TRAINING A NON-MATCH RESPONSE: TOWARD A TECHNOLOGY FOR
DETERMINING CONTROLLING STIMULUS DIMENSIONS FOR
TWO CHILDREN WITH AUTISM

Tanya Yvonne Baynham, B. A.

Thesis Prepared for the Degree of
MASTER OF SCIENCE

UNIVERSITY OF NORTH TEXAS

December 2007

APPROVED:

Jesús Rosales-Ruiz, Major Professor
Shahla Ala‘i-Rosales, Committee Member
Manish Vaidya, Committee Member
Richard Smith, Chair of the Department of Behavior Analysis
Thomas Evenson, Interim Dean of the College of Public Affairs and Community Service
Sandra L. Terrell, Dean of the Robert B. Toulouse School of Graduate Studies

Match-to-sample procedures traditionally train match responses but do not train non-match responses that could allow researchers and practitioners to better infer stimulus control. This study evaluated the effectiveness of 2 stimulus control shaping procedures in training a non-match response to 2 children with autism. Two-step prompting and response blocking was successful in producing criterion matching and non-matching performance for 1 participant. Correct responding did not maintain after feedback and praise was removed. A stimulus control shaping program was implemented with a second participant to transfer the non-match response of head shaking to the criterion response, placing a stimulus into a box. She learned to perform the criterion non-match response in the presence of 3-dimensional stimuli but not in the presence of 2-dimensional stimuli. The utility of training a non-match response and considerations when developing training methodologies are discussed.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF FIGURES</td>
<td>iv</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>GENERAL METHOD</td>
<td>11</td>
</tr>
<tr>
<td>CHRIS: PROGRAM 1</td>
<td>20</td>
</tr>
<tr>
<td>ABBY: PROGRAM 2</td>
<td>32</td>
</tr>
<tr>
<td>GENERAL DISCUSSION</td>
<td>42</td>
</tr>
<tr>
<td>APPENDICES</td>
<td>48</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>59</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>Chris' baseline phase data</td>
<td>22</td>
</tr>
<tr>
<td>2.</td>
<td>Chris' pretraining phase data</td>
<td>27</td>
</tr>
<tr>
<td>3.</td>
<td>Chris' non-match training phase data</td>
<td>29</td>
</tr>
<tr>
<td>4.</td>
<td>Abby’s baseline phase data</td>
<td>32</td>
</tr>
<tr>
<td>5.</td>
<td>Abby’s pretraining phase data</td>
<td>36</td>
</tr>
<tr>
<td>6.</td>
<td>Abby’s non-match training phase data</td>
<td>40</td>
</tr>
</tbody>
</table>
INTRODUCTION

Stimulus overselectivity refers to responding under control of one or a small number of dimensions, characteristics, or properties, of a complex stimulus (Lovaas, Koegel, & Schreibman, 1979). It is commonly observed in persons with autism, particularly those of low mental age (Wilhelm & Lovaas, 1976). The rationale for remediation of this deficit has been well noted; responding to a minimal number of stimuli in the environment has been associated with difficulties in acquiring simple and conditional discriminations (e.g., Stromer, McIlvane, & Serna, 1993), language (e.g., Rincover & Koegel, 1975), and social behavior (e.g., Schreibman & Lovaas, 1973). In fact, overselectivity is such a prevalent and pervasive problem that training responsivity to multiple cues within a complex stimulus is considered pivotal in the treatment of children with autism (Burke & Cerniglia, 1990; Koegel, Koegel, & McNerney, 2001).

The measurement of stimulus overselectivity has a long and illustrious history within behavior analysis. In initial demonstrations of the phenomenon of overselectivity, behavior analysts set up conditions that were likely to produce overselectivity and then measured its existence. In a seminal study, Reynolds (1961) differentially reinforced two pigeons’ responding in the presence of two complex stimuli: a triangle with red background and a circle with green background. He then presented the four stimuli (the colors red and green and the shapes triangle and circle) separately during test trials and found that one pigeon’s responding was allocated to the triangle and the other pigeon’s responding was allocated to the red stimulus. Both pigeons’ responses met the programmed reinforcement contingency but different dimensions of the complex stimulus controlled their responses.
Lovaas, Schreibman, Koegel, and Rehm (1971) demonstrated similar results with children with autism. They initially reinforced responding in the presence of a complex stimulus composed of auditory, visual, and tactile components. Then, they presented each component separately and found children with autism responded primarily in the presence of one of the components (though the component to which they responded was not consistent across participants). Using the same procedures, Lovaas and Schreibman (1971) demonstrated that overselective responding could occur even when complex stimuli had only two components.

Later, behavior analysts applied similar procedures to first assess the existence of overselectivity, and then to utilize information derived from the assessment to improve intervention. For example, Rincover and Koegel (1975) assessed the stimulus control of four participants’ responding when, following training, responding failed to transfer to novel settings. After isolating potential controlling variables, as Reynolds had, the authors discovered that a different, unintended variable controlled each child’s response (e.g., an incidental hand movement or an initial prompt). Rincover and Koegel not only ascertained the controlling stimulus-response relations, but also programmed for responding to come under the control of experimenter-intended discriminative stimuli.

Clinically, one may argue that Rincover and Koegel’s analysis is not necessary when generalization reliably occurs. However, when a learner isn’t performing as the teacher intends, it becomes necessary to determine the existing, unintended, stimulus controls. In the Rincover and Koegel study, experienced clinicians identified the potential controlling variables after observing multiple training sessions. Clinicians with
such expertise are not common, which may be one reason that the established method of determining controlling variables has not been adopted in popular teaching practice. Clearly, measurement technologies that are easier for general practitioners to implement are required.

There are two general theories related to the measurement of overselectivity that may affect what types of technologies will be devised. The traditional approach to measurement is consistent with a descriptive stimulus control interpretation. This approach assumes a continuous relation between a dimension of the training stimulus and measured behavior (Bickel, 1987). In practice, overselectivity has been defined within continuous relations as responding to one of two elements of a complex stimulus on 100% of the test trials and at “chance” levels (25% - 75%) to the other element. The interpretation is that the stimulus control is restricted to one S+ element while the other S+ element does not exert reliable control (i.e., it occasions random responding).

The second approach to the measurement of overselectivity is the quantal interpretation of stimulus control. The primary tenet suggests that there is never random responding, just control that is difficult to measure because it is not control specified by the experimenter (Ray, 1969; Sidman, 1980). Bickel, Stella, and Etzel (1984) applied this interpretation to demonstrate and describe hierarchies of stimulus control with preschoolers. After training 27 children to respond to one complex auditory stimulus and not to another complex auditory stimulus, six combinations of the four elements of the complex were presented as stimuli pairs. Children were labeled overselective when their responding occurred in the presence of certain S- elements more than other S+ elements, which resulted in S- elements being placed higher on a hierarchy of stimulus
control. This study approached a quantal interpretation of stimulus control because stimulus control was measured in a binary way on each trial before the placement of the stimulus elements on the hierarchy.

In a systematic replication conducted with adults with mental retardation, Bickel, Richmond, Bell, and Brown (1986) again showed responding that appeared as chance, or random responding, when averaged together but were actually contextually controlled stimulus-response relations not associated with the physical properties of the stimulus. These relations included responses to stimulus position, responses to S-elements, and intraexperimental historical intrusion (e.g., discrimination reversals). The authors concluded that this contextual control confirmed that “properties of the S+ and S- elements are neither inherent nor static, but are instead a relative property in which the type of control exerted depends on both the stimulus and the context of its presentation” (Bickel et al., 1986, pp. 235). The additional consideration of contextual control allows for the assertion that control is exerted by the S+ elements under Conditions A and by the S- elements under Conditions B (a binary analysis).

There are interpretations of data that support a quantal interpretation of stimulus control, such as those presented above, and there are data interpretations that support a descriptive analysis of stimulus control. The value of each interpretation is still in question (see Schneider & Morris, 1988 for a discussion). However, autism practitioners will likely ignore conceptual aspects of the argument in favor of a more pragmatic approach: which interpretation will lead to more effective treatment?

Stimulus overselectivity has been observed, and remediation has been attempted, during matching tasks with visual stimuli (e.g., Dube & McIlvane, 1999;
Hedbring & Newsom, 1985). Matching is fundamental to stimulus equivalence class formation (e.g., Sidman, 1986; Maguire, Stromer, MacKay, & Demis, 1994), the acquisition of verbal behavior (e.g., Sidman, 1986), and early reading skills (e.g., Saunders, Johnston, & Brady, 2000). It is not surprising then that identity matching (i.e., matching based on all physical properties of the stimulus) is outlined in most published early intervention training texts and curricula for children with autism (e.g., Leaf & McEachin, 1999; Lovaas, 2003; Taylor & McDonough, 1996).

Within these curricula, most recommended matching training programs are based on standard identity matching to sample (IMTS) procedures (Cumming & Berryman, 1965). During these procedures, a sample stimulus is presented to the child. The child then engages in some kind of observing response (e.g., taking the card or touching the sample on a computer screen) and is directed to a number of comparison stimuli. The teacher provides differential consequences; if the child matches to a physically identical comparison (S+), reinforcement is delivered, if the child matches to a non-identical comparison (S-), no reinforcement, or a correction procedure, is provided. A child demonstrates generalized identity matching when correct responding occurs in the presence of novel stimuli and stimuli can serve as both S+ and S- stimuli in quick succession, a preparation described as conditional-function identity matching (Dube & Serna, 1998).

In many cases, this simple procedure is successful in training identity matching. However, when typical training procedures fail, practitioners must wade through the myriad of other intricate and time intensive approaches that produce correct responding under MTS conditions. Examples of alternative training procedures include criterion-
related (or within-stimulus) shaping (e.g., Schilmoeller & Etzel, 1977; Schreibman, 1975), verbal observing responses (e.g., naming) and nonverbal differential observing responses (e.g., touching each component of a compound stimulus or signing; Dube & McIlvane, 1999; Lowenkron, 1988), blocked trial training (e.g., Saunders & Spradlin, 1989), and reinforcement schedule manipulation (e.g., Dube & McIlvane, 1997; Koegel, Schreibman, Britten, & Laitinen, 1979). These alternative training procedures either obviate or assume stimulus control that may not exist. The procedures help to achieve stimulus control transfer, which is important from a therapeutic standpoint. However, it is also imperative to present demonstrations of how and why the transfer has occurred. Determining how and why transfer has occurred allows practitioners to better choose from the various procedures based on the client’s current repertoire and needs.

Recently, a particular quantal interpretation of stimulus control has been presented that may lead to effective tactics for choosing procedures to remediate stimulus overselectivity. McIlvane and Dube (1992) have adopted the term stimulus control topographies (SCTs) from a 1969 paper by Barbara Ray and suggest conceptualizing different controlling relations as different SCTs. It is suggested that there are many SCTs to which a learner can respond, and determining these controlling relations can lead to more effective intervention.

Suppose, for example, in a traditional IMTS task, a learner is given a picture of a pig and must match it either to a pig, cow, or horse presented in that serial order on the table. If the child places the sample on top of the pig, the teacher could surmise that the learner matched on the basis of all features common to both sample and comparison. True, the child’s matching might be controlled by this teacher-intended contingency.
However, she also might be matching to the position associated with the most recent reinforcement history, to an upright tail, or as ample research has shown (e.g., McIlvane, Kledaras, Lowry, & Stoddard, 1992; McIlvane, Kledaras, Munson, King, De Rose, & Stoddard, 1987) by exclusion of the cow and horse stimuli.

In reference to persistent responding based on position, Sidman (1992) pointed out that experimenters don’t realize that these are not causes of subjects’ failures to learn, but are usually products of their failure to learn (p. 173). This can be said for all unintended SCTs. We cannot know whether to use, for example, a nonverbal differential observing response or a strategy to control for position without assessing the current control. Unfortunately, we cannot easily perform assessments with the prevailing methodologies for training identity matching. Assessment of stimulus control is particularly difficult to conduct within the traditional IMTS procedure because of the forced choice; we risk creating unintended stimulus control as we try to identify it.

Ray (1972) offers some advice on how to distinguish controlling stimuli from those stimuli that are coincidental. She states eloquently:

We have to ask the experimental subject. We ask by arranging features to which the subject can selectively respond. To the extent that an experimenter limits the subject’s choice, he deprives the subject of his communication device. The range of features the experimenter presents, his lack of prejudice, and his imagination will all hasten valid measurement of stimulus control. (p. 297)

An MTS variation that includes a default-response option may be less likely to force specific experimenter intended (or unintended) stimulus control and is less likely to limit subject choice (e.g., Fields, Reeve, Adam, Brown, & Verhave, 1997; Innis, Lane, Miller, & Critchfield, 1998; Duarte, Eikeseth, Rosales-Ruiz, & Baer, 1998; Smith, Schull, Strote, McGee, Egnor, & Erb, 1995). Default-response options provide the subject with
an alternative response to selecting the comparison stimuli presented. In terms of verbal behavior, at least as revealed by the instructions used in the experiments, default-response options can indicate uncertainty (e.g., I don’t know or can’t answer), rejection of the stimuli present (e.g., neither, none of the above, or no match), or an alternate relation between the sample and comparison (e.g., different than). Previously cited studies demonstrate that it is likely that other controlling relations will be revealed when a default-response option is present. Unfortunately, in most existing research the controlling relation often remains ambiguous because, as in the case of the can’t answer response described by Duarte and colleagues (1998), the response is not explicitly trained and is not reinforced as a correct response. In this way, it is difficult for the experimenter to infer specific controls. Therefore, it is beneficial to train the default option to have a specific function so that it can be used to assess existing but unrevealed stimulus-response relations following training using traditional MTS procedures, such as those used in typical early intervention programs.

Rosales-Ruiz (1983) was the first to employ such an assessment technology. An apparatus was composed of three boxes and Kanji letters, including a center box where children placed any samples that did not match two Kanji comparisons on either side. Typical children were trained to use the apparatus aided by verbal instructions (i.e., Put in this box the ones that are the same,” and “Put in this box the ones that are different,” and an explanation that the experimenter would not say whether answers were correct or incorrect). The apparatus revealed (a) consistent response patterns for each subject and (b) that incorrect matches to the Kanji comparisons were based on particular
dimensions of the stimulus (e.g., similarities between the top or bottom portions of the letters).

Rosales, Pasley, and Potucek (1992) employed a tabletop sorting task similar to the one used in the current study and demonstrated how a default option (in this case, the middle box indicating different than, or neither of, the comparisons present) can reveal interference by properties of preceding trials. A 23-year-old verbal man (with a history of instruction following) with autism correctly sorted samples to the neither box if that sample had not appeared as a sample with a physically identical comparison in previous trials. The study illustrates the context-specific nature of control by identity. The environmental context always enters, explicitly or implicitly, into the relation. This is consistent with Bickel and colleagues (1986) assumption that stimulus control is context-dependent. In the current study, I evaluate training procedures appropriate to persons who are nonverbal and discuss how the sorting task described by Rosales and colleagues (1992) can be used to assess overselective responding in children with autism.

Serna, Wilkinson, and Mcllvane (1998) used a default-response option to assess identity-matching stimulus relations of adolescents with mental retardation. They trained a selection-based method of assessing yes/no relations using a blank-comparison MTS task (BC-MTS). In BC-MTS, participants view a display containing the sample and at least one comparison stimulus plus a blank-comparison (often a black square). If there is a comparison present that matches the sample, the participant selects it. If there is not a correct comparison present, the participant defaults to the blank comparison. A stimulus-control shaping procedure was used to establish BC-MTS whereby the blank
comparison initially was a physically identical correct comparison on some trials. Across trials, the choice faded to black while alternating between an S- and an S+ function. The authors framed their analysis in terms of participants’ similarity/dissimilarity judgments. The current study closely resembles a BC-MTS procedure except that the participants are nonverbal children with autism, a blank index card serves as the blank comparison, the stimuli are presented on a tabletop instead of a computer, and the training procedure differs from the one described above.

To summarize, published studies have demonstrated the efficacy of an MTS variation that includes a default-response option as an assessment tool for revealing stimulus overselectivity. However, given the correlation between overselectivity and mental age and the direct relationship between mental age and verbal skills, it is imperative to devise teaching technologies without verbal instructions that provide information about the controlling relations acquired during traditional MTS tasks.

The methodology presented in this study provides more information about measurement of stimulus control in general and its extension to overselective responding of children with autism in particular. Thus, the purpose of this study is to present potential stimulus control shaping (see McIlvane & Dube, 1992 for a discussion of the term) programs to establish a non-match response that will allow teachers to make informed inferences about sources of control for matching responses of persons with limited verbal repertoires.
GENERAL METHOD

Participants

Both participants were chosen for inclusion in this study because they had a limited verbal repertoire and they were working with me, the primary investigator and experimenter, while I provided applied behavior analysis (ABA) services within their home-based treatment program. Both parents and the behavioral consultant agreed that the component skills trained during the study (e.g., sorting stimuli, the non-match response, tracking a stimulus) would be clinically relevant and assessment results from the non-match phase could help with future skill building.

Participant A (Chris) was a 3-year-old boy diagnosed with autism by a general practitioner at 2.6 years of age. Throughout the study, Chris received between 15-20 hours a week of ABA therapy and attended a typical preschool two mornings a week with support. Treatment emphasis was on increasing language and social interactions, and expanding play skills. At the onset of the study, Chris spontaneously used 3 words consistently: up, no, and chicken. During the course of the study, the number of words increased to 11. Chris requested items using approximations of words or by pointing to the item. He followed six verbal instructions (e.g., wave bye-bye, blow kisses, sit down) and could pick eight items (e.g., chips, soda, shoe, book) from an array on a table when given the instruction, “Give me item name.” Chris also had a large imitative repertoire and readily learned new actions in 1:1 training situations without intensive intervention (both gross and fine motor). A matching program was introduced in November 2003 and he demonstrated conditional identity matching in June 2004. Conditional generalized identity matching is defined as matching based on physical properties of the stimulus
when discriminative function reversals (i.e., the S+ was sometimes the S-) occur frequently and with novel stimuli (for a discussion, see Dube, McIlvane, & Green, 1992).

Participant B (Abby) was an 8-year-old girl diagnosed with mental retardation, brain anomaly, and autistic tendencies. She had also recently been diagnosed with extropia and hypertropia which produces blurred vision. (Abby began tilting her head and holding objects close to her eyes.) During pretraining she started receiving eye-drops that dilated her left eye for part of the day. Later in the pretraining phase, Abby had strabismus surgery on her left eye but this did not correct the condition. At the time the study was conducted she had been receiving 15 hours a week of ABA services in her home for two years. Ten hours a week in a clinic setting had recently been added to her program. The primary treatment goals were increasing functional communication and self-help skills, and increasing opportunities for, and appropriateness of, social interactions with family and members of the community. She also attended public school in a special education classroom two days a week.

Through intensive intervention, Abby had acquired a verbal repertoire of seven modified signs that she used when requesting items (e.g., movie, book, jump). Of potential importance, she learned to shake her head, “No” in the presence of unwanted items during the study. She followed five verbal instructions (e.g., stand up, sit down, clap your hands) without contextual cues. Abby could imitate over 25 actions but did not demonstrate generalized imitation as defined by Baer, Peterson, and Sherman (1967). She demonstrated generalized conditional identity matching, as defined previously, with both 3D and 2D objects after two months of training.
Setting

Each session was conducted in various rooms in the participants’ homes or in a small observation room in the Department of Behavior Analysis at the University of North Texas. All rooms contained one table and a chair. A digital video camera was set up on a tripod after the participant had entered the room of choice. An assortment of toys and treats were always present. The specific items in the rooms varied throughout the study. Training was conducted loosely (Stokes and Baer, 1977) to increase the likelihood of generalization to multiple contexts. Many stimuli were varied across phases; family members and dogs were sometimes present, setting changed based on participant choice, and sessions were held at various times of day based on participant schedule.

Stimuli and Apparatus

The stimuli used (see Appendix A) were nine picture cards from the Pooh Memory™ Game (Milton Bradley Company, www.hasbro.com): Pooh with inner tube (PI), Pooh as farmer (PF), Pooh with backpack (PB), Owl with sandcastle (OS), Owl with kite (OK), Owl under an umbrella (OU), Piglet in a sailboat (PigS), Piglet with carrot (PigC), and Piglet with cornucopia (PigCo). Chris had no history with the cards. Prior to the experiment onset, Abby had difficulty matching the stimuli when they were presented during leisure activities (i.e., while playing Memory). In addition to the Memory cards and boxes, in the non-match training phase there was a blank card the same size and shape as the Memory cards. During Abby’s non-match training phase, over 20 pairs of toys were also used (e.g., balls, blocks, jacks, and cars) as matching stimuli. As well, a large toy bin, a blue grocery bin, a red plate, and a red 5 X 5-inch box
with a 2.5 X 1-inch slot served as the non-match apparatus. During Chris’ reinstatement of control phase, Jay-Jay the Jet Plane and Arthur Memory™ cards (Milton Bradley Company, www.hasbro.com) were used.

The apparatus consisted of three 5.1 X 5.1-inch square boxes covered in beige newsprint. The top of each box had a 2.5 X 1-inch slot (big enough to minimize errors due to fine motor problems and small enough to conceal prior matches to the box) where the cards were placed. The bottoms of the boxes were cut out so that the cards could be retrieved quickly by lifting the box.

Measurement

The behaviors of interest in this study were matching and non-matching. Multiple measures of this behavior were taken. Primary measures included correct and incorrect responding, trial duration, and session duration. Data were also taken on assent and the number of prompts and error corrections. Data sheets and a scoring protocol containing detailed definitions, with inclusion and exclusion examples can be found in Appendix B.

A correct match was defined as either (a) placing the sample on some part of physically identical comparison (in baseline) or (b) placing at least half of the sample in the box corresponding to the physically identical comparison (in pretraining and training conditions).

A correct non-match was defined as placing the sample in at least half of the box corresponding to the blank card, or during Abby’s non-match training phase, a red plate/box.

An incorrect match was defined as placing the sample any place not defined as a correct match or not taking the card within 5-s of a verbal direction to match.
Duration data were taken from video-files viewed after the session. Trial duration was recorded beginning the second that the child’s hand touched the sample and ending the second that the child was no longer touching the sample and the sample had contacted the correct comparison or its corresponding box, depending on the phase. Session duration was calculated by adding the total duration of all trials in the session.

Assent was defined as coming to the table in the absence of physical prompts following no more than one verbal direction (e.g., “Come here,” or “Sit down.”) and staying in proximity of the materials without any whining or protest (e.g., banging hands against head) lasting more than 5 s. If the participant said, “No,” when presented with the sample at any time, this was defined as non-assent. Assent data informed clinical decisions but are not presented in the body of this text.

A prompt was defined as behavior emitted by the experimenter other than the initial verbal instruction to match (e.g., blocking the boxes, pointing to the correct comparison) occurring prior to the participant’s first response of each trial. An error correction was defined as behavior emitted by the experimenter (e.g., modeling the 2-step prompt + response blocking sequence, physically guiding the child to touch the sample to the correct comparison) following the child’s incorrect match with the same sample.

Interobserver Agreement

I served as the experimenter and primary data-taker. A second observer was also trained to score reliability data from session footage using the Apple® QuickTime® audio/visual playback software (Apple Inc., registered in the U.S. and other countries, www.apple.com). He was naïve to behavior analysis. During training, I provided the
scoring protocol (see Appendix B), read the protocol with him, and answered any questions. Then, the reliability observer and I watched and scored one session in each phase to allow for further questions and adaptation of the protocol if necessary. When observations matched 100% of those made by the experimenter, for two consecutive sessions, the reliability observer independently scored the remaining taped sessions. Interobserver agreement (IOA) for correct/incorrect data was calculated by dividing the number of agreements by the number of agreements plus disagreements (trial by trial) and multiplying by 100. For session duration, IOA was calculated by adding the total number of seconds scored by each observer, and dividing the shorter duration by the longer duration and multiplying by 100. IOA was assessed between 33 - 43% of sessions across pretraining and non-match training phases. During baseline 25% of Abby’s sessions and 23% of Chris’ sessions were scored due to taping difficulties. Mean agreement for correct/incorrect data across participants was 94.8% (range, 66% to 100%). Duration data were not scored during baseline or during some sessions of Abby’s non-match training due to videotaping difficulties or availability. IOA on duration measures was scored during the same sessions that correct/incorrect IOA was scored. Mean agreement for duration data across participants was 96.8% (range, 75% to 100%).

Design

An AB design partially replicated across participants was used to evaluate the effects of training procedures on match and non-match responses.
Procedure

Procedural Sequence and Rationales

Experimental procedures were carried out in three phases (baseline, pretraining, and training) and each phase consisted of several steps. Each step was designed with particular teaching targets in mind. Therefore, the number of steps within each phase depended upon the child’s skill set. Baseline was conducted to show that both participants could demonstrate identity matching, at least under certain conditions. Also, previous research has shown that overselective responding can be reduced by repeated exposure to testing (e.g., Shreibman, Koegel, and Craig, 1977) so a relatively long baseline was employed to determine whether this would happen with these participants. The purpose of pretraining was to familiarize the participants with the box apparatus and the 2-step sorting task. The goal of the training phase was to teach the non-match response using the same apparatus and sorting skill that the participants were familiar with from the pretraining phase.

Procedures Consistent Across All Sessions

At the beginning of each session, I set up the camera and placed materials on the table or floor. During this time, the child had access to all toys in the room and the experimenter delivered no demands. The experimental stimuli remained available for the entire therapy session.

Experimental sessions began when the participant sat down in front of the stimuli that were placed either on a table or the floor. Sitting down occurred spontaneously or after the child selected a putative reinforcer (e.g., by saying, “chicken,” or reaching for music) and I gave a verbal direction (e.g., “Okay, come sit down.”) Three procedural
components were employed to minimize potential coercion to complete the specific task and to assess the difficulty of the training procedures compared to other programming targets. First, the participant could leave the table at any time without penalty. That is, the child could access all tangibles during other current teaching targets. If the child left the table before completing nine trials, the same session was continued upon return. Second, if the participant left the table without completing a trial or did not come to the table following a verbal direction three consecutive times, the experimental session was terminated for the day. Third, the child provided assent (as defined in the measurement section) throughout the entire session. If at any point, the child did not provide assent (e.g., whined for 6 s or said, “NO!” when given a sample), the child was prompted to leave the experimental area and was required to complete other programming targets (e.g., labeling objects, following instructions) before returning to the experimental session. All of these procedural components were considered measures of assent, were analyzed each session, and influenced whether I initiated another session on the same day or made procedural changes. Assent data are not presented except when non-assent resulted in session termination. This is indicated on the graphs by sessions consisting of fewer than nine trials.

During all phases that included tangible consequences, tangibles were chosen that (a) had been manipulated during the period of free access, (b) had been requested immediately prior to task, or (c) had reliably acted as reinforcers for teaching targets in previous sessions.

Each experimental session consisted of nine trials unless the child terminated the session. Only the author conducted experimental sessions during therapy. Therapy
sessions were between 1-3 hours in duration. All experimental sessions were conducted during scheduled therapy hours 1-3 times per week.
Baseline Procedures

Baseline consisted of two conditions running concurrently and presented in a semi-random multi-element fashion. In Condition 1, when the sample was a specific character (i.e., Pooh, Piglet, or Owl), the comparisons depicted the other two characters. In Condition 2, the character was the same across all trials within a session but the comparisons differed based on feature. For example, Pooh was inner tubing, farming, or backpacking. Position of the comparisons rotated every three trials.

During baseline, the teaching procedures in both conditions were identical. Rather than an unreinforced baseline, a minimal teaching procedure was used to determine whether the skill could emerge with some teaching (i.e., contingent error correction and praise). For a flowchart depicting a minimal teaching session, see Appendix C. Each trial began with the participant sitting facing a field of three comparison stimuli. I sat facing the child. Then, I either gave the sample to the participant or placed it in front of the participant, often providing the verbal instruction, “Match.” The participant took the card and placed it on one of the comparison cards. When the child made a correct match, I provided a social consequence (e.g., descriptive praise, banging the table, cheering, high fives) on a continuous (FR1) schedule. A tangible consequence (e.g., cheetos, a clip of a movie or song, a toy) was provided on an FR3 schedule. After the second correct match (regardless of whether it followed a prompt), I announced that the reinforcer was coming (e.g., “We’re going to get music!”). This schedule remained constant for Abby throughout baseline but was
thinned across three sessions for Chris until he was matching nine cards before a tangible item was delivered.

Following an incorrect response, I modeled the correct response. That is, I took the card off the S- and held it up, ensuring that the child’s eyes oriented towards the card, and labeled the relevant stimulus dimension: character in Condition 1 (i.e., “Pooh,” “Owl,” or “Piglet,”) and feature in Condition 2 (e.g., “kite,” “farmer” or “carrot”). Then, I held the sample against the correct comparison and repeated the relevant dimension before placing it on the card. I handed the card to the participant to complete the trial. If incorrect, I repeated the model and asked the participant to match again. After the third incorrect response, I physically prompted the child to complete the trial and a new trial (with a new sample) was initiated. If, at any time during the correction procedure, the participant appeared to stop tracking the model, the error correction procedure was immediately restarted.

If the child matched the sample to the correct comparison following a prompt, I responded in the same way as if the response had been correct on the first trial. It is possible that nondifferential reinforcement could lead to slower acquisition of the target skill (Cunningham, Cooper, Plummer, & LeBlanc, 1975). However, Koegel, Koegel, and Carter (1999) suggest that a more pervasive, pressing issue in autism treatment may be a general lack of motivation. The term motivation encompasses observable behaviors such as “increases in the number of responses a child makes to teaching stimuli, decreases in response latency, and changes in affect” (p. 21). These authors recommend reinforcing attempts, to increase the likelihood of future responding to tasks and improve the child’s learning during social and academic tasks.
When Chris (and Abby) terminated experimental sessions or left the area, it was most often during the error correction sequence. Therefore, I reinforced correct responses and attempts, defined as responding to prompts or error corrections. I also provided consequences following a response that was a current target in the participant’s program intermittently throughout the session.

**Baseline Results and Discussion**

![Graph showing baseline results for Chris](image)

*Figure 1. Closed circles represent the number of correct matches for Chris during Condition 1 of baseline. Open circles represent the number of correct matches for Chris during Condition 2 of baseline.*

Figure 1 represents Chris’ data during baseline for Condition 1 and 2. Nine sessions were conducted to establish the existence of accurate identity matching, and the remaining sessions were conducted to assess maintenance of matching while the other training phases were implemented. Chris showed high accuracy during the first
nine sessions of both baseline conditions. There was very little variability within and between the two conditions. The data for Condition 1 ranged between 8 and 9 correct matches per session with a mean of 8.7 out of 9 across the first nine sessions. The results obtained in Condition 2 were only slightly lower, ranging between 6 and 9 correct responses per session with a mean of 8.3 across nine sessions. Of interest, it was during the first session of Condition 2, when Chris matched with the least accuracy; matching only six samples correct. Perhaps he did require one day of minimal teaching to establish identity matching when the negative comparisons had dimensions common to the positive comparison (i.e., character).

Because of rapid acquisition during Condition 2, and the data falling within a range considered clinically and experimentally acceptable, no further error analyses were conducted and Chris began pretraining probes. After one month, during which Christmas break and probe sessions within the pretraining phase were conducted, but prior to pretraining, I ran another session in both baseline conditions. Chris did not respond correctly on any of the trials. In fact, he did not respond to the task at all. Potentially, pretraining probes had resulted in some interference of control over the baseline task. After one session of training using the minimal teaching procedure, Chris responded at levels identical to those before the introduction of probes. Overall, the results suggest that Chris demonstrated generalized identity matching without baseline training.
Pretraining Procedures

Probe Sessions for Chris: Because Chris’ baseline data across both Conditions 1 and 2 were stable and accurate, various procedural components were probed with Chris prior to developing the 2-step + response blocking procedure described later in the text and also presented as a schematic in the appendix. A total of seven probe sessions were conducted. During Probe Sessions 1 and 2, two large brown boxes with lids were placed on the table. A comparison was taped to one box and there was nothing on the other box. The experimenter provided tangibles and praise following all correct responses. Following incorrect responses, the experimenter held the sample beside the comparison stimulus. If it was the S+, the experimenter said, “Yes!” and placed the sample in the corresponding box. If it was the S-, the experimenter said, “No!” in a happy tone and then placed it the box without a comparison on it.

“Yes” and “no” were initially chosen because they were close to Chris’ current language level. However, the model “no” evoked emotional responding (e.g., whining, leaving the table, head hitting) four times. Therefore, during Probe Sessions 3 and 4 the experimenter said “same” instead of “yes” and “different” instead of “no.” The boxes were also shortened to be the height ultimately used during the non-match training phase (though they were not covered with newsprint until the non-match training phase) because lifting the lids was more tedious than lifting the box to retrieve the cards.

Initially, the comparisons were placed on top of the boxes because the experimenter reasoned that it might be easier to shape a response if the slot was close in proximity to the comparison. Ultimately the comparisons were placed on the table in front of the box so that the first step in the 2-step sequence was identical to baseline.
Initially, the field consisted of two boxes but a third box was added after two sessions because increasing the field minimized the likelihood that selections based on exclusion (e.g., McIlvane, Kledaras, Lowry, and Stoddard, 1987) would be reinforced. It is also well established that increasing the number of potential choices in a field can decrease the likelihood that the participant’s behavior will enter into contingencies not anticipated by the experimenter (Sidman, 1980). (Within this condition, I also probed density of reinforcement required to establish and maintain matching behavior.)

**Match with Boxes Training:** Following probe sessions, match with boxes training was introduced. Materials present were the nine memory cards used during Condition 1 of baseline and the 3-box apparatus. The comparisons were situated in front of the boxes. Comparisons did not change positions within session, but did change across sessions. Error correction and reinforcement procedures were similar to those in baseline. When Chris placed the sample in the box corresponding to the correct comparison, praise was always delivered and tangibles were delivered intermittently. (The schedule of tangible delivery was not systematic.) Following an incorrect response, I modeled the correct response. That is, I took the sample from the box where the child had placed it and held it up, ensuring that the child’s eyes oriented towards the card. Then, I placed it in the box corresponding to the S+ and handed the card to the participant to complete the trial. If this occasioned an incorrect response, I repeated the model and asked the participant to match again. After the third incorrect response, I physically prompted the child to complete the trial and a new trial (with a new sample) was initiated.
2-Step Prompt + Response Blocking Strategy: One comparison was placed in front of each of the three boxes on the tabletop. I either gave Chris the sample or placed it within arm’s reach. If Chris placed it in the correct box, I provided descriptive praise and a tangible item. If he placed the sample on top of the correct comparison (Step 1), I provided praise and pointed to the box behind the comparison saying, “It goes here!” in an excited tone. If Chris placed the sample in the corresponding box (Step 2), then I provided praise and presented a tangible consequence.

If Chris placed the sample in the incorrect box, I blocked access to the boxes and re-presented the sample to him. Initially, I kept my arm over the box until Chris touched the sample to the correct comparison before removing my arm. However, as Chris’ accuracy increased, I required only card movement or eye orientation towards the correct comparison before lifting my arm. If I removed my arm and Chris either did not move the sample to the box or placed it in the incorrect box, I blocked the boxes again and pointed to the correct box. This completed the trial (ending on a prompted response) and a new sample was presented. Changing samples appeared to decrease emotional responding (e.g., non-assent, banging the table, shrieks) despite the absence of a tangible consequence.

During initial sessions of the condition, I immediately placed her arm over the box holds, blocking access until Chris had touched the sample to the comparison on the table (Step 1). I provided social praise and lifted my hand so that Chris could place it in the box to complete the trial (Step 2). As performance stabilized and number of corrects with the block increase, I faded the block prompt and then blocked the boxes only as an error correction strategy following an incorrect match. To reinforce persistence with the
task, tangible reinforcers were initially delivered following both prompted and correct responses. Across sessions, the schedule was thinned until tangible consequences were provided only following correct, unprompted matches. Appendix C represents a flowchart of a training session for 2-step prompt + response blocking.

Pretraining Results and Discussion

Figure 2. Closed circles represent the number of correct matches for Chris during the Pretraining Phase. Grey bars represent session duration in seconds. The number 3 above the Session 11 bar indicates the number of trials in that session due to lack of assent.

Figure 2 represents Chris’ data during pretraining following seven probe sessions. After four sessions of match with boxes training, correct matches decreased to zero and Chris did not assent to continue after the third trial. Therefore, the experimenter implemented a new strategy: 2-step prompt + response blocking. Four
sessions were required for Chris to meet the mastery criterion of 9 of 9 correct for two consecutive sessions. Although there were no programmed contingencies in place for decreased latency, there was also a decrease in session duration across the condition. The results suggest that the 2-step prompt + response blocking strategy was effective in training matching with the boxes.

**Non-match Training Procedures**

As in the 2-step + response blocking phase, a comparison was placed in front of each of the three boxes. The third comparison, however, was always a blank card that served as the non-match response. That is, if the sample was not identical to the comparison in the left or middle position, the blank card was the correct comparison (S+). Each comparison served as the S+ three times during a session. I either gave Chris the sample or placed it within arm’s reach. If Chris placed the sample in the corresponding box, I provided praise and presented a tangible consequence.

If Chris placed the sample in the incorrect box, I modeled a problem solving sequence. That is, I removed the card from the box and placed it on the comparison on the far right. If the sample and comparison were identical, I said, “Same; it goes here!” If they were not identical, I casually said, “No,” and placed the sample on top of the middle comparison. If the sample and middle comparison were identical, I said, “Same; it goes here!” If they were not identical, I moved the sample over the blank comparison before silently placing it in the corresponding box. Then, I took the card from the box and handed it to Chris to complete the trial. If Chris modeled the sequence and placed the sample in the correct box, praise was delivered. Tangibles were delivered during the first three sessions when Chris completed the motor component of the problem solving
sequence (whether prompted or independent). To increase fluency and reduce the likelihood that correct responses were under the control of praise or error correction on prior trials, I began fading praise during Session 5. All other procedural aspects remained constant throughout non-match training.

*Non-match Training Results and Discussion*

![Graph showing non-match training results](image)

*Figure 3.* Closed circles represent the number of correct matches. Grey bars represent session duration in seconds. Numbers above the bars indicate number of trials in a session for any session consisting of fewer than nine trials due to lack of assent.

Results of non-match training are depicted in Figure 3. Introduction of the non-match stimuli did not disrupt stimulus control established in the pretraining phase. That is, when a correct comparison was present, Chris placed the sample in the corresponding box 6 of 6 opportunities. However, during the first two sessions, when a sample corresponding to the non-match box was presented, Chris engaged in the non-
match response only 1 of 3 opportunities. After Session 4, when Chris responded at
100% accuracy, I started fading the praise and tangible consequences until, within two
sessions, it was delivered only at the end of the 9-trial session. The rationale for
thinning the schedule was that reinforcer consumption interrupted the session
unnecessarily, reinforcement and corrections could interfere with future probe stimuli,
and the breaks may act as an unintended source of intertrial stimulus control.

Williams, Johnston, and Saunders (2006) demonstrated a direct relation between
accuracy of delayed-MTS performance and increases in intertrial interval duration for
eight adults with mental disabilities. This finding could apply to the current study. For
example, if Chris watched a movie clip between some responses and not others the
stimulus control might differ across trials. If performance was analogous to that
described by Williams and colleagues (2006), Chris’ performance would be more
accurate after the presentation of a movie clip than after social praise delivery as a
function of the longer intertrial interval. In addition, any verbal feedback could have had
an effect other than maintaining responding in general (e.g., reinforcing a particular
stimulus control topography).

Despite rationales for eliminating praise, it is likely that the praise was faded too
quickly. This is demonstrated by the decreasing trend in corrects and the increasing
session duration across the condition. It may be prudent to develop a systematic
reinforcement fading program for mastered skills or to forgo the benefits gained from
withholding feedback in order to maintain responding to the task. Following 10 sessions
with no praise, and 3 days when Chris did not assent to the task at all, measures were
taken to reinstate stimulus control.
For the procedure to be used as an assessment method in Chris’ applied behavioral therapy program, it was imperative that he assent to the task and maintain correct responding. Several attempts were made to reinstate rates of correct responding that Chris had demonstrated when praise was delivered during the non-match training phase. First, the original rate of reinforcement was re-established. However, Chris still made errors or did not assent to the task at all. The pretraining task was re-introduced and still, no assent. Etzel, LeBlanc, Schilmoeller, and Stella (1981) reported that in their clinical experience “returning children to an easier task (or denser schedule) does not necessarily reduce inappropriate responses” (p. 9) and this observation was consistent with Chris’ responding. The Pooh cards that had been associated with the non-match task were replaced with ones that displayed different characters, Jay-Jay the Jet Plane and Arthur. Both characters were ones that Chris approached often when they were presented in other contexts (e.g., on TV, on t-shirts and toys). Unfortunately, Chris only assented to participate in 5 experimental sessions during this phase and correct responses continued to decrease. By this time, it was summer vacation and experimental sessions were terminated to accommodate the participant’s changing schedule.
**Baseline Procedures**

Baseline procedures implemented with Abby consisted of the minimal teaching procedure and did not differ from Chris except that the schedule of reinforcement was not thinned for Abby’s responding during baseline. Instead, the experimenter reminded her of the upcoming tangible after the second correct response and then delivered a tangible following the third correct match (prompted or not).

**Baseline Results and Discussion**

*Figure 4.* Closed circles represent the number of correct matches for Abby during Condition 1 of baseline. Open circles represent the number of correct matches for Abby during Condition 2 of baseline.

Abby showed no overlap between conditions and high variability both within and across conditions for the duration of baseline (see Figure 2). During the 15 sessions
conducted in Condition 1, correct responses ranged between 7 and 9, with a mean of 8.6 of 9 correct matches per session. Correct responses during Condition 2 were significantly lower, ranging between 0 and 8 correct responses per session with a mean of 5 across seventeen sessions. There was no overlap between Condition 1 and Condition 2 throughout baseline. Despite the gradually increasing trend in baseline, progress was inconsistent and slow enough to warrant clinical intervention. The pretraining phase had been successful with Chris, and so it was introduced with some modifications with Abby.

**Pretraining Procedures**

*Tracking Phase:* Only Abby participated in the tracking phase. Because she had a history of difficulty with fine motor tasks and was currently having vision problems, the experimenter implemented this condition to isolate the skill of putting the cards into small slots. The primary goal was to build rate of responding. To minimize the potential of a position bias in the subsequent condition, the experimenter touched her index finger to the box (often with such force to make a noise) in each of the three positions 3 times during the session.

Tangibles were delivered intermittently throughout the session until Abby had completed a session with 100% accuracy. The following session, the experimenter praised all correct responses and delivered a tangible only at the end of 9 trials. After 2 days at 100% accuracy, Abby began the match with boxes training. During the tracking condition, Bob the Builder Memory™ cards (HIT Entertainment PLC and Keith Chapman, [www.hasbro.com](http://www.hasbro.com)) and My Little Pony Memory™ cards (Milton Bradley
were used to avoid creating a history of placing the Pooh cards in boxes following a point prompt.

**Match with Boxes Training:** Following the tracking phase, match with boxes training was introduced. Procedures were similar to those employed with Chris during the match with boxes training described in Program 1 except that a point prompt was delivered during the first error correction because of Abby’s experience following the prompt during the tracking phase. Therefore, when Abby matched the sample to the box corresponding to the correct comparison, praise was always delivered and tangibles were delivered intermittently.

Following an incorrect response, I pointed to the correct box and handed the card to Abby to complete the trial. If incorrect, I modeled the correct response and asked the participant to match again. After the third incorrect response, I physically prompted the child to complete the trial and a new trial (with a new sample) was initiated.

**2-Step Prompt + Response Blocking Strategy:** This procedure was identical to the 2-step prompt + response blocking strategy used in Chris’ pretraining phase. The materials used during match with boxes training were also used during the 2-step prompt + response blocking strategy phase (i.e., memory cards and the 3-box apparatus). A comparison stimulus was placed in front of each of the three boxes on the tabletop. I either gave Abby the sample or placed it within arm’s reach. During initial sessions of the condition, I immediately placed my arm over the box holes, blocking access until Abby touched the sample to the comparison on the table (Step 1). I provided social praise and lifted my hand so that Abby could place it in the box to complete the trial (Step 2). As performance stabilized and number of corrects with the
block increased, I faded the block prompt within sessions. Then, I blocked the boxes only as an error correction strategy following the incorrect match. To reinforce persistence with the task, tangible consequences were initially delivered following both prompted and correct responses. Later, if Abby placed the sample in the correct box, I provided descriptive praise and a tangible item. If Abby placed the sample on top of the correct comparison (Step 1), I provided praise and pointed to the box behind the comparison saying, “It goes here!” If Abby placed the sample in the corresponding box (Step 2) then I provided praise and a tangible consequence.

If Abby placed the sample in the incorrect box, I blocked access to the boxes and re-presented the sample. This completed the trial (ending on a prompted response) and a new sample was provided. Changing samples appeared to decrease emotional responding despite the absence of a tangible consequence.

Across sessions, the schedule was thinned until tangible consequences were provided only following correct, unprompted matches. Appendix B represents a flowchart of a training session for 2-step prompt + response blocking.
Pretraining Results and Discussion

Figure 5. Closed circles represent the number of correct matches. Grey bars represent session duration in seconds. Numbers above the bars indicate number of trials in a session for all sessions consisting of fewer than nine trials due to lack of assent.

Figure 5 shows data for the pretraining phase with the two-dimensional memory cards. Abby rapidly learned to place a card in the box to which the experimenter was pointing during the tracking condition. After Abby had met criterion (2 days at 100% accuracy) on this condition, she underwent an eye operation to correct her hypertropia which may, or may not, have affected responding in future sessions. It is also possible that the tracking condition was actually detrimental to responding in subsequent conditions because the introduction of the box was paired with an extra-stimulus, or non-criterion related, point prompt (see Schreibman, Charlop, & Koegel, 1982 and Koegel & Rincover, 1976 for demonstrations of detrimental effects of extra-stimulus prompts).
Following the tracking phase, match with boxes training was introduced. Correct responding was initially low followed by a gradually increasing trend that then decreased to zero levels. Introduction of the 2-step prompt + response blocking phase quickly produced 10 sessions (i.e., sessions 20-29) of responding that ranged between 6 and 9 correct responses per session. However, analysis of the raw data demonstrated that the variability within this range was not a result of clear, experimenter-unintended stimulus control (e.g., position or incorrect responding in the presence of one particular sample). Sessions 30-33 showed an upward trend. Therefore, the experimenter could have continued running this phase if data had remained high and stable. However, following Session 33, Abby handed an item back to me during a therapy session when the match was accidentally not present. This prompted me to try shaping the non-match response already in Abby’s repertoire (handing an item back) into putting the item in an alternative box. The results of this shaping procedure are described in the results and discussion section for non-match training.

**Non-Match Training Procedures**

A serendipitous event prompted an alternative training procedure for Abby than was effective for Chris. While Abby was still in the pretraining phase, I was conducting a maintenance 3D-matching task during a therapy session. When Abby picked up an item that was not part of the matching field, she paused over the stimuli and shook her head, “no.” Though the typography was different than the intended trained response and materials were different than the ones currently used for pretraining, there did appear to be a non-match response in Abby’s repertoire. Therefore, a stimulus fading procedure was designed to gradually transform the context in which she was currently successful
into the training context that had worked to establish the non-match response with 2D stimuli with Chris.

During Step 1, the experimenter placed a large box (the non-match container) used for storing all the 3D matching materials beside four small bowls, each containing a 3D object. If Abby placed an object in the bowl that was physically identical to the sample, I delivered praise on every trial and a tangible intermittently. If Abby did not complete a match response, I returned the item to Abby and said, “Try again,” or “Oops, not that one.” When I gave Abby an item that was a non-match, often though not always, Abby held the sample above each comparison before pausing with it in the air. At this point, I pointed to the large storage box. If Abby placed it in the box, she received a tangible consequence. If she did not place it in the box, I restarted the trial with the same sample and immediately pointed to the box.

After non-matching was established under these conditions, Step 2 was implemented whereby the type of container used for the non-match response varied, a blue toy grocery bin and then a red plate was used. The red plate (and later a red box) was chosen because the color red is often associated with rejection/no responses (e.g., a red X, stop light, Do not enter sign) and might therefore be beneficial if the contextual properties of this stimulus generalized to other events in Abby’s life.

During Step 3, I gradually faded out the 3D toys and replaced them with the 9 2D Pooh Memory cards. Initially, the field consisted of two 2D cards and two 3D toys. When accuracy was 100% for two sessions, I removed the 3D stimuli and the bowls until there was just a field of two cards and the big, red, plate acting as the non-match comparison.
Following four sessions with little success (while maintaining perfect performance in the 3D condition described as Step 2), we discontinued Step 3 and implemented Step 4.

During Step 4, I replaced the red plate with a 5.1 X 5.1 inch slotted square box covered with red construction paper and added the 2 beige boxes of the same size behind the left and middle comparisons. I also replaced the Pooh Memory cards (that all had a blue background and a reinforcement history established throughout the experiment) with traditional Memory Game™ (Milton Bradley Company, www.hasbro.com) cards. The cards had white backgrounds and pictures considered less complex (e.g., a hat, a ball, and a bird). Stimuli were always chosen from among a set of 20 cards, randomly arranged except that stimuli of similar color or category were not included in the same comparison array (e.g., a red ball was not placed in the same field as a red bird and a horse was not placed in the same field as a cow.) This step was taken to minimize the possibility that errors in previous sessions were due to either stimulus complexity or the common background interfering with the acquisition of the appropriate stimulus control.
Non-match Training Results and Discussion

Figure 6. Close circles represent the number of correct matches. Grey bars represent session duration in seconds.

Figure 6 depicts Abby's responding during non-match training. (Duration data were not collected during sessions 9-11 and 17-23 due to technical difficulties.) There was an increasing trend during Step 1 and she met mastery criterion (2 consecutive sessions at 100% accuracy) after seven sessions. Correct matches did not decrease significantly during Step 2 when the non-match apparatus, the bin, was switched to a red plate and Abby met mastery criterion after five sessions. The purpose of Step 3 was to gradually remove some of the 3D stimuli and replace it with the 2D stimuli used during baseline and pretraining phases. First, two 3D comparisons and two 2D comparisons were presented. After three sessions, Abby had responded correctly (i.e., matching a sample to its corresponding comparison if it was present, and placing the sample on the red plate if the S+ comparison was absent) on 100% of trials for two
consecutive sessions. Therefore, a 2D comparison was added and a 3D comparison was removed from the field. Responding did not improve as rapidly as it had previously. After four sessions correct responses decreased from six to five and so Step 4 was implemented. During Step 4, less complex memory cards were used. Responding remained variable throughout this phase and while there was an upward trend at the end of the phase, correct responding decreased to 4 of 9 trials when the final session was conducted.
GENERAL DISCUSSION

The goal of this study was to examine the efficacy of a technology that teaches children with autism and limited verbal skills a procedure to respond to an apparatus that can reveal types of stimulus control. One stimulus shaping procedure (i.e., 2-step response blocking) produced discriminated responding for one child, though responding did not maintain when reinforcement and feedback was removed. An adaptation of the stimulus control shaping procedure was partially successful in that it produced discriminated responding with 3D objects, but not with the original 2D stimuli, for a second child.

Utility of the Current Technology

The data sheet was designed to reveal, upon visual inspection, some common controlling stimulus-response relations (e.g., control by position, proactive interference, or control by character or feature). Also, a sample could be matched to an incorrect comparison dependent upon the specific features of the S- or upon the serial ordering of the comparisons. SCTs may change within session following events such as reinforcer delivery and error correction; these events can be monitored easily with the data sheet.

While accuracy and duration were the primary dependent variables graphed and analyzed in this study, the data sheet provides additional information that can better direct teacher behavior. For example, during the first sessions of the non-match training condition, Chris’ accuracy decreased from 9 of 9 to 6 of 9 trials. The data sheet shows that Chris was continuing to respond when the S+ was present and was responding incorrectly when a comparison that was physically similar to the sample was not present (i.e., the presence of the non-matching box did not disrupt the stimulus-response
relations established during the pretraining phase, but the box did not immediately exert control over the non-match response).

In short, by accepting a quantal interpretation of stimulus control, teachers of children with autism may be more likely to conduct within session analyses rather than assuming a lack of stimulus control or attributing stimulus control patterns to the child (e.g., labeling the child prompt dependent or describing a pattern of responding as a “position preference”).

**The Use of Boxes as Instructional Stimuli**

Lowenkron (1989) conditioned different colored backgrounds as instructional stimuli that signaled whether oddity or identity matching would result in reinforcement. However, he used verbal behavior to establish instructional control with the 4-year-old children and this strategy is not an option when teaching children with minimal verbal skills. Fortunately, instructions need not be verbal. The current study’s apparatus allows matching based on *all* relevant features of the stimuli rather than a specific stimulus dimension, as when matching based on color. With a verbal child, one could use an instruction; “Match if exactly the same.” For a non-verbal child, the presence of the red non-matching box serves as the instructional stimulus to match, “same.” That is, all stimulus dimensions of the S+ must be identical to the stimulus dimensions of the S-. Absence of the red box is an instructional stimulus to match, “similar.” That is, at least one stimulus dimensions (e.g., color or character) must be common to the S+ and the S-.

The current study designed the pre-training phase to establish the boxes (and the red box in particular) as instructional stimuli that indicated identity matching would
result in reinforcement. This is important because the behavior of children with autism has shown to be influenced, in part, by recent behavioral history (Garcia, 1994). Clear, novel instructional cues may minimize the likelihood that the child will generalize a previous history of matching by specific dimensions of the stimulus (e.g., color or category) to all contexts.

The Use of a Tabletop Method Instead of a Computer

The current tabletop apparatus may have more utility than the computer apparatus described in other studies (e.g., Gutowski & Stromer, 2003; Serna et al., 1993). First and foremost, a tabletop procedure can be adapted quickly and easily as participant responding shapes the developing technology. In the current study, the experimenter varied the materials while attempting to gain assent during the reinstatement of stimulus control phase. Jay-Jay the Jet Plane™ and Arthur™ character cards were introduced during different sessions on the same day and neither character resulted in increased attending to stimuli or sitting at the table. Displaying the characters on the computer may have added days to the stimuli probe because of the high response effort involved in programming.

There are of course benefits to using a computer preparation. Computers can be programmed to automatically measure multiple dimensions of behavior (e.g., latency, rate, duration), and randomize and counterbalance trials (Saunders & Williams, 1998). They also eliminate the potential of inadvertent cueing by the teacher (Augustson & Dougher, 1991). In addition, the child may enjoy manipulating the computer, which may increase motivation towards the task. However, the skill of responding conditionally to one dimension within a complex stimulus is more likely to generalize after training within
a number of different teaching situations. The priority for teachers and autism interventionists is that there are a number of empirically tested options for teaching a skill at a table or on the computer. Despite these positive benefits, when determining potential teaching methods, it may be most productive to probe with stimuli at hand. In other words, one needs to know what to teach before programming the computer to teach it.

Limitations of the Current Investigation

There are several obvious limitations of the current study. First, because the procedures were individualized to each participant, the modified AB design does not demonstrate control as intended. However, given that Abby was 8 years old and had a limited repertoire, it is unlikely that some variable other than the procedures were responsible for skill acquisition. In addition, the correlation between removal of reinforcement and feedback, and increased task duration and decreased session accuracy, provides support for the assertion that the procedures were responsible for both the acquisition, and the breakdown, in Chris’ responding.

Second, the present analysis tells us only that the described methods are adequate to teach the skill; they do not tell us (a) if all steps are necessary, (b) if the arbitrary mastery condition imposed for progression to the next step impeded progress, or (c) if a stricter criterion might have prevented the breakdown of stimulus control seen in Chris’ data. Future studies should analyze the components to determine the most efficient way to train the skill.

Another limitation of the study is that Abby’s program was only partially successful in that she learned a non-match response with 3D objects but did not
engage in a non-match response with 2D objects at the study’s conclusion. Reasons for the lack of success need to be investigated. One potential explanation for Abby’s lack of success with 2D objects relates to her visual deficits. However, other explanations should be considered. Serna, Dube, and McIlvane (1997) describe tabletop IDMTS tasks with 3D objects as a preceding, if not requisite step, to 2D matching in their program designed to train generalized identity matching. It is possible that Abby was missing unidentified requisites related to the 3D task that interfered with acquisition of the non-matching response with 2D items.

Of pragmatic importance, the program took significantly longer to complete than the BC-MTS procedure (Serna et al., 1998), which took 1-2 sessions per participant. Perhaps the computer procedure could be implemented as a primer, prior to training the tabletop method with stimuli of more relevance to the individual.

Conclusion

Despite the presented limitations, the results suggest a potential avenue of research. This study focused on developing a technology that taught children with autism to complete the assessment procedure. Future studies should examine the SCTs revealed by the procedure and then apply the assessment information to train appropriate stimulus control.

In conclusion, replication with additional participants and stimuli is required to draw conclusions about the efficacy of the 2-step prompt + response blocking strategy as a method of teaching non-matching to children with autism who have a limited verbal repertoire. However, this study suggests a promising teaching method that establishes a
way of revealing stimulus control that could be applied to children who display overselective responding, the source of which is difficult to identify.
APPENDIX A

STIMULI
SCORING PROTOCOL FOR NON-MATCHING PROGRAM

GENERAL INSTRUCTIONS

Score correct/incorrect data for the entire session. Then, score duration (fast-forwarding through reinforcement periods if possible). The data sheet is divided into 9 trials. Three boxes for each trial indicate position of the stimuli (e.g., left, middle, right). Record the stimulus name (e.g., OS, PF or PigC) in each position on every trial under each box or card position (L = left, M = middle, R = right).

CORRECT/INCORRECT MATCH RESPONSE

Baseline

During baseline, score a + in the corresponding box on the data sheet when the participant touches the sample in her hand to the identical comparison on the table. Score the trial as incorrect (-) if the participant touches the sample to any position other than the correct comparison, or if the experimenter intervenes in any way EXCEPT verbally.

Pretraining and Non-match Training Phases

During the pretraining and non-match training phases, place a + in the corresponding box on the data sheet when the participant places the sample at least half way in the box or on the plate that corresponds to the correct comparison. Place an E (i.e., experimenter extended the participant’s response) in the corresponding box on the data sheet if the participant matches the sample to the correct comparison independently but puts it in the corresponding box following an experimenter prompt. Score the trial as incorrect if the participant places the sample anywhere that is not a correct response or an E.
Examples of Correct Match:
The participant takes the card from the experimenter and places it in the pile on the floor, the experimenter picks up the card and gives it back, saying, "match," and the participant places the sample on the S+.

Examples of Incorrect Match:
The participant takes the card from the experimenter and places it in the pile on the floor, the experimenter picks up the card and gives it back, POINTS TO THE CORRECT COMPARISON, saying, “match,” and the participant places the sample on the S+.

DURATION OF MATCH AND NON-MATCH RESPONSES

Baseline
Drag the arrow under the QuickTime movie under the frame slowly until it shows the first time the child’s hand contacts the card. This will be the beginning of the first trial of the session. Record the time code ONLY in seconds and drag the arrow until the participant completes the trial. Record the time code in seconds as it appears on the movie file. Do not round up or down; this is a gross measure of time. Start timing when the participant’s hand contacts the memory card. Stop timing when the card in the participant’s hand touches the S+ card on the table. For incorrect matches, record the duration starting when the participant first touches the card and continue timing until s/he has touched the card to the S+ (this includes when the trainer is touching the card, when the card is on the table, etc). If the participant walks away in the middle of a trial, stop timing when s/he is 3 feet from the table or out of the frame whichever occurs first.

Note: If the experimenter and participant’s hands are not in the frame, look for a
shadow on the carpet. Start timing when the two shadowed hands are not separated by space. If you do not see the shadow, start recording 1 s after the card leaves the frame and continue timing until the participant places the sample on the correct comparison.

Exceptions due to filming quality:
If the participant’s hand is out of the frame at the point that the trial begins, start timing the second that his hand has left the frame.
If you cannot see either the beginning or completion of the trial (e.g., because the experimenter’s hand is covering the participant’s hand or the participant's hand is not in the frame, record the time at the first frame that you can see that the item that was not in view.

Session duration
Total the time for 9 trials (fewer than 9 if the participant leaves the session and does not return that day).

*Pretraining and Non-match Training Phases*
The only difference between baseline and these phases is that the observer stops timing when there is no contact between the card and the participant’s hand and at least half of the card is in the box that corresponds to the S+.

Examples of Inclusions:
The participant takes the card from the experimenter and drops it on the floor then completes the trial.
The participant touches the card in the experimenter’s hand and then signs, “movie,”
The participant touches the card that the experimenter has placed on the table, then grabs a piece of popcorn and eats it before completing the trial.
Examples of exclusions:
Following the verbal direction, “Match,” and before contact with the card, the participant signs the word, “movie.”

**Interobserver Agreement (IOA) for Correct/Incorrect Match**

Two observers will be scoring the student’s performance independently to establish the percentage of agreement between observers. Upon completion of data collection, IOA data will be compared trial by trial. When both observers score the trial as correct, agreement is scored. When both observers score the target behavior as incorrect, agreement is scored. When one of the observers scores the target behavior as occurring and the other as non-occurring or when one of the observers score the trial as correct and the other scores it as incorrect, disagreement is scored. Add the number of agreements, divide by the total number of agreements and disagreements and multiply by 100.

If the score is under 85% during more than one non-consecutive session, the primary investigator will intervene and make changes (e.g., to the behavioral definition, re-train the observer, etc).
Client came to table □ after only sorting stimuli was presented □ with neutral physical guidance
(record order of events) □ after reinforcer was established □ after verbal instruction

<table>
<thead>
<tr>
<th>Trial</th>
<th>L</th>
<th>M</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trial 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trial 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trial 3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trial 4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trial 5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trial 6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trial 7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trial 8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trial 9</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

# of prompts | Duration of Match Response

Session Length: __________________ (Add together total duration of match responses)
Total Correct: __________________
Total Incorrect: __________________
# of Corrections: __________________
# of times child left table: __________________
APPENDIX C

FLOWCHARTS OF BASELINE
AND 2-STEP PROMPT + RESPONSE
BLOCKING PRETRAINING
MINIMAL TEACHING PROCEDURE SESSION

Child sits in front of field to start

Give sample to child

NO

Match correct

NO

3rd incorrect

Model Correct Match

NO

Eyes track

Deliver social consequence and descriptive praise

HOH physical prompt

Field of 3 complete

Tangible Consequence

YES

3rd field complete

End session

YES

all 3 matches correct

Rearrange field

Select new sample
REFERENCES


