding consequences were controlled by hyperlinks, animations, and slide transitions. In other words, once the slideshow was started the participant could click through independently and the experiment was presented appropriately for all 16 trials. Participants used a BIGtrack™ trackball mouse to navigate through the slides.

The visual stimuli consisted of 12 color photos of birds against a blue background, 8 color images of country flags against a gray background, and 8 color photos of state flowers against a background of foliage. All stimuli were 240 x 240 pixels when the slide show was in presentation mode. Auditory stimuli were recordings of the English names associated with the chosen visual stimuli. A female voice was used for the recordings of bird and state flower names, and a male voice was used for the recordings of country flag names. All sound recordings were edited to be 1670 ms in duration.

The stimuli were split into groups of four visual stimuli with their corresponding auditory stimuli, with three groups of birds, two groups of country flags, and two groups of state flowers (shown in Table 1). Each participant was assigned one group of birds and one group of flags for

both the sample first and the comparison first experimental conditions, equating to four stimulus groups total. Groups 1 and 2 of birds and flags were initially counterbalanced to both conditions across all participants. However due to significantly high levels of responding during baseline, Peter had his Group 2 birds replaced with Group 3 birds and both of his groups of flags replaced with Groups 1 and 2 of state flowers. Table 2 depicts the assignment of stimulus groups across participants and conditions.

Procedure

Experimental Design. Acquisition in both conditions for Matthew, Rory, and Danny was compared in an adapted alternating-treatments design (Sindelar, Rosenberg, & Wilson, 1985) within a multiple baseline design across stimulus groups, which was identical to what was done in the initial study (Petursdottir & Aguilar, 2016). For Peter, however, his responding above chance levels in baseline sessions for the flags stimuli resulted in the experimenters having to create two new stimulus groups (state flowers). This meant that his acquisition for the sample-first condition and the comparison-first condition was compared in an adapted alternating-treatments design (Sindelar, Rosenberg, & Wilson, 1985) within two separate AB experimental designs.

Sample-First Condition. Baseline and instruction trials started with a blue square presented in the middle of the upper third of the computer screen with a white background. If the participant did not respond by clicking the blue square within 5 s of the slide show starting, the experimenter verbally instructed him to “click the blue square”. When the participant clicked the square, a blank white screen appeared and the computer presented the auditory sample stimulus (e.g., “China”). Once the auditory stimulus ended, four visual comparison stimuli (e.g., four

flags) were presented horizontally across the bottom third of the white screen. These consisted of one positive and three negative comparisons evenly spaced (1 cm) apart.

Comparison-First Condition. Baseline and instruction trials began similarly to the sample first conditions, however the blue square was positioned in the middle of the bottom third of the screen. When the participant clicked the square it disappeared and four visual comparison stimuli, also one positive and three negative, were presented horizontally across the bottom third of the screen spaced equidistantly (1 cm). 1670 s after the presentation of the comparison stimuli, the auditory stimulus was presented while the visual stimuli remained on the screen. If the participant clicked one of the comparison stimuli before the auditory sample was fully presented, the experimenter restarted the trial and instructed the participant to "listen to the whole name before you pick". These early responses were noted but only the responses that occurred after the auditory sample was presented were recorded as correct or incorrect.

Mouse Training. A short four-trial mouse training was implemented before the study began in order to ensure that participants could use the BIGtrack™ trackball mouse appropriately. The sample stimuli for these trials consisted of everyday objects (e.g., basketball and blocks). Two of the trials were presented in the sample first condition and two in the comparison first condition as described above, with the exception that only two comparison visual stimuli were present. Clicking the positive comparison in each trial resulted in a 4 s presentation of a positive reinforcement slide (e.g., computer animation of Spiderman and a trumpet sound). Clicking a negative comparison resulted in a 4 s presentation of a plain black screen. For the first trial, the experimenter modeled using the trackball to move the cursor around the screen, and then verbally instructed the participant to "click the blue square". After the comparisons and auditory sample had been presented, the experimenter instructed the participant

to “click the picture that matches what you heard”. No modeling or instruction was given for the other three trials, except for prompting the participant to click the blue square if necessary (in the case of no responding within 5 s).

Baseline. No feedback was given for correct or incorrect responses during the baseline trials. Clicking on any of the four comparison stimuli simply advanced the presentation on to the next trial. A correct response was recorded if the participant clicked on the positive comparison. An incorrect response was recorded if the participant clicked on any of the negative comparisons. If the participant failed to respond within 10 s after the end of the auditory sample, the presentation advanced on to the next trial and it was recorded as incorrect response. All baseline sessions were composed of 16 trials, with four blocks of four trials each. The stimuli that made up one of these 16 trial sessions were all from the same stimulus group (birds, flags, or state flowers). Each of the four visual stimuli was used as a positive comparison 4 times and as a negative comparison 12 times, and the order of presentation was counterbalanced throughout the trials. It was also ensured that each visual stimulus was shown in each of the four possible comparison positions: once in every position as a positive comparison and three times in every position as a negative comparison. To ensure that participants could not memorize the order of trials, two session types were created where the order of auditory sample presentation and positive and negative comparison locations were different. Thus, within both the sample-first and the comparison first conditions, any two consecutive sessions consisted of both types to guarantee neither type of session was conducted more than twice in a row.

Instruction. Training was conducted in the same fashion as baseline sessions, except differential reinforcement was provided for correct and incorrect responses. A correct response, defined as clicking the positive comparison, resulted in the presentation of a computer animation

and a sound clip for 4 s. The animation and sound clip used were specific to each trial in each session type, but were identical across both of the sample-first and the comparison-first conditions. An incorrect response, defined as clicking the negative comparison, and no response within 10 s of the end of the auditory sample resulted in the presentation of an all-black screen for 4 s. Once the black screen timed out, the presentation proceeded to an error correction trial. Initially, clicking outside of the comparison stimuli borders resulted in being recorded and receiving feedback as an incorrect response and thus the participant was advanced to the black screen for 4 s and then on to error correction. Matthew and Rory went through the experiment while this protocol was still in place. However for Danny and Peter, clicking outside of the parameters of the comparison stimuli resulted in no action, and thus the slide trial would not progress unless either a negative or positive comparison was selected or there was no response within 10 s.

Error Correction. The order of stimulus presentation in the error correction trials for the sample-first condition and the comparison-first condition was the same as the instruction trials respectively, except a sample visual stimulus of the positive comparison was added to the sequence. This error correction procedure followed all incorrect responses during instruction. For the sample-first condition, the black screen disappeared and was replaced with a white background. The auditory sample was then presented and immediately following its cessation (1670 ms) a sample visual stimulus identical to the positive comparison was presented at middle of the top third of the screen along followed by the same four comparison stimuli at the bottom of the screen from the instruction trial. Figure 1 shows an example of the structure of an error correction slide used during instruction trials. Clicking the positive comparison resulted in the presentation of a computer animation and a sound clip lasting 4 s. Clicking anywhere besides the

positive comparison made no change to the screen and the error correction trial remained. If the participant did not click after 5 s or clicked anywhere besides the positive comparison, the experimenter instructed them to "click the one that matches". For the comparison-first condition, the black screen disappeared and was replaced with a blank white screen. The four comparison stimuli were presented in the same position as the instruction trial, at the bottom third of the screen. The auditory sample was then presented, followed by the appearance of the sample visual stimulus in the middle of the top third of the screen. All consequences were the same as the sample-first condition described above. Initially, clicking anywhere besides the positive comparison resulted in immediately being advanced to the next trial. Matthew and Rory went through the experiment while this protocol was in place. However for Danny and Peter, clicking anything besides the positive comparison resulted in the procedures described above.

Mastery was defined as the participant responding correctly in at least 14 of the 16 trials (87.5% of trials) in three consecutive sessions. Training then stopped in that condition and continued in the unmastered condition until responding either met mastery criterion or there was no upward trend for at least four consecutive sessions.

Response Measurement and Reliability. The dependent measure in this study was the number or correct responses selected during baseline and instruction sessions. As mentioned above, a correct response was defined as clicking within a positive visual comparison's borders. An incorrect response was defined as clicking within a negative visual comparison's borders or not clicking any stimuli within 10 s of the auditory sample ending. For Matthew and Rory an incorrect response was also defined as clicking anywhere besides the comparison stimuli, however this definition changed before Danny and Peter began instruction and thus did not apply to them. Additional data were also collected during the comparison-first condition to record any

early responding, which was defined as clicking any of the comparison stimuli before the auditory sample was fully presented. A primary experimenter recorded these responses on a data sheet with a second observer independently scoring 54.2-100% of all sessions for each condition and for each participant either live or by watching a video recording. This interobserver agreement (IOA) was calculated on a trial-by-trial basis and agreement was defined as a trial that was scored identically by both observers for both accuracy of response selection and early responding. IOA was calculated by dividing the number of agreements by the total number of trials and multiplied by 100. The mean agreement for all sessions was 97.6% for Matthew (94-100%), 99.6% for Rory (94-100%), 98.8% for Danny (88-100%), and 97.8% for Peter (88-100%).

Results

Figures 2 and 3 shows the results of all four participants' data by representing the number of correct responses for each baseline and instruction trial. Trials conducted in the sample-first condition are represented by the filled circles. Trials conducted in the comparison-first condition are represented by the open circles. The percentage of early responding that occurred during comparison-first trials is depicted by the gray bars along the x-axis.

Matthew responded at or below chance levels in baseline sessions. For the bird stimuli, Matthew reached mastery in the comparison-first condition after 11 sessions, or 176 trials. He reached mastery in the sample-first condition after 14 sessions, or 224 trials. Early responding occurred in 22.2% of baseline sessions (2.08% of trials) and 9.1% of instruction sessions (0.6% of trials) for the comparison-first condition. Instruction for the flags stimuli never took place because of time constraints.

Rory responded at or below chance levels in baseline sessions. For the bird stimuli, Rory reached mastery in the sample-first condition after three sessions, or 48 trials. He reached mastery in the comparison-first condition after four sessions, or 64 trials. For the flags stimuli, Rory reach mastery in both conditions after five sessions, or 80 trials in each. Early responding occurred in 0% of all comparison-first trials for the both the birds and the flags stimuli.

Danny responded below chance levels in baseline sessions. For the flags stimuli, Danny reached mastery in the comparison-first condition after 17 sessions, or 272 trials. He did not reach mastery in the sample-first condition before criteria for terminating instruction were met. For the birds stimuli, Danny reached mastery in the sample-first condition after four sessions, or 64 trials. He also reached mastery in the comparison-first condition after six sessions, or 96 trials. Early responding occurred in 100% of the baseline sessions (20% of trials) and 20.8% of the instruction sessions (1.6% of trials) for the comparison-first condition.

Peter responded below chances levels for the sample-first sets of birds and flags stimuli. He responded high above chance levels (62.5-81.25%) in baseline sessions of the assigned Group 2 for birds stimuli and thus these had to be replaced with Group 3. He responded below chance levels for this new group in baseline sessions. Peter also responded high above chance levels (81.25%) in baseline sessions of the assigned Group 2 for flags stimuli and thus these also had to be replaced. The new group, Group 3, also had high responding in baseline sessions (93.75%) and a final group of flags stimuli was created. However Peter continued to high accuracy (91.67%) in responding for one of the comparison stimuli in this Group 4, the flag Iceland. The continued high responding in Group 4 flags resulted in it being replaced completely with new stimuli, Groups 1 and 2 state flowers. For the birds stimuli, Peter reached mastery in the sample-first condition in four sessions, or 64 trials. He reached mastery in the comparison-

first condition in five sessions, or 80 trials. For the flowers stimuli, he reached mastery in the sample-first condition in five sessions, or 80 trials. In the comparison-first condition, he reached mastery in eight sessions, or 128 trials. Early responding occurred in 28.6% of the baseline sessions (2.2% of trials) and 8.3% of instruction sessions (0.5% of trials) trials for the comparison-first condition.

Discussion

Out of the seven evaluations that were completed by all four participants, four showed faster mastery in the sample-first condition (though in one case, they differed by only one session), two of them showed faster mastery in the comparison-first condition, and equal number of trials was required in one. Of the three participants who completed two evaluations, only Peter showed a sample-first advantage in both evaluations. Although faster mastery in the sample-first condition was the most common scenario, the results were less clear than in Petursdottir and Aguilar (2016), where a sample-first advantage was seen for all participants and evaluations. For the two evaluations where there was faster mastery in the comparison-first condition, the discrepancy of the number of sessions and trials between the two conditions was larger than any of the evaluations where sample-first mastery was met first; in fact, Danny never mastered the sample-first condition for flags. Overall, the findings do not imply a consistent advantage to either condition.

There are at least three possible reasons why the results of the present study differed from those of Petursdottir and Aguilar (2016), which employed the same procedures but without the picture prompt error correction. First, the error correction procedure may have diminished the importance of stimulus presentation order. For example, it is possible that pairing of auditory and visual stimuli (i.e., the picture prompt) involved in the error correction procedure itself was a

major contributor to learning. If this is the case, then the order of stimulus presentation may have become less relevant, as auditory and visual stimuli were paired in the same manner during error correction in both conditions. Future research should investigate the amount of learning that happens within the error correction trials themselves.

Second, given that these were single-case design studies, individual differences (e.g., different instructional histories and personal characteristics) between participants in the two studies could account for the results. Future research should evaluate the two presentation orders with and without error correction within subjects. Third, error correction trials in the sample-first condition could be less effective than in the comparison-first condition. In the sample-first condition, the insertion of the picture prompt into error correction trials resulted in a slightly increased delay between the presentation of the sample and the comparisons compared to instructional trials. In the comparison-first condition, by contrast, the comparisons were present on the screen when the sample was presented in error correction trials as in instructional trials. It is possible that the added delay in the sample-first condition made the error correction procedure less effective in that condition and therefore, removed some of the advantage of sample-first presentation seen in the earlier study (Petursdottir & Aguilar, 2016).

Early responding was recorded but did not prove to be of concern, and this is shown by the minimal number of times it occurred within each session. The maximum number of early responses that occurred in a session was 5, and this only occurred in one session for one participant (Danny). Danny exhibited the highest rate of early responding, with at least one premature response occurring in all of his baseline sessions. However this may have been due to his unfamiliarity with the procedure and his history of making an observing response for a previous study in which he was involved. Due to the decline in his early responses during

instruction (from 20% of trials in baseline to 1.6% of trials in instruction) it is plausible to assume that this premature responding was not necessarily due to interferences in stimulus control. Early or impulsive responding in the comparison-first condition may be taken as a possible indicator that the comparison display alone is evoking selection behavior irrespective of the sample stimulus (McIlvane, Kledaras, Stoddard, & Dube, 1990). Such a stimulus control interference might be particularly likely to affect performance in the sample-first condition, where an impulsive response to the display alone will likely result in an error each time. In the present study, however, the evaluations in which mastery occurred first in the comparison-first condition were not associated with more early responding than the remainder of the evaluations.

Although the current study expanded on some of the recommendations and limitations from Petursdottir and Aguilar (2016), some issues remained. One of these was that the participants in this study were typically developing school-age children who showed average to advanced language skills, and thus the tasks were not immediately necessary or educationally relevant to them. This is also evident when examining the data of two of the participants, Rory and Peter, who seemed to be experiencing a ceiling effect during at least one of their evaluations, suggesting that the task was easily mastered for them. In accordance with recommendations in the past, research of the order of stimulus presentation for receptive language tasks should be extended to the populations that could benefit most from more efficient and effective training methods such as those with autism spectrum disorder and other developmental language delays (Petursdottir & Aguilar, 2016). While the data from this study suggests no discernable advantage of either presentation condition when an error correction picture prompt procedure is added, it is possible that different results would be obtained with children in need of language intervention.

A procedural restriction made by the experimenters was the discontinuation of a condition if it began to show no upward trend after the first condition was mastered. This procedure was chosen in order to ensure that time would permit conducting two evaluations with each participant. While this only occurred in Danny's case, it still prevented a full analysis of the discrepancy between mastery of the two conditions. An additional limitation was that Matthew was only able to complete one evaluation due to the school year ending before his second evaluation could commence.

In the current study, the advantage of presenting the sample first during receptive instruction trials was less clear than in Petursdottir and Aguilar (2016). Future research should also look to investigate these procedures with populations who are more representative of those who might receive receptive instruction in MTS format, such as those with ASD.

References

Barbetta, P. M., Heward, W. L., & Bradley, D. M. (1993). Relative effects of whole-word and phonetic-prompt error correction on the acquisition and maintenance of sight words by students with developmental disabilities. *Journal Of Applied Behavior Analysis*, *26*(1), 99-110. doi:10.1901/jaba.1993.26-99

Carp, C. L., Peterson, S. P., Arkel, A. J., Petursdottir, A. I., & Ingvarsson, E. T. (2012). A further evaluation of picture prompts during auditory-visual conditional discrimination training. *Journal of Applied Behavior Analysis, 45*(4), 737-751.

Carroll, R. A., Joachim, B. T., St. Peter, C. C., & Robinson, N. (2015). A comparison of error correction procedures on skill acquisition during discrete-trial instruction. *Journal of Applied Behavioral Analysis, 48*(2), 257-273.

Doughty, A. H., & Saunders, K. J. (2009). Decreasing errors in reading-related matching to sample using a delayed-sample procedure. *Journal of Applied Behavior Analysis, 42*, 717-721.

Fisher, W. W., Kodak, T., & Moore, J. W. (2007). Embedding an identity-matching task within a prompting hierarchy to facilitate acquisition of conditional discriminations in children with autism. *Journal of Applied Behavior Analysis, 40*(3), 489–499.

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

Grow, L., & LeBlanc, L. (2013). Teaching receptive language skills: Recommendations for instructors. *Behavior Analysis in Practice, 6*, 56-75.

Hart, B. & Risley, T. R. (1975). Incidental teaching of language in the preschool. *Journal of Applied Behavioral Analysis, 8*(4), 411-420.

Leaf, J. B., Alcalay, A., Leaf, J. A., Tsuji, K., Kassardjian, A., Dale, S., & ... Leaf, R. (2014). Comparison of most-to-least to error correction for teaching receptive labelling for two children diagnosed with autism. *Journal Of Research In Special Educational Needs*, doi:10.1111/1471-3802.12067

Leaf, R., & McEachin, J. (1999). *A work in progress. Behavior management strategies and a curriculum for intensive behavioral treatment of autism*. New York: DRL.

Leaf, J. B., Leaf, R., Taubman, M., McEachin, J., & Delmolino, L. (2014). Comparison of flexible prompt fading to error correction for children with autism spectrum disorder. *Journal Of Developmental And Physical Disabilities*, *26*(2), 203-224. doi:10.1007/s10882-013-9354-0

Kodak, T., Fuchtman, R., & Paden, A. (2012). A comparison of intraverbal training procedures for children with autism. *Journal Of Applied Behavior Analysis*, *45*(1), 155-160.

MacDuff, G. S., Krantz, P. J., & McClannahan, L. E. (2001). Prompts and prompt-fading strategies for people with autism. In Maurice, C., Grren, G., & Foxx, R. M. (Eds.), *Making a difference: Behavioral intervention for autism* (37-50). Austin, TX: PRO-ED, Inc.

McGhan, A. C. & Lerman, D. C. (2013). An assessment of error-correction procedures for learners with autism. *Journal for Applied Behavioral Analysis, 46*(3), 626-639.

McIlvane, W. J., Kledaras, J. B., Stoddard, L. T., & Dube, W. V. (1990). Delayed sample presentation in MTS: Some possible advantages for teaching individuals with developmental limitations. *Experimental Analysis of Human Behavior Bulletin, 8*, 31-33.

Petursdottir, A. I. & Aguilar, G. (2016). Order of stimulus presentation influences children's acquisition in receptive identification tasks. *Journal of Applied Behavioral Analysis*, *49*, 58-68 doi: 10.1002/jaba.264

Rodgers, T. A. & Iwata, B. A. (1991). An analysis of error-correction procedures during discrimination training. *Journal of Applied Behavioral Analysis, 24*(4), 775-781.

Sindelar, P.T., Rosenberg, M.S., & Wilson, R.J. (1985). An adapted alternating treatments design for instructional research. *Education and Treatment of Children, 8*, 67-76.

Taylor, B. A., & McDonough, K. A. (1996). Selecting teaching programs. In C. Maurice, G. Green, & S. Luce. (Eds.), *Behavioral intervention for young children with autism: A manual for parents and professionals* (pp. 63-177). Austin, TX: PRO-ED.

Table 1

Stimulus Groups

	Category		
	Birds	Flags	State Flowers
Group 1	Scrub Jay, Cassins Finch, Crested Caracara, Western Medowlark	Togo, Denmark, Yemen, Peru	Alabama, Illinois, Colorado, Maryland
Group 2	Bobolink, Bells Vireo, American Kestrel, Pigmy Nuthatch	Finland, Brazil, China, Mali	Alaska, Iowa, Vermont, Georgia
Group 3	Hepatic Tanager, Common Redpoll, Canyon Towhee, Chukar	Turkey, Sweden, Benin, Ukraine	-
Group 4	-	Iceland, Congo, Gabon, Jordan	-

Table 2

Assignment of Stimulus Groups to Participants and Conditions

	Sample-First			Comparison-First		
	Birds	Flags	Flowers	Birds	Flags	Flowers
Matthew	Group 1	-	-	Group 2	-	-
Rory	Group 2	Group 1	-	Group 1	Group 2	-
Danny	Group 1	Group 2	-	Group 2	Group 1	-
Peter	Group 1	-	Group 1	Group 3	-	Group 2

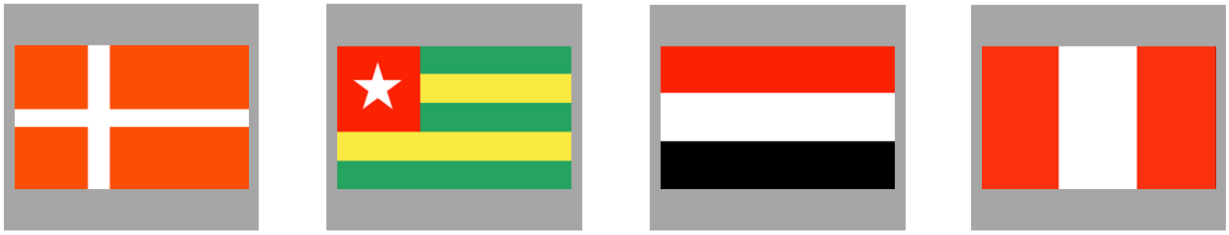

Figure 1. An example of the presentation organization for an error correction picture prompt procedure.

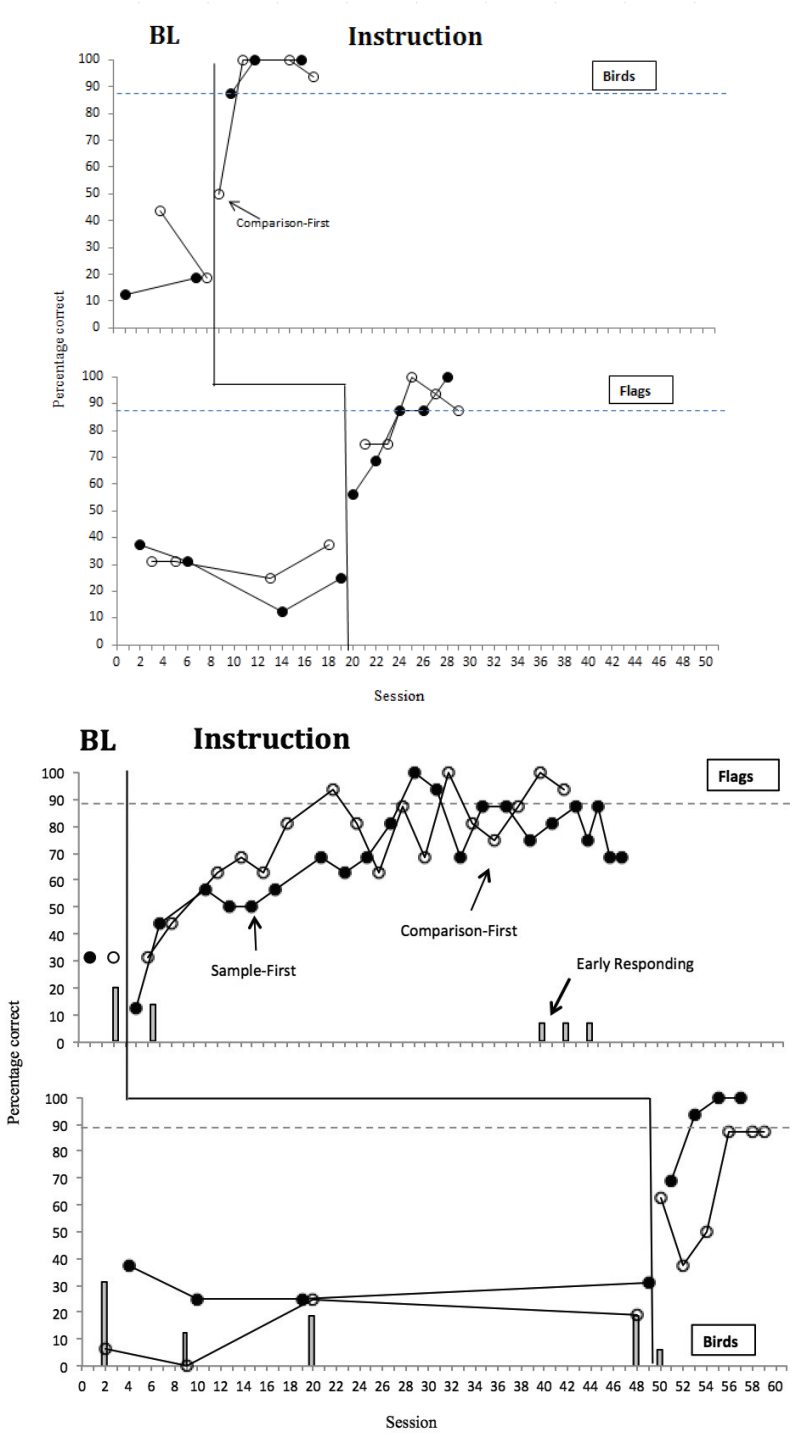

Figure 2. Response accuracy in the sample-first condition (closed circles) and the comparison-first condition (open circles) during baseline and instruction for Rory (top two panels) and Danny (bottom two panels). Gray bars represent percent of early responses within each session in the comparison-first condition. BL = Baseline. The dotted line represents mastery criterion percentage (87.5%).

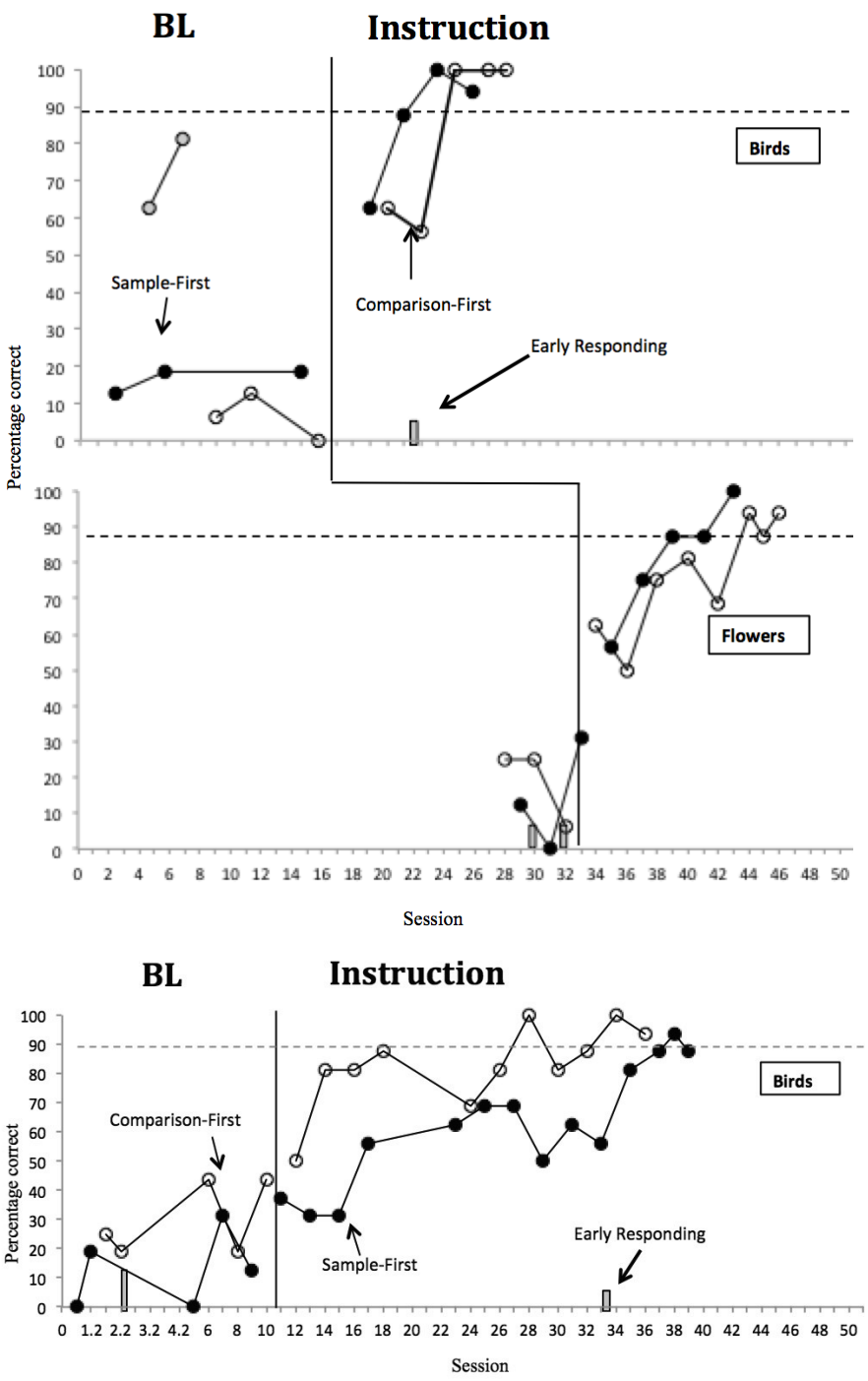

Figure 3. Response accuracy in the sample-first condition (closed circles) and the comparison-first condition (open circles) during baseline and instruction for Peter (top two panels) and Matthew (bottom panel). Gray bars represent percent of early responses within each session in the comparison-first condition. BL = Baseline. The dotted line represents mastery criterion percentage (87.5%).