FURTHER EVALUATION OF BLOCKED TRIALS TO TEACH INTRAVERBAL
CONDITIONAL DISCRIMINATIONS: EFFECTS

OF CRITERION-LEVEL PROBES

Jennifer Lynn Haggar, B.B.A.

Thesis Prepared for the Degree of

MASTER OF SCIENCE

UNIVERSITY OF NORTH TEXAS

May 2014

APPROVED:

Einar Ingvarsson, Committee Chair
Shahla Ala’i-Rosales, Committee Member
Karen Toussaint, Committee Member
Richard Smith, Chair of the Department of Behavior Analysis
Thomas Evenson, Dean of the College of Public Affairs and Community Service
Mark Wardell, Dean of the Toulouse Graduate School
Individuals with autism often have deficient intraverbal repertoires. Previous research has found success in using a blocked trials procedure to facilitate discrimination training. A previous study (unpublished) from our laboratory extended this procedure to intraverbal training. The current study continued this line of research by exploring the outcomes of probing the criterion performance more frequently. Three children with autism, ages 7-13, participated. Eight question pairs were taught. One question was presented repeatedly until a specified number of consecutive correct responses occurred, then the other question was presented. Contingent on specific mastery criteria, the trial blocks were faded into smaller blocks until the questions were presented in quasi-random order. Between each step, a criterion probe was conducted to determine if further steps were necessary. The procedure has been successful for two of the three participants. Criterion probe performance showed that not all teaching steps were needed every time. The procedure may have facilitated acquisition over time, because the number of trials to mastery generally decreased over successive targets.
Copyright 2014

by

Jennifer Lynn Haggar
ACKNOWLEDGEMENTS

First, I would like to thank Dr. Einar Ingvarsson for being my thesis advisor. I feel very privileged to have had you as my advisor. Your input and guidance along the way in the research and writing process is very valuable to me. I would also like to thank you for being a great supervisor for me at the Child Study Center. The skills I have gained from working as your intern are irreplaceable, and I greatly appreciate all of the guidance and support you have provided to me. To Duy Le, thank you for giving me the opportunity to earn my supervision hours at the Child Study Center. I have learned more in my year here than I could have ever imagined. I appreciate all of the opportunities that you have given me to work in various capacities at the center. These experiences have shaped me and will continue to shape me into the best behavior analyst that I can be. To Dr. Shahla Ala’i-Rosales and Dr. Karen Toussaint, thank you both very much for being on my thesis committee. Hearing your inputs and comments about my thesis and other behavior analysis components was very valuable to me. To the Child Study Center Research Lab, thank you all for all of your support along the way. I especially appreciate all of you who attended my practice (“mock”) thesis defense. That experience was extremely beneficial to me in preparation for the real defense. To Emily Ferris, thank you for sitting through hours and hours of videos collecting my IOA and TI data. I know that that could not have been the most fun thing to do, and appreciate you sticking with it to the end. And finally, I would like to thank my family and friends for their support throughout this process. Thank you for listening to me endlessly talk about my research and for being my ear and shoulder when the stress of it all would get to me. I appreciate each and every one of you and am so glad to have your support.
TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>ACKNOWLEDGMENTS</th>
<th>iii</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF TABLES</td>
<td>v</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>vi</td>
</tr>
<tr>
<td>Chapters</td>
<td></td>
</tr>
<tr>
<td>1. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>2. METHODS</td>
<td>12</td>
</tr>
<tr>
<td>Participants</td>
<td>12</td>
</tr>
<tr>
<td>Setting</td>
<td>13</td>
</tr>
<tr>
<td>Measurement</td>
<td>14</td>
</tr>
<tr>
<td>Experimental Design</td>
<td>15</td>
</tr>
<tr>
<td>Interobserver Agreement</td>
<td>15</td>
</tr>
<tr>
<td>Treatment Integrity</td>
<td>16</td>
</tr>
<tr>
<td>Procedures</td>
<td>16</td>
</tr>
<tr>
<td>3. RESULTS</td>
<td>26</td>
</tr>
<tr>
<td>4. DISCUSSION</td>
<td>32</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>60</td>
</tr>
</tbody>
</table>
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Question Pairs and Sets by Participant</td>
<td>41</td>
</tr>
</tbody>
</table>
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Flow of conditions</td>
<td>42</td>
</tr>
<tr>
<td>2</td>
<td>Random presentation pretest 1 Weldon</td>
<td>43</td>
</tr>
<tr>
<td>3</td>
<td>Baseline probe Weldon</td>
<td>43</td>
</tr>
<tr>
<td>4</td>
<td>Random presentation pretest 2 Weldon</td>
<td>44</td>
</tr>
<tr>
<td>5</td>
<td>Blocked trials Weldon</td>
<td>45</td>
</tr>
<tr>
<td>6</td>
<td>Trials to criterion Weldon</td>
<td>46</td>
</tr>
<tr>
<td>7</td>
<td>Random presentation pretest Michael</td>
<td>47</td>
</tr>
<tr>
<td>8</td>
<td>Baseline probe Michael</td>
<td>48</td>
</tr>
<tr>
<td>9</td>
<td>Blocked trials Michael</td>
<td>49</td>
</tr>
<tr>
<td>10</td>
<td>Trials to criterion Michael</td>
<td>50</td>
</tr>
<tr>
<td>11</td>
<td>Random presentation pretest Grayson</td>
<td>51</td>
</tr>
<tr>
<td>12</td>
<td>Baseline probe Grayson</td>
<td>52</td>
</tr>
<tr>
<td>13</td>
<td>Blocked trials Grayson</td>
<td>53</td>
</tr>
<tr>
<td>14</td>
<td>Trials to criterion Grayson</td>
<td>54</td>
</tr>
</tbody>
</table>
CHAPTER 1

INTRODUCTION

Conversations are a regularly occurring part of people's lives—from asking someone's name, to asking and giving directions, to answering people's questions, to discussing to how someone feels. These exchanges involve responding to someone else's verbal behavior. Skinner (1957) defined verbal operants included in these types of exchanges in terms of the antecedents and consequences that are functionally related to the behavior. These types of exchanges are very complex and often involve multiple verbal operants. The intraverbal is one of these verbal operants; it involves a verbal response that is evoked by a verbal antecedent, maintained by generalized reinforcement, and has no formal similarity or point to point correspondence with the verbal stimuli that evoked it (unlike echoic behavior, which is considered a special case; Skinner, 1957). A common deficit in the repertoires of many individuals with autism and other developmental disabilities is the failure to acquire an intraverbal repertoire that allows the individual to engage in normal conversations (Sundberg & Sundberg, 2011). It is important to teach the skills necessary to have conversations because of the importance that these skills have in our lives. Deficient intraverbal repertoires can greatly affect an individual's social repertoire.

Sundberg and Sundberg (2011) suggested that the deficient intraverbal repertoires of children with autism might be due to the failure of the verbal behavior of these individuals to come under conditional stimulus control of complex verbal antecedents. A conditional discrimination is a type of discrimination “…in which a response to a discriminative stimulus produces reinforcement only if another stimulus is present” (Slocum, Miller, & Tiger, 2012, p. 619). Tasks requiring these types of discriminations can be presented in various forms. They can be visual-visual discriminations as in identity matching-to-sample, or auditory-visual, as in
receptive labeling. They can also be in an auditory-auditory form, in the case of complex stimulus control of intraverbal behavior.

Eikeseth and Smith (2013) discussed compound and conditional discriminations in the context of intraverbal relations. The authors cited Sidman (2000) in describing a conditional stimulus: “the conditional stimulus determines the function of the antecedent stimulus in a discriminated operant” (Eikeseth & Smith, 2013, p. 126). Therefore, a conditional discrimination involves a discrimination in which the function of one stimulus will change in the presence of another stimulus. Whereas with a compound discrimination, the two (or more) stimuli or stimulus components would function together as a unit to evoke a different response than if one or the other stimuli were presented separately (Eikeseth & Smith, 2013).

When a conditional discrimination is involved, the listener has to attend to more than one stimulus in order to respond correctly. This skill can at times be difficult for an individual with autism to acquire, even when they are able to make simple discriminations. The failure to respond to all relevant stimuli or parts of stimuli has been referred to as stimulus overselectivity (Lovaas, Koegel, & Schreibman, 1979). This occurs when a limited or restricted range of stimuli or stimulus components controls responding, leading to deficient repertoires. This becomes a problem in acquiring more complex verbal behavior because speech “…is a complex stimulus input for which adequate responding necessitates the child’s attention to a number of stimulus dimensions” (Lovaas et al., 1979, p. 1241). An example of this would be in responding to the question “what animal has a long neck?” When responding to this question using all of the relevant stimulus components, an appropriate response would be a giraffe. A few examples of responses that may result from only attending to a limited number stimuli would be giving the answers short (from attending to the stimulus “long”), dog (from attending to the stimulus
“animal”), head (from attending to the stimulus “neck”), and many other nonvocal responses may occur (e.g., touching the neck).

When an individual cannot attend to all of the relevant stimuli in speech, he or she may not be able to respond in ways that are typically reinforced by his or her verbal community. For example, if a child with autism is with his typically developing peers and is asked “do you want play with us?” and the child with autism responds with “play with us.” This could result in the other children laughing or making fun of the child with autism. It could also result in those children not asking him to play with them again. Another example could be in a classroom when the teacher asks for everyone who is finished with their work to sit on one side of the room, and those who need to do work to sit on the other side of the room. These types of instructions will be very hard to follow if the child is not able to attend to multiple parts of the auditory stimulus. It is for reasons like these that it is important to teach children with autism how to respond appropriately to conditional stimuli.

Previous research has explored the use of an approach known as the “blocked trials procedure”1 to overcome failures to acquire conditional discriminations. These studies have almost exclusively focused on methods to teach visual-visual and auditory-visual conditional discriminations. Saunders and Spradlin (1989) conducted groundbreaking research on conditional discriminations with adults with intellectual disabilities. The participants had acquired generalized identity matching, but had failed to acquire visual-visual arbitrary matching. In the first experiment, the experimenters used sample-schedule training, which required the participant to press a button when the sample stimuli appeared on the screen. They

---

1 This procedure used to be called a blocking procedure, but the experiments chose to refer to it as the blocked trials procedure to avoid potential confusion with the blocking phenomenon (i.e., failure to establish stimulus control due to a previously established SD).
successfully taught successive discriminations of sample stimuli using the sample-schedule training, but failed to see acquisition of the conditional discrimination. In the next two-part experiment they examined if the participants would acquire the conditional discrimination after learning to discriminate between the comparison stimuli. They first exposed each participant to the comparison discrimination to train them in isolation before continuing with the sample schedule conditions from the first experiment, but this showed no effect on conditional discrimination acquisition.

In the second part of the two-part experiment, Saunders and Spradlin (1989) used alternating blocks of trials in which each block contained only one of the sample-schedule stimuli. At first, they taught only one of the two blocks of sample-schedule stimuli per session until 90% correct responding occurred. They then taught two blocks of trials per session using a fixed number of trials per block, one of each component of the conditional discrimination, for 16 trials each. Contingent on mastery criteria, they continued to decrease the number of trials per block until the two stimulus relations were targeted in randomly alternating trials. The experimenters moved back to the previous session criterion upon “failure to make progress at any point in the sequence” (Saunders & Spradlin, 1989, p. 6-7). They added a response cost contingency with one participant due to apparent variability in reinforcer effectiveness. Both participants acquired the conditional discriminations following the blocked trials presentation modification.

Saunders and Spradlin (1990) replicated and extended the previous study by examining if the amount of training necessary (i.e., trials to criterion) would decrease following the acquisition of conditional discriminations using the blocked trials. They used a combination of the blocked trials procedure, trial-and-error, and sample naming to teach the discriminations.
One participant acquired the conditional discriminations in the baseline trial-and-error condition after the training package of trial-and-error plus sample naming had been used to teach the first two discriminations. The blocked trials procedure plus sample naming were used for the second participant for the first target, and blocked trials only were used with the next two discriminations. The subsequent discriminations were all acquired in the baseline trial-and-error condition. These results suggest a learning to learn effect as successive targets were acquired in the baseline condition following the discrimination training.

Saunders and Spradlin (1993) extended their previous research by attempting to create an instructional program using the training components of their previous experiments, with a focus on teaching prerequisite skills. They started by establishing two component simple discriminations by reinforcing one comparison stimulus until errors were only made in the first two trials. This was followed by reversal training in which selection of the other comparison stimulus was reinforced. They then used sample discriminations to establish differential responding, which they called differential sample naming, to the sample stimuli. Finally, after the simple and sample discriminations were established, arbitrary matching trials were conducted using the blocked trials procedure and differential sample naming.

Saunders and Spradlin (1993) found that two of the three participants who were unable to acquire the discrimination with the trial-and-error training procedure were able to acquire the discrimination with the training procedure that first established the prerequisite component skills. The blocked trials procedure in this study involved a fixed number of trials per trial block before the switching to the other stimulus. The authors suggested for future research to use a number of consecutive correct trials to determine when to switch to the other stimuli, instead of the fixed
number of trials. They also suggested to decrease the size of the blocks in one or two trial
increments, instead of the larger, abrupt changes in the block size that were used in their study.

Smeets and Striefel (1994) tested the efficacy of the blocked trials procedure to teach
arbitrary conditional discriminations to typically developing 5-year-old children. They compared
different procedures involving a simple-discrimination (SD) and conditional responding (CR)
task. They required the participant to either point to the comparison or place a sample on the
comparison. They used a five step procedure in which the locations of the comparisons were
fixed in the first step. Over successive steps, the location of the comparison became more
variable, until it was randomly changed over trials. They evaluated the effects of different
teaching methods by using a combination of pointing or placing with the SD or CR
discrimination tasks. They found that the blocked trials method involving the CR-placing was
superior to the SD-placing and CR-pointing, which were all superior to the SD-pointing.

Perez-Gonzalez and Williams (2002) extended this research by testing the efficacy of
what they named a combined blocked trials procedure. This procedure was a combination of the
original blocked trials procedure (Saunders & Spradlin, 1989, 1990, 1993) and the revised
blocked trials procedure (Smeets & Striefel, 1994). The authors first targeted a visual-visual
matching-to-sample task, and then taught an auditory-visual object discrimination to five
children with autism and mental retardation ranging in age from 5 to 9-years-old. They were able
to teach the conditional discrimination to all five participants using the combined blocking
procedure. Further, they found that “mastering the first conditional discrimination with the
combined [blocked trials] procedure made it possible for the children to learn new conditional
discriminations under standard trial-and-error conditions” (Perez-Gonzalez & Williams, 2002, p.
299).
The combined blocked trials procedure used in Perez-Gonzalez and Williams, 2002, consisted of a six step procedure to teach the visual-visual discrimination. Step 1 started by presenting the same comparison stimulus in the same location until 10 consecutive correct responses occurred; this constituted a “block” of trials. After 10 consecutive correct responses, the other comparison location was presented until 10 consecutive correct responses occurred. This procedure was repeated until the participant responded with 10 consecutive correct responses in four consecutive blocks, allowing for the first trial after the switch to be incorrect. Step 2 followed the same procedures, except that five consecutive correct responses were required. Errors were not allowed to pass this step (i.e., four consecutive switches with no errors occurred before moving to the next step).

In Step 3, the trials were presented in blocks of two or three. Criterion to move to the next step was two consecutive 10-trial sessions with one or no errors. Step 4 included random sample presentation with the comparison stimuli in the same location. Criterion to move to the next step was identical to Step 3. Step 5 included random sample presentation and reversed comparison locations, with one session with 10 correct responses as the criterion to advance steps. The final step involved random sample stimulus presentation and random comparison locations. The final mastery criterion was two consecutive 10-trial sessions with no errors, or three consecutive 10-trial sessions with no more than one error. The auditory-visual object discriminations were taught using an identical sequence of steps, except Step 5 was not included. In the auditory-visual discrimination, the sample stimuli were the names of the objects.

Williams, Perez-Gonzalez, and Queiroz (2005) continued to extend the research on the effects of using the combined blocking trials procedure to teach color discriminations to a 14-year-old boy with autism. The procedures replicated those of Perez-Gonzalez and Williams
(2002) by systematically fading the block sizes while keeping the cards in a fixed location until the last step. The results indicated success in teaching a color discrimination with the blocking procedure. They suggested future research to look into necessity of each step in effort to develop a more efficient procedure (Williams et al., p.558).

Slocum, Miller, and Tiger (2012) extended these studies by teaching identity matching to a 25-month-old boy diagnosed with autism and examining “the necessity of fading the block size after initial acquisition of the conditional discrimination” (Slocum et al., 2012, p. 620). They evaluated this by presenting two samples in alternating fixed-size 10-trial blocks. They continued to alternate the 10-trial blocks until six consecutive sessions occurred with at least 90% of responses correct. After criterion was met, they returned to a mixed trial procedure with differential reinforcement in which both sample stimuli were presented in each session. They found that fading the number of blocked trials was unnecessary for the participant because he was able to respond correctly in the mixed-trials condition without first experiencing the gradual fading of the number of trials per block. The authors suggested that future researchers could conduct brief probes with random stimulus presentation before fading the size of the blocks to determine the necessity of fading.

Up to this point in the research line, the blocked trials procedure had only been applied to visual-visual and auditory-visual conditional discriminations. As discussed earlier, the auditory-auditory discrimination which is found in an intraverbal is another type of discrimination that can be difficult to teach. Axe (2008) wrote a review with recommendations for teaching intraverbal relations. He described methods of transfer of stimulus control and errorless teaching methods that have been successful in teaching intraverbal relations. However, there are also many studies
that have not been successful in making substantial improvements in teaching the more complex stimulus control over intraverbal responses.

One study has been successful in teaching more complex intraverbal relations to children with autism. Braam and Poling (1983) used a transfer of stimulus control procedure to teach simple and complex intraverbal relations to three individuals with developmental disabilities. In the first two experiments, they taught two participants to respond correctly to questions involving categories (e.g., name a food). In the third experiment, they taught two participants with hearing impairments and mental retardation (a 17-year-old girl and 23-year-old man) to make conditional discriminations using transfer of stimulus control procedures. They used a word flashcard to prompt the response to the question if a correct response did not occur within 10 s. The experimenter instructed the participant to name a thing from home, for example. The stimulus sets included stimuli that overlapped between the sets (e.g., home things, school things, home people, school people). The overlapped stimuli are important because it ensures the control by both components of the antecedent (e.g., home vs. school; things vs. people) to make a correct response. The transfer of stimulus control procedure was successful in teaching the intraverbal relations to both participants. At the conclusion of the study, the participants were able to respond correctly to all intraverbal relations. They also saw emergence of untrained correct responses with both participants.

Sundberg and Sundberg (2011) examined the intraverbal and verbal conditional discrimination repertoires of typically developing children and children with autism. They created an 80-question intraverbal assessment, and each participant was given a score based on the number of correct responses. They found that “there was a general correlation between the age of the child and the number of correct intraverbal responses” (Sundberg & Sundberg, 2011,
They found that errors due to lack of conditional stimulus control were common in younger typically developing children, but these children generally overcame these difficulties starting around the age of 4. However, some children with autism did not overcome these difficulties in responding to questions involving conditional stimulus control. Nevertheless, they tended to make the same types of errors as younger typically developing children who received the same score. Due to the importance of intraverbal behavior to academic and social repertoires, this study gives us a reminder of how important it is to examine different procedures for teaching intraverbal responses to children with autism to determine the most effective procedures.

Ingvarsson, Macias, Koelker, Carp, and Petursdottir (2014) used the previous research on teaching conditional discriminations using the blocked trials procedures, and applied it to teaching an intraverbal relation. Because this method had shown success with other types of conditional relations, this application of the procedure was a logical step into a new line of research to teach an auditory-auditory conditional discrimination. The authors examined the effects of using the blocked trials procedure to teach conditional intraverbal discriminations deriving methods from Perez-Gonzales and Williams (2002).

The modified procedures in Ingvarsson et al. (2014) contain five total steps; however, teaching began at the third step and reverted to previous steps only when necessary. They designed the training this way in attempts to reduce the amount of trials to teach the discriminations. They taught answers to eight question pairs requiring a conditional discrimination using a multiple probe design across question pairs. The procedure was eventually successful for all participants (four children with autism). However three participants required some procedural modifications. For one participant, the final two discriminations were acquired during the baseline condition (differential reinforcement and error correction). They
recommended future research to analyze the error patterns to determine the best procedural modification (e.g., adding in distracter trials due to incorrect responses to the first trial in each block).

The current study replicated and extended Ingvarsson et al. (2014) by adding in probes to test for mastery after the completion of each step. If mastery is reached before going through all of the steps, less time would be spent in training. This could help in avoiding potential overtraining that could in effect cause the participants to respond too rigidly. By discontinuing the training at the appropriate point and reducing the potential rigidity in responding, the discriminations might be more likely to be acquired in baseline probes. The study will also test for retention by conducting one month maintenance probes following the end of training and baseline probes.

The following were the experimental questions: Will the modified blocked trials method be effective in teaching intraverbal conditional discriminations? Are all of the steps and trials necessary in teaching the discriminations using the blocked trials? Will adding in distractor trials be an effective way to teach discriminations when the child shows little progress? Will the discriminations be more likely to be acquired in baseline probes if fewer training trials are implemented?
CHAPTER 2

METHODS

Participants

Three children between the ages of 7 and 13 years participated in the study. All the participants were diagnosed with autism spectrum disorder and were enrolled in a treatment center for children with developmental disabilities. Each participant had mastered many tact and listener skills programs in the center’s curriculum. Each was also able to answer simple questions (e.g., “what’s your name?”). Despite the participants’ extensive tact and listener repertoires, they continued to have difficulty in answering more complex questions requiring conditional discriminations. The participants chosen for this study exhibited restricted stimulus control in answering the types of questions. For example, they gave the same response of “food” or “fork” (or something similar) to the questions “what do you eat?” and “what do you eat with?”

Weldon was an 8-year-old Caucasian boy who was enrolled in the autism services program at the center, but received therapy in one of the school classrooms. He had been enrolled in the center for almost three years. He had mastered many of the listener and speaker skills programs and was able to expressively and receptively identify over 200 items. He had also mastered many intraverbal programs involving simple discriminations, but he had not been able to master more complex intraverbal targets. Weldon was able to use 3-4 word sentences to ask for items and engage in simple intraverbal exchanges.

Michael was a 13-year-old Hispanic boy who was enrolled in the center’s autism services program. He had been receiving treatment from the center for 9 months. He had mastered many of the listener and speaker skills programs and was able to expressively and receptively identify over 200 items. He was able to read simple words. Although Michael had been able to a few
intraverbal programs involving simple discriminations, he had not been able to master more complex intraverbal targets. Michael was working on using phrases and sentences when manding.

Grayson was a 7-year-old African American boy who was enrolled in the center’s school for children with developmental disabilities. He had been enrolled in the school for six months. Prior to his enrollment in the school, he received therapy in the center’s autism services program for one year. He had mastered many of the listener and speaker skills programs and is able to expressively and receptively identify over 200 items. He was also able to follow two- and three-step instructions. Although Grayson had been able to a few intraverbal programs involving simple discriminations, he had not been able to master more complex intraverbal targets. Grayson also had a strong textual repertoire and was able to read many words. Grayson was able to use 5-8 word sentences to ask for items and have short, simple intraverbal exchanges.

**Setting**

Sessions occurred in an individual treatment room at the center in which the participants were enrolled. The individual treatment room contained a child-sized table with two chairs, a token board, preferred items for which the participants could trade their token boards (e.g., iPad), a clipboard, a timer, and a video camera set up on a second (larger) table.

The participant and experimenter sat across from one another at the smaller table in the treatment room. The experimenter kept the clipboard angled away from the participant at all times, because two of the participants had fairly developed textual repertoires and might otherwise have been able to read the correct answers. The participants’ reinforcer consumption time occurred in the treatment room as well.
Measurement

The data were collected using a data sheet that was created specifically for each session. The order of the questions asked each session was predetermined and specified on the data sheet (with the exception of the data sheets used during steps 1-3, see below, all of which contained the same question on each data sheet). The experimenter asked the questions in the predetermined order, and switched between data sheets according to the criterion for that step when applicable (i.e., Steps 1-3). The “target answer” was provided on the data sheet for each question as the response to be prompted in event of an error.

Across all conditions, data were collected for correct or incorrect responses. Because the targets were open-ended questions, there were multiple responses that could qualify as correct answer. A correct response was defined as any response that was written in the “target answer” box on the data sheet. In addition, responses that were not written in the “target answer” box, but still answered the question correctly were counted as correct responses. When these types of responses occurred, the response was written down and later discussed to determine whether to change that response to the target answer. For example, Michael answered the question “what do you wash?” with “a dish” instead of the target response written, “hands.” This response was reinforced, written down, and discussed after the session. The target answer for Michael was changed to “a dish” from that session on.

The participants were not required to answer the question using a sentence frame. Responses that contained the correct answer, but did not have the correct sentence frame were counted correct. For example if the participant’s response was “with a ball” for the question “what do you kick,” the response was scored as correct, but the correct sentence structure was also modeled after praise was delivered. Self-corrections were counted correct if the participant
corrected himself before saying the entire incorrect word. For example, if the participant responded “a ba-, a leg” for the question “what do you kick with,” the response would be scored as a correct response.

An incorrect response was a response that was not written in the “target answer” box and/or was not another acceptable answer to the question. Responses that were not initiated within 5 s were counted incorrect. Self-corrections were scored as an incorrect response if the participant corrected himself after saying the entire word. For example, if the participant responded “a ball, a leg” for the question “what do you kick with,” the response would be scored as an incorrect response.

Experimental Design

We used a multiple probe across question pairs design to evaluate the effects of training on intraverbal responses (Horner and Baer, 1978). The experimenter taught each participant two sets of four question pairs (for a total of eight question pairs, or 16 questions total). The baseline probes were conducted at the beginning of the experiment for each participant. Baseline probes were also conducted following mastery criterion for each question pair. The experimenter introduced a question pair to the blocked trials procedure only after the probe data from the previously mastered question pairs were stable. The design also contained criterion probes to check for mastery before completing all steps of the intervention.

Interobserver Agreement

A second observer collected data independently during at least 30% of all sessions in all conditions. IOA was calculated by comparing the data of the experimenter and the second observer trial-by-trial. The calculations were made by dividing the number of agreements (i.e., the data from both observers was identical for a given trial) by the total number of trials in a
given session. That number was then multiplied by 100 to obtain the percentage of agreement. IOA for Weldon was 99.6% (range: 97.8-100%), for Michael was 100%, and for Grayson 99.6% (range: 97.1-100%).

Treatment Integrity

The second observer also scored treatment integrity (TI) during at least 30% of sessions in each condition using a checklist listing the procedures that the experimenter should follow in each condition. The observer marked either a check for a completed step or an X for a step that was implemented incorrectly. TI was calculated by dividing the number of correct steps by the total number of steps and multiplying the resulting number by 100. TI for Weldon was 99.7% (range: 98.4-100%, for Michael was 100%, and for Grayson 98.1% (range: 92.1-100%).

Procedures

Preference Assessment

The experimenters implemented a multiple stimulus without replacement (MSWO; DeLeon & Iwata, 1996) to determine highly preferred items to be used as backup reinforcers for token boards for two of the participants. The experimenters did not use an MSWO with Michael due to reports of a clear preference for one particular item during his treatment at the center. In the first trial of each MSWO session, six stimuli were placed in front of the participant, and he was asked to choose one of the stimuli. After the stimulus was chosen, the participant had 30 s with which to play or consume (if an edible is used) the stimulus. Subsequently, the chosen stimulus was removed from the array, and the remaining stimuli placed in front of the participant. The position of the stimuli was rotated from trial to trial. This continued until there were no stimuli left in the array, or if the participant did not choose a stimulus within 30 s. The
MSWO was conducted six times in consecutive sessions. The top ranked stimuli from all six sessions combined were identified as the high preference items.

**Question Sets**

The experimenter determined the question pairs during pretests (see below). One question pair consisted of two questions: “what do you ____” and “what do you ____ with.” The experimenter selected a total of eight question pairs (16 questions) for each participant. These eight pairs were broken into two sets of four pairs for the baseline probes. The question pairs and sets for each participant are displayed in Table 1.

**Pretests**

The experimenter asked a series of questions targeting different types of discriminations. For example, “what day comes before Thursday?” and “what day comes after Thursday?” or “when does the sun rise?” and “when does the sun set?” The question pairs selected for the study included pairs in which the participant gave the same response for both questions. For example, when asked “what do you kick?” and “what do you kick with?” the participant might respond with “a leg.” If participants did not show this error pattern for a sufficient number of question pairs, they were not eligible to participate in this study.

**Random Presentation Pretests (RP pretest)**

Figure 1 displays the flow between the different conditions (explained below) that may be used to navigate through the procedures. To determine if the participant could acquire the discriminations with the traditional, less-intensive method of teaching using quasi-random question presentation, differential reinforcement, and prompting, the experimenter first conducted random presentation pretests with a randomly chosen question pair. The experimenter asked each question eight times; therefore each RP pretest session contained 16 questions. The
experimenter asked the questions in a quasi-random order determined prior to the session. The questions were arranged randomly with the exception that each question could not be asked more than three times consecutively.

Contingent on correct answers, the experimenter delivered descriptive praise on a continuous reinforcement (CRF) schedule and tokens on a variable ratio (VR) schedule individualized per participant. A VR 3 schedule was used for Weldon, a VR 4 schedule for Michael, and a VR 2.5 for Grayson. This reinforcement schedule was used across all conditions. Weldon and Michael were always required to earn ten tokens. Grayson first earned ten tokens for all conditions, but later earned three tokens in the teaching conditions (i.e., blocked trials and remedial conditions). When the participants responded incorrectly, the experimenter prompted the correct response and moved to the next question. When prompting, the experimenter stated the correct response following the participant’s incorrect response, and required the participant to repeat the response. Following the prompted response, the experimenter said “that’s right” and then repeated the sentence (e.g., “you kick with a leg”) in a neutral tone. The experimenter waited 3 s before moving on to the next question.

The experimenter conducted at least six RP pretest sessions of 16 trials each before determining if the participant would continue in the study. More than six RP pretest sessions were conducted if the errors were on a decreasing trend after the six sessions. If the participant still made more than three errors during the RP pretests and did not have a decreasing trend of errors, we concluded that learning was not occurring through the random presentation method, and use of the blocked trial procedure was therefore indicated. The participant therefore continued in the study. If participants had learned the discrimination during the RP probe, they would not have continued in the study; however, this did not occur.
Baseline Probes

Following the RP pretests, the experimenter began the baseline probes. The experimenter asked the participants all questions from the two sets. The baseline probe session contained two data sheets with one set of question pairs per data sheet. For each set, each question pair was presented two times in a quasi-random order resulting in a session consisting of 16 trials per set, or 32 trials per baseline session. The experimenter asked questions in a quasi-random order, with the exception that the same question could not be asked three times consecutively. The order was determined prior to session. The consequences for correct responses were identical to the RP pretests. When the participant responded incorrectly, the experimenter provided no consequence and waited 3 s before moving to the next question.

To pass the baseline probe and continue to the next question set, the participant had to respond correctly to all four questions per mastered pair in the baseline probe. If the participant responded incorrectly to any of the questions in a mastered question pair, the baseline probes were either conducted again, or remedial training was used for the question pair. One baseline probe session with an incorrect response on a mastered question pair was allowed before implementing the remedial training. However, if this occurred a second time, remedial training was implemented. An exception was made if several errors occurred in the first baseline probe, and fewer errors occurred in the second probe (i.e., probe date were on an upward trend). In this case, baseline probe sessions continued until an upward trend was no longer observed. In some cases, it was necessary to make a decision whether it was necessary to accept more variability in the baseline probes than originally planned. If the participant had shown to be able to answer 100% correct in the past, remedial procedures were not always used at the point described above.
(i.e., greater variability was accepted). These decisions were made on an individual basis by participant and by question pair.

**Blocked Trials Procedure**

In this procedure two questions were targeted in separate blocks of trials at first and then gradually interspersed contingent on correct responding until quasi-random rotation was achieved. The first step started with a block consisting of repeated trials of one question until a criterion of a certain number of consecutive corrects was met (the block criterion), at which point an identical trial block consisting of the other question was presented. The other question was then asked until the same criterion of consecutive corrects was met, at which point the first question was presented again in an identical manner. This process was repeated until a mastery criterion of consecutive blocks with specific numbers of errors allowed per block was met (the step criterion; e.g., 0 errors in four consecutive blocks). Once a step was mastered, the block criterion was modified. The modifications vary per step, as described below. The procedure used in the current study contains a total of five steps. The steps were identical to those of Ingvarsson et al. (2014), which were derived from Perez-Gonzales and Williams (2002). As in Ingvarsson et al., we introduced the block trials procedure starting at Step 3. Steps 1 and 2 were implemented only if Step 3 was initially unsuccessful.

**Step 1**

The experimenter asked one question from the first question pair (e.g., “what do you eat”) until the participant responded with ten consecutive correct answers. This constituted one block of questions. When the participant answered ten correct questions consecutively, the experimenter switched to the other question (e.g., “what do you eat with”) until the participant met the same criterion. The experimenter continued to present blocks of questions until the four
consecutive blocks with two or fewer errors per block occurred. When this criterion was met, the experimenter moved to Step 2.

**Step 2**

The experimenter asked the same two questions in the same manner as the previous step; however, criterion to switch questions was eight consecutive correct questions. The criterion to move to the next step was four consecutive blocks with one or fewer error per block.

**Step 3**

The experimenter asked the same two questions in the same manner as the previous step; however, criterion to switch questions was five consecutive correct questions. The criterion to move to the next step was four consecutive trial blocks with no errors.

**Step 4**

The experimenter alternated between the two questions in predetermined blocks of two or three questions. The criterion to move to the next step was 15 consecutive correct questions.

**Step 5**

The two questions were asked in the quasi-random order in the same manner as in the RP pretests. The criterion for mastery was 15 consecutive correct questions.

Procedural modifications were needed for Grayson. For Question Pair 1, after going back to Step 1 for Grayson and not seeing improvement, three distracter trials were added in between the trial blocks (Ingvarsson et al., 2014). These were added due to the error pattern of answering the first question after the switch with the previously reinforced response. The distracter trials consisted of mastered targets and were designed to control for Grayson repeating the last reinforced answer. The distracter trials were also used after reinforcer time before returning to the blocked trials.
A differential observing response (Kisamore, Karsten, Mann, and Conde, 2013) was also implemented with Grayson due to lack of acquisition. The observing response consisted of requiring Grayson to repeat the question before providing the answer to the question. The experimenter gave Grayson the vocal instruction to repeat the question and then to give the answer prior to presenting the first question each session. The experimenter prompted Grayson to emit the observing response first by pointing to him and repeating part of the question until Grayson repeated the question. The vocal reminder was then faded, leaving only the pointing, and eventually the experimenter provided no prompt to repeat the question.

Modifications were also used for Question Pair 2 for Grayson. The blocked trials procedures started at Step 3 with the observing response, then moved back to Step 1. The distracter trials were removed before starting Question Pair 2 because Grayson made several errors in the middle of the blocks, as well as on the first question, suggesting that the distracter trials were not effective or indicated in this particular case.

Random Presentation (RP) Probes

These probes were implemented following Steps 3 and 4, and following Steps 1 and 2 if those steps were needed, during the blocked trials intervention. The questions in the RP probes were arranged in the same manner as in the RP pretests. The questions were asked in a quasi-random order with the exception that the same question could not be asked more than three times consecutively. The experimenter delivered praise and a token on a VR schedule for correct responses. The experimenter did not prompt the correct response when an error occurred. Following incorrect responses, the experimenter delivered no consequence and waited 3 s before asking the next question. These probes determined whether baseline probes or the next step in the blocked trials procedure would be implemented next. If the participant made one or no errors
in the RP probe, baseline probes were conducted. If the participant made two or more errors, the next step in the blocked trials procedures was implemented.

**Remedial Training Procedures**

If a previously mastered pair did not pass the baseline probe, remedial procedures were used to reteach the discriminations. Remedial procedures were identical to Step 5 of the blocked trials procedure. If the participant required procedural modifications on the latest pair taught in the blocked trials, those teaching conditions were used for the pair going through remedial training. After passing the remedial training step, baseline probes were repeated. Using the same criterion, the experimenter either moved on to the next pair in the blocked trials procedure, continued with remedial training, or conducted another baseline probe.

For all participants, it was necessary to implement remedial training consisting of more than one question pair. This permutation of remedial training was implemented when a participant made errors across multiple question pairs. The combination training had the same question order requirements as Step 5 of the blocked-trials procedure. The criterion to pass this remedial combination step was 12 consecutive corrects (instead of 15 as in the programmed Step 5).

**Procedural Modifications for Grayson**

Grayson’s reinforcement schedule was modified during Step 1 for Question Pair 1 from earning 10 tokens to receive the backup reinforce, to three tokens for the blocked trails procedure. This was because our observations suggested that Grayson was not highly motivated when he was required to earn 10 tokens to access reinforcement. Grayson would look around the room, pick at his hands, play with his shirt, and mumble at times when the experimenter tried to get him to look at her. At times, she would ask if he was ready and he would not respond. The
experimenter noticed that when it was time to access reinforcement, he perked up and was extremely attentive to her as they played on the iPad together, responding to everything she said and making frequent eye contact. The reinforcement schedule was changed to allow him to earn the iPad more frequently. Subsequently, the amount of time he was given on the iPad was decreased to 1 min when this reinforcement schedule was in effect. We also added an “easy” token board (i.e., answering simple questions mastered in his program at the center) before presenting the target questions and free access to the iPad for a minute prior to session.

Grayson required more remedial training for Pair 1. The remedial training data for Step 5 were analyzed in the same manner as the RP pretests and probes. The data were graphed in blocks of 15 trials. If errors were still occurring after at least six blocks, then we moved the remedial training back to Step 3. Textual prompts were added to Step 3 in second remedial training procedures for Question Pair 1. The textual prompt was a 6.5 x 4 inch piece of white paper with the target question printed in black in 36 point Calibri font. The latter part of the question was typed in bold (e.g., “eat” and “eat with”). The experimenter held the card up in front of Grayson and at his eye level when asking the question. The experimenter waited for Grayson to look at the card before she asked the question. After proceeding through Steps 3 and 4 with the textual prompt, the textual prompt was faded during step 5. The experimenter faded the textual prompt by holding the card up for 1 s when asking the question, then putting the card down before finishing the question. The experimenter was able to say “what do you” in the 1 s while holding the card up. After five consecutive correct responses, the experimenter no longer used the cards. If an error was made, the experimenter used the cards again until 5 consecutive correct responses occurred.
Listener and Tact Probes

Listener and tact probes were conducted with Weldon and Michael after they had completed the study. The materials included environmental objects and pictures. The probes targeted the responses to all of the question pairs in each of the participant’s two sets. For example, for Weldon, two of the stimuli in the probe were a ball and a leg, which were responses to the “kick” question set. For the listener probes, the experimenter either placed a few pictures or objects in front of the participant or used objects in the environment and asked the participant to point to the target stimuli. The tact probes used the same materials and environmental objects. The experimenter pointed to one of the objects and asked “what is it?” If the participant named the target stimuli before the prompt, the experimenter used this response and did not ask the question. This occurred with Weldon. These probes were used to determine if the participant was able to either label the target responses or point to the target response when asked. These probes have not been conducted with Grayson. This is because following this study, we will continue to attempt to teach the remaining discriminations. Once he has acquired all responses, we will conduct these probes with Grayson.
RESULTS

Figure 2 shows the results from Weldon’s first RP pretests. After 135 trials, Weldon continued to make errors, making 5 errors each in the final two 15 question blocks. This showed that teaching it using a randomly alternating order, differential reinforcement, and prompts for incorrect responses was not an effective way to teach the discriminations for Weldon.

Figure 3 shows the baseline probes for Weldon. Each question pair was targeted four times in the probe; therefore, the score of a four in a baseline probe means that the participant answered all questions correctly for that particular question pair. A score of 2 usually indicates that the participant gave the same answer to all four questions. Question Pair 1 (bottom of Figure 3) was included in the first RP pretest. After six baseline probes, Weldon was answering three or four of the four questions correctly for Question Pair 1. Due to the greater than chance level responding for this question pair in the baseline probes, we repeated the RP pretest with another question pair. Figure 4 shows the next round of RP pretests, after 90 RP pretest trials (six 15 question sessions), errors remained high (nine errors in the final 15 question session).

Baseline probes continued after the failure to acquire the discrimination with the second question pair in the RP pretest. Weldon responded correctly to all four questions from the first question pair in the next baseline probe. However, when five baseline probes had been conducted, he only answered two of the four questions from the second question pair correctly. At this point, with the data for the second question pair stable, the blocked trials procedure was introduced with Question Pair 2.

Figure 5 shows the results of the blocked trials teaching procedure for Weldon for each pair taught. The figure shows the number of prompted responses (or number of errors) per blocked session. In Step 3, the mastery criterion was four consecutive blocks with no errors,
which is represented by the last four data points in that step. The mastery criterion for Steps 4 and 5 did not specify the number of errors, but rather the number of consecutive correct responses. The total number of errors for Step 4 and Step 5 are represented on the graph. Graphs that do not show a Step 4 or Step 5 indicate that the pair was mastered in the RP probe prior to that step.

The question pairs were taught sequentially, starting with Pair 2 and ending with Pair 8. For Pair 2, Weldon passed Step 3 in 37 trial blocks, and made six and eight errors total in Steps 4 and Step 5, respectively. For five of the remaining six pairs, Weldon passed Step 3 in fewer than 12 trial blocks, with only 5 trial blocks needed to master Pair 4. Weldon passed Step 3 for Pair 7 in 27 trial blocks, which is 10 fewer trial blocks than the first pair that was taught using the blocked trials procedure (Pair 2). Weldon’s errors during Step 4 ranged from 1-24. Subsequent to Question Pair 2, Step 5 was only necessary for two additional question pairs. Weldon passed Step 5 with 16 errors for Question Pair 4 and 0 errors for Question Pair 6. The blocked trials procedure was not used for question pair 1.

An RP probe was conducted following mastery of each step of the blocked trials procedure. The question pair was mastered if one or no errors were made during the probe. The next scheduled teaching step was conducted if the participant made two or more errors. Weldon responded with one to five errors in the RP probes following Step 3 and zero to three errors in the RP probes following Step 4.

Figure 6 shows the trials to criterion and number of steps necessary for each pair. Weldon acquired the discrimination for Question Pair 2, the first question pair taught using the blocked trials procedure, in 322 trials. The trials to criterion for all but one of the remaining question pairs decreased by 50-80%, with the highest number of trials being 198 and the lowest 62 trials.
Three of the eight question pairs required going through all three steps, three pairs were mastered in the probe after two steps, and one pair was mastered following the first step in the teaching procedure. The eighth question pair was acquired during baseline probes. However, this was the question pair that was initially used in the random presentation pretest. In the last baseline probe, Weldon answered all questions correctly.

Remedial training procedures were needed seven times across five different pairs throughout the baseline probes for Weldon. The remedial procedures were used in between the baseline probes when responding fell below a certain point (details above in methods section). A total of 353 trials were needed during the seven remedial training procedures. Remedial training trials to criterion ranged from 15 to 205 trials, with an average of 50 trials.

A one-month maintenance probe was conducted a month after the last baseline probe. In the first one-month probe, Weldon answered four of four questions per pair correctly for five of the eight question pairs and two of four on one pair. Another probe was conducted the next day. Weldon answered the four questions correctly for six of the pairs and two of the four questions for one of the pairs (different from the one in the first one-month probe). Results from the listener and tact probes showed that Weldon was able to receptively or expressively label stimuli representing the answers to all the questions used in the study.

Figure 7 shows the results from Michael’s RP pretests. After 120 trials, Michael continued to make errors, making 6 errors in the final 15 question block. Figure 8 shows the baseline probes for Michael. Question Pair 1 (bottom of Figure 8) was included in the RP pretest. In the first three of four baseline probes, Michael was answering three of the four questions for Pair 1. For the probes following, he was answering two of the four questions. Because of the
variability in responding to Question Pair 1, the blocked trials procedure was introduced with Question Pair 2.

Figure 9 shows the results of the blocked trials procedure for Michael for each pair taught. The pairs were taught sequentially by pair number, starting with Question Pair 2 and ending with Question Pair 8. For Question Pair 2, Michael passed Step 3 in 12 trial blocks, and made four and zero errors total for Step 4 and Step 5, respectively. For four of the remaining six pairs, Michael passed Step 3 in 5 trial blocks, with only one error made total per question pair. Michael passed Step 3 for the other two pairs, Question Pair 2 and Question Pair 5, in eight and nine blocked trials, respectively. He made two errors in both RP probes for Question Pair 2. Michael made one or no errors in the RP probes following Step 3 for the remaining pairs.

Figure 10 shows the trials to criterion and number of steps necessary for each question pair. Michael acquired the discrimination for Question Pair 2, the first question pair taught using the blocked trials procedure, in 128 trials. The trials to criterion for all but one of the remaining pairs decreased by 63-80%, with the highest number of trials being 47 and lowest of 26 trials. The first pair required all steps in the blocked trials procedure. The remaining six pairs probed to mastery after the first step, with four of those pairs only having one error total before mastery. The eighth question pair was acquired during baseline probes. However, as with Weldon, this was the question pair that was initially used in the random presentation pretest.

The baseline probes for Michael prior to intervention were very stable for all but one pair (the pair used in the RP pretest). Only four times in the baseline probes following mastery for a pair did correct responses fall below three of four. Remedial procedures were needed three times for three pairs in the baseline probes for Michael. A total of 63 trials were needed during the three remedial training procedures. Remedial training trials to criterion ranged from 12 to 36
trials, with an average of 21 trials. Michael answered all questions correctly in the last two baseline probes.

In the one-month probe, Michael answered four of four questions per pair correctly for six of the eight pairs. He answered three of four on the other two pairs, both of which had been through remedial training. Results from the listener and tact probes showed that Michael was able to receptively or expressively label stimuli representing the answers to all the questions used in the study.

Figure 11 shows the results from Grayson’s RP pretests. After 150 trials, Grayson continued to make many errors, after an initial downward error trend, making 12 errors in the final 15 question block. Question Pair 1 was used for the RP pretest. Figure 12 shows the baseline probes for Grayson. Across many of the question pairs, the data fluctuated from two of four correct responses per pair to zero correct responses across baseline probes. The blocked trials procedure was introduced for Question Pair 1.

Figure 13 shows the results of the blocked trials procedure for Grayson. Question Pair 1 began in Step 3, but then moved back to Step 1 after 27 trial blocks during Step 3 with errors on an increasing trend. Due to lack of progress with the original procedures for Step 1, distracter trials were added between the switch questions. Distracter trials were used due to the data indicating that Grayson frequently missed the switch question. Distracter trials after each reinforcer break were added in after there was no change in the number of errors. Finally, a differential observing response was added, and after an initial increase in incorrect responses, Grayson’s responding met criterion for mastery of Step 1. Step 2 was mastered in 6 blocked trials. Question Pair 2 also moved back to Step 1 after inadequate progress during Step 3. Grayson moved through this pair with fewer errors and fewer blocked trials than with Question
Pair 1. Grayson mastered both pairs during the RP probes. He made only one error following Step 2 for Question Pair 1. After many errors in the RP probe after Steps 1 and 2, he made no errors in the RP probe after Step 3 for Question Pair 2. Figure 14 shows the trials to criterion and number of steps necessary for each pair. Grayson acquired the first pair after 1549 trials and the second pair after 481 trials.

The baseline probes, shown in Figure 12, after the first pair was mastered showed that the discrimination was not maintained and remedial procedures were required. Following remedial procedures, in the first baseline probe, Grayson responded to four of the four questions for that pair correctly; however, in the next baseline probe, responding fell to three of four. Given the difficulty Grayson had with acquiring the discrimination and the fluctuation in his data, variability in his baseline probes (i.e., moving on with the prior pair at a three) was accepted. During the baseline probes following mastery of Question Pair 2, Question Pair 1 fell back to two of four correct responses and remedial training was used. A total of 780 trials were needed during the three remedial training procedures. Remedial training trials to criterion ranged from 56 to 603 trials, with an average of 260 trials. Following remedial procedures; however, Grayson still made errors in the baseline probes after passing out of the remedial steps. At this point, we decided that this procedure was not successful enough to warrant more of Grayson’s time. However, we will continue to work with him, using modified teaching procedures, in an attempt to teach him the discriminations.
DISCUSSION

The results show that the blocked trials procedure successfully established conditional stimulus control over intraverbal responses for two of the three participants. Weldon and Michael acquired all eight question pairs. Further, these participants did not need all of the steps in the blocked trials procedure for every question pair. Weldon probed out of one question pair following one step, three question pairs following two steps, and three question pairs following three steps. Michael probed out of all question pairs (except the first pair taught) following one step of the blocked trials procedure. For Michael, the number of trials to criterion decreased following the first question pair with only one exception. The procedure has not been successful so far with Grayson. However, even though he required many trials to reach the criterion for the first question pair, he required significantly fewer trials to criterion for the second pair. Weldon and Michael both acquired one pair during baseline probes without going through the blocked trials intervention. For both participants, this question pair had been used in the random presentation (RP) pretest.

This study adds to and extends previous literature by continuing to test the efficacy of the blocked trials procedure in teaching intraverbal responses. It added to previous research through the use of a pretest to ensure that the participants would not learn using random presentation of questions, prompts, and differential reinforcement. It also addressed suggestions from prior research to determine the optimal block size (Williams et al., 2005; Slocum et al., 2012) by starting in the third step of the procedure. It extended the research of Slocum et al. by using the RP probes to determine the necessity of remaining steps. Most importantly, it extended on Ingvarsson et al. (2014) by applying the procedures used with visual-visual (Saunders & Srapdlin, 1989, 1990, 1993; Perez-Gonzalez & Williams, 2002; Williams et al., 2005; Slocum et
al., 2012) and auditory-visual (Perez-Gonzalez & Williams, 2002) discriminations and applying them to an auditory-auditory discrimination.

The complex auditory-auditory discrimination is an important discrimination to master, especially because of its social importance. These types of discriminations are used constantly in speech (e.g., conversations). As Lovaas et al. (1979) described, the listener needs to be able to respond to all dimensions of speech to be able to understand what was said and respond correctly. The practical implications of teaching complex auditory-auditory discriminations are infinite. With this skill, an individual would be able to respond to questions and/or statements that would enable him or her to participate in the social activities. The individual would also be able to have a conversation about something that could change his or her future behavior (e.g., engaging in a conversation about politics, which leads to the individual doing more research that causes him or her to vote in a certain way in an election). This repertoire is especially important to children and adolescents who are continually learning and developing social skills. Without the ability to make conditional discriminations, the child may not respond to a peer’s question or statement appropriately, which could cause that peer to no longer engage in conversation with that child. These types of interactions could be detrimental to a child’s social development.

The blocked trials procedure was successful in teaching conditional discriminations to two of the three participants who had previously failed to acquire the discriminations using random presentation, differential reinforcement, and prompts, as in the RP pretest condition. There are many different possibilities for why the trials presented in blocks facilitated acquisition. An error pattern that was observed during the RP pretest condition was randomly alternating between the two responses. For example, when the questions “what do you kick” and “what do you kick with” were presented randomly, the responses given were sometimes “ball”
and sometimes “leg.” When the trials were presented in the blocks, that error pattern was broken by requiring and reinforcing the same response over consecutive trials. The switch between the two antecedent stimuli may also become more salient when the trials are presented in blocks. Finally, there was an increase in the density of reinforcement due to fewer errors. These are a few possibilities as to why the blocked trials procedure was successful.

As discussed above, the RP pretests were conducted to determine whether the blocked trials procedure was needed for each participant. From the RP pretest results, it was clear that presenting the question in a quasi-random order was not an effective and efficient teaching procedure. This shows that they were good candidates for a more systematic method of teaching. If the participants had been able to acquire the skills in the RP pretests, then the blocked trials procedure would not have been indicated, because this method may require a greater number of trials.

The procedure may have facilitated acquisition over time, because the number of trials to criterion generally decreased over successive targets. For all participants, the trials to criterion were significantly higher for the first question pair than the remaining pairs. The criterion probe performance (RP probe) showed that not all teaching steps were needed every time. In addition to fewer trials to criterion for successive targets, some of the pairs were mastered during the probe without going through each step of the procedure.

The RP pretest was not conducted again after baseline probes began. In future research, the procedures could be altered such that after teaching a certain number of question pairs (e.g., 2-3), RP pretest conditions would be replicated with a novel question pair. This would provide an evaluation of to what extent it is necessary to use the blocked trials procedure across multiple discrimination training targets. However, it might be the case that presenting the questions in
blocks continues to facilitate acquisition beyond this point in training. If the questions were no longer organized in blocks, the rate of acquisition might be detrimentally affected.

To illustrate the previous point: Michael acquired four of the final five question pairs in 26 trials, which is the minimum number of trials that could occur in Step 3, unless no errors were made. If at some point before all of these questions were taught, the teaching procedures had been changed to the quasi-random question presentation used in the RP pretest, would he have been able to acquire the pair in fewer than 26 trials? This is a good question for future research to consider: Is the blocked trials teaching procedure a procedure that could facilitate future acquisition of skills as long as the blocking procedure is still used, or could it facilitate future acquisition of skills using a different procedure to teach the same type of skill? Finding the most effective and efficient teaching procedure is very important, in order to ensure that valuable time is being used in the best way when working on skill acquisition.

Past researchers suggested that future research should evaluate the necessity of each step and the optimal size of trial blocks (Williams et al., 2005; Slocum et al., 2012). The current study evaluated both of these issues. The RP probes tested for the necessity of each step by presenting criterion-level probes. The results suggest that not all of the steps were always necessary, but this finding was somewhat inconsistent. For Michael, Steps 4 and 5 of the blocked trials procedure were only needed for the first pair, but for Weldon, these steps were needed beyond that point. Remedial training was needed for at least a few question pairs for all participants. This might call into question whether the RP probes are a good indication of mastery; however, in most cases, the remedial training was implemented for question pairs that had been mastered prior to the most recent target. Therefore, the issue is likely one of maintenance rather than acquisition. RP probes should be included in future research to help evaluate the necessity of each step.
The optimal size of the blocks was evaluated in this study by starting in the middle of a five step procedure. When starting in the middle, there is always the opportunity to go back to the first step when it is needed, but when it is not needed, the learner will not need to go through unnecessary steps. Two of the three participants in this study did not need the first two steps of the procedure, potentially cutting back the number of trials necessary to acquire the targets. Steps 1 and 2 require at least 40 and 32 trials each, and this is only possible when no errors are made. Having to go through those steps would add another 72 trials at minimum. Starting at Step 3 and moving back only when necessary could potentially cut back the total number of trials to criterion, which would decrease the amount of time spent on each target. Future research should continue to evaluate the optimal number of trials per block.

When moving between the different phases of the study, sometimes more than one phase was completed in the same session. A few times a question pair met criterion in a blocked trials condition, which would result in the immediate presentation of an RP probe or baseline probe in the same session. Because the blocked trials procedure included prompts following incorrect responses, it is possible that the temporal proximity of the prompts may have influenced subsequent probe performance. Future studies could include a break (e.g., one day) between mastery during blocked trials and the RP or baseline probes. Conducting the probes after a break, or conducting the probes both immediately following mastery and after a break, might provide for a better test of mastery and retention. For Grayson, this rule was put into place after a pattern of results was observed in which question pairs reached criterion in probes if they were conducted immediately following the blocked-trials procedure, but poorer performance was observed if the probes were conducted the next day.
A potential deficit in attending (i.e., listening to the questions) was a problem for one of our participants throughout the study. When selecting participants prior to the study, the participants’ attending did not appear to be a problem; however, after starting the blocked trials procedure with Grayson, lack of attending (e.g., moving around in the chair, staring at the wall or past the experimenter, mumbling words, etc.) started to cause problems in sessions. It became very difficult for the experimenter to gain Grayson’s attention (i.e., eye contact and being quiet), and even when eye contact was made, it was still unclear if he was attending completely each time. The experimenter would attempt to test for attending, by moving slightly to the left or right. If the participant’s eyes followed her, this was the best indication that he was attending.

Anecdotally, it was apparent to the experimenter that when the textual prompts were used, Grayson attended to the card (i.e., looked at the card and followed the card if it moved). There was much less looking off around the room, picking at his hands or clothes, or mumbling words. Evaluating the participant’s attending and prerequisite skills (e.g., being able to sit at a table for at least 10 minutes doing work, with one or two reinforcer breaks in the same room) would be suggested in future research. It was during the condition using the textual prompts that Grayson made the fewest errors. Identifying the most effective prompts for each participant before starting the procedure could help maximize the effectiveness and efficiency of the blocked trials procedure.

A limitation to the study is that the listener and tact probes were conducted after the completion of the study. These probes were intended to evaluate whether the participants would be able to respond as listeners and/or speakers to stimuli representing the answers to the target questions. Since the probes were conducted at the conclusion of the study, it is unclear whether the participants had these responses in their repertoire prior to the study. However, it should be
noted that behavioral and educational assessments showed that all the participants had quite extensive listener and tact repertoires prior to the study. Therefore, it is most likely that the participants had learned these responses prior to the study. Nevertheless, future research should conduct these probes when conducting the pretests to qualify for the study. Another limitation is that interobserver agreement was not obtained for treatment integrity data. Future research should include IOA for this measure.

The nature of the stimulus control established in the current study is worth considering. As discussed in the introduction, a conditional discrimination is one in which one stimulus affects the function of another stimulus. Compound stimulus control occurs when two (or more) stimuli or stimulus components function together to evoke a response that is different than the responses evoked by the stimuli or stimulus components in isolation (Eikeseth & Smith, 2013). The discriminations taught in this study were identified as a conditional discrimination because the addition of a stimulus (i.e., “with”) appears to change the function of another stimulus (e.g., “kick”). It could be argued, however, that we did not teach the discrimination in the way that one stimulus changes the function of the other, but rather we taught the stimuli together as a compound. For example, we taught the participants to respond to “kick with” as a compound stimulus working together as one stimuli.

Axe (2008) and Sundberg and Sundberg (2011) refer to this type of discrimination as a conditional discrimination, which is why we will use this term to describe the discrimination in this study. However, it could be argued that the discriminations required in the study were in fact conditional because the same stimulus “with” was used across many different sentences, and in all of those the function of a stimulus (i.e., a word in the sentence) changed with the addition of this word. Further, with Michael, only one total error was made for many of the questions taught.
The fact that only one trial was necessary could indicate that the conditional stimulus “with” was indeed functioning as such. However, it is not clear that interpretation in terms of conditional vs. compound stimulus control makes a practical difference when assessing and teaching these types of complex auditory discriminations.

As discussed above, the social implications of developing a strong auditory-auditory repertoire and the ability to make conditional discriminations are significant. The significance of teaching and expanding these repertoires for children with autism is even more significant because of the lack of social skills that they typically present. Most children with autism have a general lack of interest in social interactions. This combined with the inability to answer complex questions and respond to comments could add to the lack of interest in social interactions because these interactions will most likely not be reinforcing. Finding the most effective way to teach these discriminations and applying these methods in the teaching environment could facilitate more social interactions. This could occur because once the child is able to interpret and respond to language requiring conditional discriminations, these types of interactions may be more reinforcing to the child.

Michael’s mother, when she was being informed that he had completed the study, told the experimenter that she had noticed an improvement in Michael’s question answering at home over the last couple of months. She stated that she used to tell him what she was going to do and ask him questions about it, but that he would not respond to her questions. She stated that at the conclusion of the study, he would respond to the questions that he previously did not respond to. We cannot know whether the teaching procedure facilitated these new intraverbal responses, but it is something worth noting. A suggestion to future researchers would be to assess these types of
responses in the home (or other) environment(s) to be able to accurately measure potential indirect effects.

In summary, the blocked trials procedure was successful in teaching conditional discriminations to two of the three participants. The procedure was especially effective for one participant, who after acquiring the first question pair, only needed the first step of the procedure to acquire all subsequent pairs. The RP probes were valuable in indicating if all of the steps were needed to teach each question pair. The results also showed that starting at Step 3 with a block of five consecutive correct responses as the criterion was sufficient in many cases. Future research should continue to evaluate the effectiveness and efficiency of the blocked trials procedure for teaching conditional discriminations.
Table 1

*Question Pairs and Sets by Participant*

<table>
<thead>
<tr>
<th>Participant</th>
<th>Set 1</th>
<th>Set 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weldon</td>
<td>1. Eat (food, fork)</td>
<td>5. Dig (dirt, shovel)</td>
</tr>
<tr>
<td></td>
<td>2. Play (games, toys)</td>
<td>6. Kick (ball, leg)</td>
</tr>
<tr>
<td></td>
<td>3. Write (words, pencil)</td>
<td>7. Taste (food, mouth)</td>
</tr>
<tr>
<td></td>
<td>4. Sweep (floor, broom)</td>
<td>8. Color (pictures, crayons)</td>
</tr>
<tr>
<td>Michael</td>
<td>1. Play (games, ball)</td>
<td>5. Kick (ball, leg)</td>
</tr>
<tr>
<td></td>
<td>2. Dig (sand, shovel)</td>
<td>6. Read (book, eyes)</td>
</tr>
<tr>
<td></td>
<td>3. Wash (dish, soap)</td>
<td>7. Eat (lunch, fork)</td>
</tr>
<tr>
<td></td>
<td>4. Cut (paper, scissors)</td>
<td>8. Write (words, pencil)</td>
</tr>
<tr>
<td>Grayson</td>
<td>1. Kick (ball, leg)</td>
<td>5. Color (pictures, crayons)</td>
</tr>
<tr>
<td></td>
<td>2. Play (games, toys)</td>
<td>6. Sweep (floor, broom)</td>
</tr>
<tr>
<td></td>
<td>3. Eat (food, fork)</td>
<td>7. Taste (food, mouth)</td>
</tr>
<tr>
<td></td>
<td>4. Write (words, pencil)</td>
<td>8. Hear (sounds, ears)</td>
</tr>
</tbody>
</table>

*Note.* The number in front of the set name indicates the pair number. The two words in parenthesis are the answers to the two questions (i.e., “what do you ___?” and “what do you ___ with?”) given by each participant. These are the actual responses given by the participants.
Figure 1. This figure displays the flow between the conditions. The arrows indicate what condition occurs next. The text on the errors indicate what the criterion is to move to each condition. For the arrows moving to and from the steps and RP probe, follow the errors from left to right beginning at the arrow from the current step.
**Figure 2.** The number of prompted responses (number of errors) during the first random presentation pretest for Weldon are displayed.

**Figure 3.** Baseline probes for Weldon are displayed. The number in each box represents the number of correct responses per question pair. Four questions per question pair were asked in each probe. A “4” in a box indicates that all of the four responses were answered correctly. Boxes with a “3” or “4” are shaded with a light shade and dark shade, respectively. Boxes with no shade indicate a “2” or lower for number of correct responses. The thick line through the middle of the data indicates the implementation of the blocked trials procedure. The boxes to the left of the line have not been through the blocked trials procedure; boxes to the right of the line have been through the blocked trials procedure.
Figure 4. The number of prompted responses (number of errors) during the second random presentation pretest for Weldon are displayed.
Figure 5. The number of prompted responses (number of errors) are displayed by question pair for Weldon.
**Figure 6.** The trials to criterion for each question pair and steps necessary to complete the blocked trials training for Weldon are displayed. The last step that was required for each pair is written above the bar for each question pair.
Figure 7. The number of prompted responses (number of errors) during the random presentation pretest for Michael are displayed.
Figure 8. Baseline probes for Michael are displayed. The number in each box represents the number of correct responses per question pair. Four questions per question pair were asked in each probe. A “4” in a box indicates that all of the four responses were answered correctly. Boxes with a “3” or “4” are shaded with a light shade and dark shade, respectively. Boxes with no shade indicate a “2” or lower for number of correct responses. The thick line through the middle of the data indicates the implementation of the blocked trials procedure. The boxes to the left of the line have not been through the blocked trials procedure; boxes to the right of the line have been through the blocked trials procedure. Question Pair 7 was added to Set 2 during the fourth baseline probe to replace a pair that was taken out.
Figure 9. The number of prompted responses (number of errors) are displayed by question pair for Michael.
Figure 10. The trials to criterion for each question pair and steps necessary to complete the blocked trials training for Michael are displayed. The last step that was required for each pair is written above the bar for each question pair.
Figure 11. The number of prompted responses (number of errors) during the random presentation pretest for Grayson are displayed.
Figure 12. Baseline probes for Grayson are displayed. The number in each box represents the number of correct responses per question pair. Four questions per question pair were asked in each probe. A “4” in a box indicates that all of the four responses were answered correctly. Boxes with a “3” or “4” are shaded with a light shade and dark shade, respectively. Boxes with no shade indicate a “2” or lower for number of correct responses. The thick line through the middle of the data indicates the implementation of the blocked trials procedure. The boxes to the left of the line have not been through the blocked trials procedure; boxes to the right of the line have been through the blocked trials procedure.
Figure 13. The number of prompted responses (number of errors) are displayed by question pair for Grayson.
Figure 14. The trials to criterion for each question pair and steps necessary to complete the blocked trials training for Grayson are displayed. The last step that was required for each pair is written above the bar for each question pair.
REFERENCES


