Are Symmetric and Generalized Matching-to-Sample Skills Associated with Picture Preference Assessments for People with Developmental Disabilities?

by

Leslie Maxine Elizabeth Thorne

A Thesis submitted to the Faculty of Graduate Studies of

The University of Manitoba

in partial fulfilment of the requirements of the degree of

DOCTOR OF PHILOSOPHY

Department of Psychology

University of Manitoba

Winnipeg

Acknowledgements

I express my sincere gratitude to the participants and staff at St. Amant without whom, this research would not be possible. I express my appreciation to Dr. C.T. Yu for his support and guidance in this study; to Drs. Garry Martin, Angela Cornick, and Dennis Hrycaiko for their support and recommendations. Finally, I thank Carly Thiessen, Alejandra Zaragoza Scherman, Quinn Senkow, Tiffany Pang, and Brigitte Jeanson for their help with data collection.

Abstract

When assessing preferences of individuals with developmental disabilities, choices can be described vocally or presented using objects or pictures in preference assessments. For individuals who are unable to perform auditory-visual discriminations and visual identity matching, objects instead of pictures or vocalizations need to be used for preference assessments to be effective. Considering the practical advantages of using pictures over objects, recent research has begun to focus on identifying and teaching skills needed for picture preference assessments. Although object-to-picture, picture-to-object, and generalized matching have been implicated as possible skills needed for picture preference assessments, further systematic studies are needed. The present study examined the relation between preference assessments (object vs. picture groups) and 5 discrimination skills. Based on direct paired-stimulus preference assessments completed at the beginning of the study, participants who could indicate their preferences with objects, but not with picture or vocal presentation, were assigned to the Object Group (n = 11); and participants who could indicate their preferences with both objects and pictures, but not vocal presentation, were assigned to the Picture Group (n = 9). The 5 discrimination tasks included: (a) object-picture matching and (b) its symmetry, picture-object matching; (c) generalized object-picture matching and (d) its symmetry, generalized picture-object matching; and (e) generalized identity picture-picture matching. All task stimuli were parts from everyday objects. Independent sample t-tests with Bonferroni correction showed that the percentages of correct responses were significantly higher in the Picture Group than in the Object Group on 4 of the 5 tasks ($p <$ .01). Individual data showed that 8 of the 9 Picture Group participants met the pass

criterion (80% or higher correct responses) on at least 1 discrimination task, with 6 participants passing 2 or more tasks. In contrast, only 1 of the 11 Object Group participants met the pass criterion on 1 discrimination task. The findings suggest that the effectiveness of picture preference assessments is not dependent on one specific discrimination, but possibly the ability to perform generalized matching.

Table of Contents

List of Tables

List of Figures

Are Symmetric and Generalized Matching-to-Sample Skills Associated with Picture Preference Assessments for People with Developmental Disabilities?

Making preferred items more available in the immediate environment allows individuals with developmental disabilities to feel happier, and ultimately their quality of life is improved (Green, Gardner, & Reid, 1997; Wehmeyer & Schwartz, 1997). Access to preferred items also reduces challenging behavior (Vollmer, Marcus, & LeBlanc, 1994) and enhances acquisition of functional skills (Green et al., 1988). Individuals with severe and profound developmental disabilities typically have limited speech and are unable to express their preferences. However, considerable research has shown that direct preference assessment is a reliable method for identifying preferred stimuli for this population (Hagopian, Long, & Rush, 2004). When assessing preferences, choices can be described vocally or presented using objects or pictures. Several studies have shown that an individual's ability to perform visual and auditory discriminations predicted which modes would be more effective. For individuals who are unable to perform auditory-visual discriminations and partial-identity visual matching, objects instead of pictures or vocal descriptions need to be used for preference assessments to be effective. Considering the practical advantages of using pictures over objects, recent research has examined the discrimination skills required in picture preference assessments and for teaching individuals in order to enable them to respond to pictures during preference assessments. Several conditional discriminations such as matching object-to-picture, picture-to-object, and generalized matching have been implicated as possible prerequisites to respond to picture preference assessments. The present study extended this line of research by examining whether several discriminations were associated with

picture preference assessments in a group design. In the ensuing sections, I define preference assessment, review research on several preference assessment procedural parameters, and present the purpose of this study. This is followed by a description of the study's methodology, findings, and a discussion.

Defining Preference Assessment and Effectiveness

Martin, Yu, Martin, and Fazzio (2006) examined the concepts of choice and preference and defined preference as "the relative strength of discriminated operants" that is measured based on a pattern of choosing. During a preference assessment, a variety of items are systematically presented to an individual. The presenter then observes the individual's approach response to each item (e.g., pointing to or touching an item) and whatever item is picked during the presentation is given to the individual. Once each item has been presented a number of times, the percentage of trials an item was chosen from all trials in which the item was available reflects the extent of preference for that item relative to the other items. A preference assessment might be considered effective if: (a) it successfully differentiates preferences among an array of stimuli and minimally identifies a preferred stimulus and (b) the most preferred stimulus is a positive reinforcer. A highly preferred stimulus is one that has been selected on at least 80% of the trials (Pace, Ivancic, Edwards, Iwata, & Page, 1985). A number of studies have shown that highly preferred items identified through direct preference assessments also function as positive reinforcers (e.g., Higbee, Carr, & Harrison, 2000; Horrocks & Higbee, 2008; Lee, Yu, Martin, & Martin, 2010).

Number of Choices and Preference Assessment Effectiveness

When conducting preference assessments, three methods of presentation of

stimuli are commonly used. The single-stimulus procedure involves presenting each stimulus separately and allowing the participant to approach the stimulus (Pace et al., 1985). In the paired-stimulus procedure, stimuli are presented in pairs on each trial and the participant selects one of the two stimuli (Fisher et al., 1992). With the multiple-stimulus presentation procedure, the participant is presented with usually 6 to 8 stimuli simultaneously on each trial and is instructed to select one (Windsor, Piche, & Locke, 1994). In the multiple-stimulus without replacement procedure (Bojak & Carr, 1999; Carr, Nicolson, & Higbee, 2000; Daly et al., 2009; DeLeon & Iwata, 1996), a stimulus is removed from the array once it has been selected and is not presented in subsequent trials.

Overall, each of the preference assessment procedures has been shown to be effective in identifying a preference hierarchy for individuals with developmental disabilities (Cannella, O'Reilly, & Lancioni, 2005) or behavioral disorders (Paramore & Higbee, 2005). The single-stimulus method is useful for individuals who have difficulty making two-choice discriminations. However, it is more time consuming than the other procedures and participants have a tendency to approach all stimuli. The paired-stimulus presentation is more sensitive in differentiating preferences than the single-stimulus presentation method and takes considerably less time. The multiple-stimulus without replacement procedure is as effective as the paired-stimulus procedure and requires less time (DeLeon & Iwata, 1996; Higbee et al., 2000). However, the multiple-stimulus procedure requires a client to scan and discriminate 6 to 8 stimuli, which may be a challenge for some individuals with severe and profound intellectual disabilities and probably for these reasons, the paired-stimulus procedure remains the most commonly

used in preference assessment research with this population.

Modes of Choices and Preference Assessment Effectiveness

Research has shown that stimulus modes (objects, pictorial, and vocal presentations) interact with discrimination skills to affect preference assessment effectiveness. Higbee, Carr, and Harrison (1999) conducted one of the first comparisons between picture and object preference assessments with persons with a diagnosis of mental retardation. Preference assessment using objects resulted in better preference differentiation than pictures among 7 stimuli across 2 adults with moderate and severe mental retardation. Conyers et al. (2002) compared the three modes with individuals with different discrimination skills. They hypothesized that in order to respond effectively to an object preference assessment, a simple visual discrimination is required. In contrast, in order to respond effectively to a picture preference assessment, a visual conditional discrimination (relating pictures to objects) is required. Lastly, with vocal preference assessments, an auditory-visual conditional discrimination (relating the vocal descriptions to objects) is necessary. Therefore, individuals who were able to make only simple visual discriminations should be able to choose their preferred items when objects were presented; individuals who were able to make simple visual and visual conditional discriminations should be able to choose their preferred items when either objects or pictures were presented; and individuals who were able to make all three discriminations should be able to indicate their preferences using any of the stimulus modes. In this study, nine participants were selected based on their ability to learn discrimination skills as measured by the *Assessment of Basic Learning Abilities* (ABLA, Kerr, Meyerson, & Flora, 1977; Martin & Yu, 2000). The ABLA measures the ease with which a person

learns to perform several visual and auditory discriminations that are hierarchical in difficulty. Using standardized prompting and reinforcement procedures, a tester attempts to teach each discrimination, individually, until a pass or fail criterion on that discrimination is met, whichever comes first. Three discrimination tasks (referred to as levels) of the ABLA were of interest in the study, specifically Level 3 (a two-choice visual discrimination), Level 4 (a two-choice matching-to-sample discrimination) and Level 6 (a two-choice auditory-visual discrimination). The results confirmed their hypotheses. Participants who passed ABLA Level 3, but failed Levels 4 and 6, selected their preferred items frequently (above 50% chance) when the choices were objects, but their selection of the same preferred items decreased to chance level when the choices were presented in pictures or using words. Participants who passed ABLA Levels 3 and 4, but failed Level 6, selected their preferred items above chance in the object and picture modes, but their preferred item selections decreased to chance in the vocal mode. Lastly, participants who passed ABLA Level 6 selected their preferred items above chance in all three modes.

de Vries et al. (2005) replicated Conyers et al.'s (2002) study with choices involving leisure activities (e.g., a large ball to represent ball playing, a CD player and CD cases to represent listening to music, etc.). Their results showed that the three ABLA levels predicted the consistency with which the participants chose their preferred leisure activities in each of the three modes.

The above findings were partially replicated by Reyer and Sturmey (2006), who examined whether the ability to learn discrimination skills, as assessed by the ABLA, could predict the effectiveness of different presentation modes for assessing preference of

work tasks. Nine participants were selected based on their ABLA discrimination levels and each participant was given a preference assessment for work tasks. Tasks that were identified as high- and low-preference were used to evaluate the three presentation modes. For 3 of the 9 participants, choice of task was predicted by their discrimination skills. The same results were observed for 2 additional participants for all trials but one.

Discrimination Skills and Picture Preference Assessment Effectiveness

Preference assessments using objects can be impractical. Describing choices vocally can be quite efficient, but it requires successive auditory discriminations, which can be challenging for many individuals with severe and profound developmental disabilities. Using pictures in preference assessments is more practical than objects and the simultaneous visual discriminations required for picture preference assessment are less challenging than auditory discriminations (Kerr et al., 1977). The above studies showed that individuals who learned to perform ABLA Level 4 (partial-identity visual matching) were able to choose their preferred items in picture preference assessments, whereas those who learned to perform ABLA Level 3 (simple visual discrimination) could choose their preferred items only in object preference assessments. These results suggest that partial-identity visual matching may be an important discrimination needed for picture preference assessments.

Clevenger and Graff (2005) examined performance on picture preference assessments and whether individuals could match objects to pictures. Picture-to-object and object-to-picture matching skills were evaluated with 6 participants and 3 of them were able to match objects to pictures and pictures to objects (mean accuracy 94%), while the other 3 were unable to do so (mean accuracy 29%). The results indicated that

participants who were able to match objects to pictures and pictures to objects showed high correspondence between object and picture preference assessments (i.e., they selected the same preferred stimuli in both assessments), whereas participants who did not demonstrate those matching skills showed low correspondence between object and picture preference assessments. However, it was not clear whether the two matching relations (object-to-picture and picture-to-object) were symmetric (i.e., involved the same stimuli).

Nguyen et al. (2009) taught object-picture matching to individuals for whom picture preference assessments were ineffective, and evaluated whether picture preference assessment effectiveness improved following training in a multiple-baseline design across training tasks. Three participants with developmental disabilities who showed preferences during assessments with objects, but not with pictures, were taught object-to-picture matching tasks unrelated to the items used during preference assessments. The authors hypothesized that participants who showed differential preferences during assessments with objects, but not with pictures of the same objects, would improve their concordance between the two modes after being taught object-picture matching. The results showed that concordance improved for two participants after mastering 2 and 3 matching tasks, respectively. However, concordance did not improve for the third participant after mastering two tasks and after additional training. Nguyen et al. suggested that training additional discrimination tasks might be necessary for responding to picture preference assessments. It is possible that "generalized matching", the ability to match new stimuli without training (Pear, 2001), may be necessary. Nguyen et al.'s study, however, did not assess this skill.

Statement of the Problem

Research has shown that the ability to learn to perform partial-identity visual matching predicted the effectiveness of picture preference assessments (Conyers et al., 2002; de Vries et al., 2005). Clevenger and Graff (2005) suggested that both object-picture and picture-object matching might be important skills for picture preference assessments. Nguyen et al. (2008) suggested that generalized matching may be necessary. A systematic study of the various matching relations and picture preference assessments could increase our understanding of the prerequisite skills necessary for successfully performing picture preference assessments, and help practitioners target important skills for training.

The purpose of this study was to examine the relations between picture preference assessments and 5 conditional discriminations. Specifically, I studied two groups of participants: one group could perform both object and picture preference assessments, but not vocal preference assessments (Picture Group) and the other group could perform object preference assessments only (Object Group). All participants were assessed on the following discriminations: (a) object-picture matching and (b) its symmetry, picture-object matching; (c) generalized object-picture matching and (d) its symmetry, generalized picture-object matching; and (e) generalized identity picture-picture matching. I hypothesized that the Picture Group would perform significantly better than the Object Group on all discrimination tasks. Ethical approval for this research was received from the University of Manitoba Psychology/Social Research Ethics Board before the study began.

Method

Participants and Settings

Twenty adults with developmental disabilities were recruited and written informed consent to participate was obtained from their legal guardians or substitute decision makers. The Object Group (n = 11) consisted of 6 males and 5 females, with a mean age of 39.6 years (range, 24 to 53). According to their health records, 2 participants were diagnosed with profound developmental disabilities, and 7 with severe developmental disabilities. The remaining 2 participants in the Object Group were diagnosed with developmental disabilities, but no information on level of functioning was indicated in their health records. The Picture Group (n = 9) included 6 males and 3 females with a mean age of 36.7 years (range, 31 to 42). Two participants were diagnosed with profound developmental disabilities and 5 with severe developmental disabilities. The remaining 2 participants in the Picture Group were diagnosed with developmental disabilities, but no information on level of functioning was indicated in their health records (see Table 1).

Participants were assigned to either the Object or Picture group based on the results of the paired-stimulus preference assessments with food items in object, picture, and vocal modes (procedure described later). To be assigned to the Object Group, a participant must have selected a high preference (HP) food item on at least 80% of the preference assessment trials with objects, and must have selected the same HP item on less than 65% of the assessment trials in picture and in vocal modes. All 11 participants in the Object Group met this criterion. To be assigned to the Picture Group, a participant must have selected the same HP food item on at least 80% of the preference assessment

trials in both object and picture modes, and must have selected the same HP item on less than 65% of the assessment trials in the vocal mode. All 9 participants in the Picture Group met this criterion (see Table 2).

Table 1. *Participant Characteristics*

Participants	Sex	Age (Yrs)	Level of Functioning	ABLA Level [a]
Object Group				
1	F	38	Severe	3
2	F	42	Profound	2
3	M	38	Severe	2
4	M	36	DD [b]	3
5	M	24	Severe	3
6	M	35	DD [b]	2
7	M	47	Severe	3
8	F	47	Severe	2
9	F	44	Severe	3
10	F	53	Severe	2
11	M	31	Profound	2
Picture Group				
12	F	39	Severe	4
13	M	36	Severe	3
14	M	36	Profound	2
15	M	31	DD [b]	2
16	M	42	Severe	4
17	M	36	DD [b]	4
18	F	40	Severe	2
19	F	39	Severe	4
20	M	31	Profound	3

[a] Highest ABLA level passed. [b] Developmental disabilities with no information on functioning level.

I administered the ABLA (procedure described later) to characterize each participant's ability to learn basic discriminations. As shown in Table 1, all Object Group participants learned to perform (pass) ABLA Level 2, a 2-choice visual discrimination

that involves position and visual cues, and 5 of the 11 participants also passed Level 3, a 2-choice visual discrimination. Participants who passed Level 2 have been shown to be able to perform paired-stimulus preference assessments with objects (Thomson, Czarnecki, Martin, Yu, & Martin, 2007). None of the Object Group participants passed Level 4 (quasi-identity visual matching) or higher. For the Picture Group, all 9 participants passed ABLA Level 2, 2 of the 9 participants also passed Level 3, and 4 of 9 participants also passed Level 4.

Throughout the study, participants were assessed individually and all sessions took place in a testing room at the St. Amant Research Centre. Participants sat behind a table, in a chair, across from the experimenter during all assessments. During some sessions, an observer was present to conduct reliability checks.

Paired-Stimulus Preference Assessments for Group Assignment

A paired-stimulus preference assessment was completed for each participant using 6 food items (objects). The procedure was deemed effective if it identified a high-preference (HP) and a low-preference (LP) food item. A high-preference item was one that was selected on at least 80% of the trials, whereas a low-preference item was selected on no more than 20% of the trials (Pace et al., 1985). Parents or caregivers were asked to nominate a list of food items for each participant and items selected for the preference assessment were considered based on ease of presentation and availability.

The paired-stimulus preference assessment procedure involved presenting two stimuli concurrently on each trial. Each stimulus was paired with every other stimulus twice and the order of presentation was randomized. Each participant received 30 trials to complete the preference assessment for the 6 food items. The two items were presented at

an equal distance from the participant and the left-right positions were counterbalanced across trials. The participant was prompted to look at each item and then asked to "pick one". An *approach* response was defined as the participant touching or pointing to an item without rejecting it within 8 s after being asked to choose. A *rejection* was recorded if the participant pushed an item away. After a rejection, an approach to the other available item was permitted on the same trial. On each trial, the item selected was recorded. If neither item was selected after 8 s, the trial was recorded as *no selection*. Immediately following an approach response, the participant was thanked for choosing and provided with the item selected. If a participant approached both items simultaneously, he/she was blocked gently and the trial was repeated.

Once the high and low-preference items had been identified using objects, the preference assessment was repeated for only these two items again in the object mode, then in picture and vocal modes. The picture preference assessment procedure was the same as the object preference assessment, except that 15 cm x 20 cm color photographs instead of the objects, were presented and the participant received the object corresponding to the chosen photograph on each trial. During the vocal preference assessment, each item was concealed in a box that looked identical. The boxes were presented to the participant one at a time and the tester named the item in the box while placing it on the table. The participant was then asked to choose. Upon choosing a box, the participant was given the item from that box.

Assessment of Basic Learning Abilities

The ABLA is a learning assessment of how rapidly an individual learns to perform several basic discrimination tasks (Kerr et al., 1977). The test consists of six

tasks (referred to as levels), which have been shown to be hierarchical in difficulty (Kerr et al.; Martin, Yu, Quinn, Patterson, 1983). Level 1 is an *imitation* task, which requires the testee to place a manipulandum into a container after the tester has demonstrated the response on each trial. Level 2 is a two-choice *position-visual* discrimination task, which requires the testee to place a piece of white foam into the container on the left when a yellow can and a red box are presented in the same left-right positions on each trial. Level 3 is a two-choice *visual* discrimination task and is similar to Level 2 except that the positions of the containers are switched randomly across trials. Level 4 is a partial identity *visual matching-to-sample* discrimination task, which requires the testee to place a yellow cylinder into the yellow can or a red cube into the red box on each trial. Across trials, the positions of the containers are changed randomly and the testee is given either the cylinder or the cube. Level 5 is a two-choice *auditory-position-visual* discrimination task, which requires the testee to place a white foam into the container requested verbally by the tester (e.g., "put it in the yellow can") on each trial. The requested container varies randomly across trials, but the left-right positions of the two containers remain the same. Lastly, Level 6 is an *auditory-visual* discrimination task, which is similar to Level 5 except that the positions of the containers change randomly across trials.

When testing a level, a testee is first given a demonstration trial, a guided trial, and a trial on which there is an opportunity for an independent response. Scoring on a level begins only after a client has responded correctly on the independent response trial. A testee receives a "pass" for a level if he/she performs that task correctly on 8 consecutive trials, or a "fail" if he/she has made 8 cumulative errors before reaching the pass criterion. Every correct response on scoring trials is followed by praise, and edibles

or toys. Each error is followed by the tester stating that the response is incorrect, a demonstration of the correct response, and an opportunity for an independent response (a correct response here would not be counted towards the pass criterion, but an incorrect response would be counted towards the fail criterion). The ABLA has demonstrated high inter-tester and test-retest reliabilities and high predictive validity (see review by Martin & Yu, 2000). Moreover, results of the ABLA have been shown to be related to preference assessment effectiveness in different modes (e.g., Conyers et al., 2002; de Vries et al., 2005; Thomson et al., 2007).

Stimuli for Experimental Tasks

To lessen the influence of history, stimuli for all matching tasks were made up of parts from everyday objects. For example, items included the axle of a toy car, part of a knob, the end of a shovel, etc. The smallest stimulus was approximately 2.5 by 7.5 by 2.5 cm and the largest stimulus was approximately 12.5 by 13 by 12.5 cm. Pictures of each item was a 15 by 20 cm color photograph, taken against a grey background. Stimuli for each task are shown in Appendices A through C, respectively.

Symmetric Object-Picture and Picture-Object Matching

An object-to-picture matching-to-sample task involves presenting objects as samples and pictures as comparisons. For example, in the presence of a cup (sample) and pictures of a cup and a plate (comparisons), selecting the picture of cup would be correct; whereas in the presence of a plate (sample) and the same picture comparisons, selecting the picture of the plate would be correct. A symmetric matching relation involves reversing the roles of the stimuli as samples and comparisons (Sidman & Tailby, 1982). Therefore, the symmetric relation of the above example would involve picture-to-object

matching with the pictures of plate and cup as samples and the objects as comparisons.

In the present study, I assessed whether participants were able to perform symmetric object-picture matching using one set of task stimuli (Appendix A), consisting of two objects and pictures of those objects. During object-picture matching, on each trial, two 15 cm by 20 cm color photographs (comparisons) were placed on the table approximately 15 cm apart in front of the participant and the participant was asked to look at each picture. An object (sample) corresponding to one of the comparisons was then held at the participant's eye level and the participant was asked to "match". Prior to test trials, participants were presented with a demonstration trial, a guided trial, and a trial on which there was an opportunity to make an independent response, using stimuli that were different from the task stimuli. A correct response was recorded if the participant approached the corresponding picture within 8 seconds. Selecting the incorrect comparison, or not responding within 8 seconds, was scored as incorrect. The positions of the comparison stimuli and the correct comparison were counterbalanced across trials such that each stimulus appeared in each position an equal number of trials and that the same sample stimulus was not presented for more than two consecutive trials. The assessment consisted of 10 trials. The experimenter said "thank you" at the end of each trial regardless of accuracy. No other programmed consequences were provided for responding. However, before presenting each trial, the participant was asked to perform a behavior unrelated to the task (e.g., roll a ball to the experimenter) and reinforced with praise and an edible in order to maintain general attending and instruction-following behaviors.

During picture-object matching, the same stimuli and procedure were used except

that the roles of the sample and comparison were reversed. That is, on each trial a picture was the sample and the objects were the comparisons.

The order of the object-picture and picture-object tasks was alternated across participants in each group. For each task (object-picture and picture-object), a "pass" was given for the task if a participant responded correctly on at least 8 of the 10 trials (80%) for that relation. The probability of obtaining 8 or more correct responses by chance, out of 10 trials, was approximately 0.055 (approximately 6 times in 100), assuming the responses across trials were independent.

Generalized Symmetric Object-Picture and Picture-Object Matching

Generalized matching is demonstrated when an individual responds correctly on the first trial when presented with a new matching task (Pear, 2001). In this study, the procedure for testing generalized symmetric object-picture matching was similar to the symmetric object-picture task described above except that 10 pairs of different stimuli were used (Appendix B). For each task (object-picture and picture-object), the stimulus pairs were presented in random order and each pair was presented for 1 trial. The positions of the correct comparisons were counterbalanced across trials. A "pass" was given for a task (object-picture and picture-object) if a participant responded correctly on at least 8 of the 10 trials (80%) for that relation. The order of object-to-picture and picture-to-object matching tasks was alternated across participants.

Generalized Picture-Picture Identity Matching

A matching-to-sample task in which the samples and comparisons are physically identical in all respects is called identity matching (Pear, 2001). In this study, generalized picture-picture identity matching was measured using 10 stimulus pairs involving

pictures that were not used in the previous tasks (Appendix C). Stimulus pairs were presented in random order, each pair for 1 trial, and the positions of the correct comparisons were counterbalanced across trials. On each trial, a picture was the sample and two pictures (one matching the sample) were comparisons. Except for the task stimuli, the definitions of correct and incorrect responses, consequences for responding, and presentation procedures were identical to those described above for generalized symmetrical matching.

Interobserver Reliability and Procedural Integrity Checks

Interobserver reliability and procedural integrity checks were conducted for each participant and for each assessment. During a reliability check, an observer independently recorded the participant's response on each trial. A trial was considered an agreement if both the experimenter and the observer recorded the same response. In contrast, a trial was considered a disagreement if the experimenter and the observer recorded different responses. Percent agreement was calculated using the following formula: number of agreements/(number of agreements + disagreements) x 100 (Martin & Pear, 2007).

An observer conducted reliability checks during: (a) 80% of the object preference assessment sessions, 75% of the picture preference assessment sessions, and 55% of the vocal preference assessment sessions; (b) 75% of the ABLA sessions; and (c) 75% of the matching task assessment sessions. Percent agreement per session was 100% across participants.

During each of the above sessions, the observer also evaluated procedural integrity by recording whether the experimenter carried out the procedures correctly on each trial using a behavior checklist (Appendices D, E, and F) that was appropriate for

the assessment being conducted. For example, the behaviors for the matching task assessments included presenting reinforcement for an alternate appropriate behavior, presenting the correct comparisons in the predetermined positions during the trial, presenting the correct sample, giving the appropriate instruction to begin the trial, and providing the appropriate consequence following a response. A trial was considered correct if all steps were carried out correctly. The percentage of trials carried out correctly per session was 100% across participants.

Results

Group Assignment

The average performance of the two groups on the paired-stimulus preference assessments is presented in Figure 1. The Object Group participants (P1 through P11) selected their HP food items during the paired-stimulus preference assessment on an average of 93.6% (range, 80 to 100%; SD = 9.2) of the trials in object mode, an average of 49.6% (range, 30 to 60%; SD = 9.1) of the trials in picture mode, and an average of 49.6% (range, 40 to 65%; SD = 8.8) of the trials in vocal mode. The Picture Group participants (P12 through P20) selected their HP items on an average of 97.8% (range, 90 to 100%; SD = 4.4) of the trials in object mode, an average of 88.3% (range, 80 to 95%; SD = 6.1) of the trials in picture mode, and an average of 52.8% (range, 45 to 60%; SD = 5.7) of the trials in vocal mode. Independent samples *t*-tests, assuming unequal variance, showed that the Object and Picture groups differed significantly on picture preference assessment effectiveness ($t(18) = 11.4$, $p < .001$, 2-tailed), and the two groups did not differ significantly on object ($p > 0.21$) or vocal ($p > 0.33$) preference assessments.

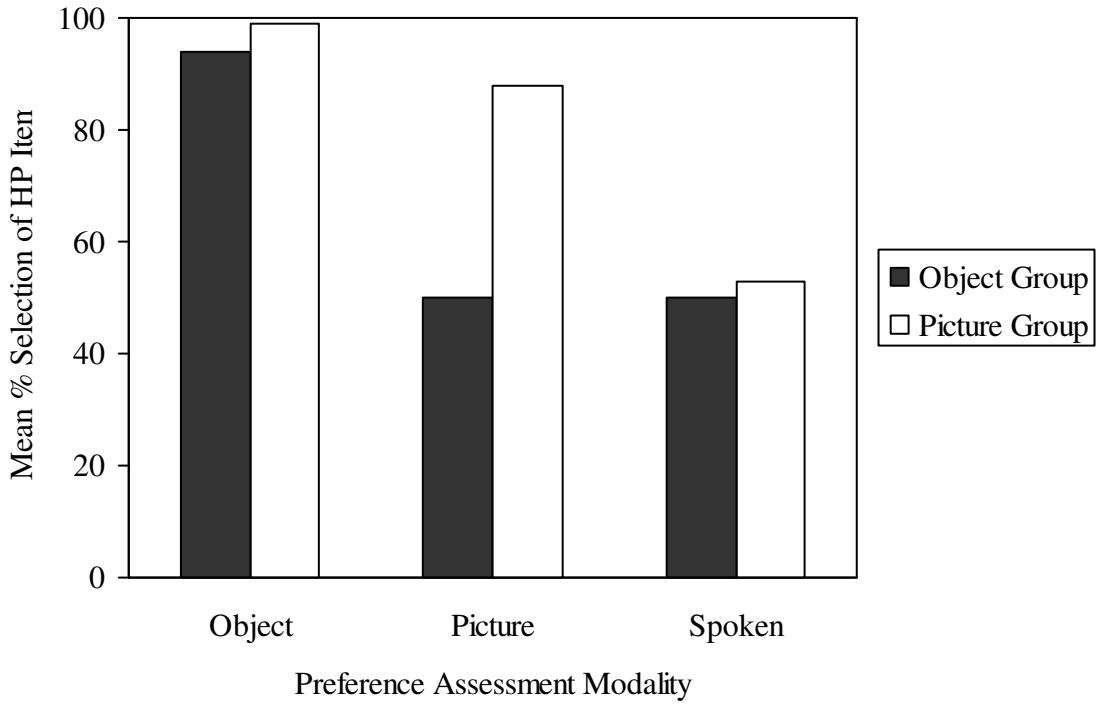


Figure 1. Mean percent selection for the high preference item in the object, picture, and spoken preference assessments for both the Object and Picture groups.

The individual data for preference assessments are shown in Table 2. As can be seen in Table 2, all participants met the criteria for group assignment – participants in the Object Group selected their HP items at or above 80% during object preference assessments and at or below 65% during picture and vocal assessments, whereas participants in the Picture Group selected their HP items at or above 80% during object and picture preference assessments and at or below 65% during vocal assessments.

Table 2. *Individual Data on Preference Assessments and Conditional Discriminations*

	Preference Assessment (% HP Selected [a])			Symmetric (% Correct)		Generalized Symmetric (% Correct)		Generalized ID Pic-Pic
Participants	Object	Picture	Vocal	Obj-Pic	Pic-Obj	Obj-Pic	Pic-Obj	(% Correct)
Object Group								
1	**90**	50	40	30	50	60	50	50
2	**100**	55	65	40	60	20	40	50
3	**100**	45	40	40	50	40	50	60
4	**100**	45	45	50	50	60	70	70
5	**100**	60	50	50	40	60	70	60
6	**100**	60	60	50	60	60	60	60
7	**100**	30	60	60	60	30	60	50
8	**80**	45	50	60	40	40	50	50
9	**80**	45	40	60	40	50	50	40
10	**100**	50	50	60	60	70	60	60
11	**80**	60	45	70	50	**80**	40	50
Picture Group								
12	**100**	**90**	50	60	70	**80**	**80**	60
13	**100**	**80**	50	70	50	60	50	**90**
14	**100**	**80**	45	70	50	70	50	60
15	**100**	**85**	60	70	70	**80**	60	**80**
16	**100**	**90**	50	**80**	50	70	**80**	**80**
17	**90**	**85**	50	**80**	60	**80**	**80**	**90**
18	**90**	**95**	50	**80**	60	**80**	**90**	50
19	**100**	**95**	60	**80**	70	**90**	**80**	**100**
20	**100**	**95**	60	**90**	**80**	60	40	**80**

[a] HP = high preference item.

Matching Task Performance

Table 3 shows the mean percentages of correct responses and standard deviations for the five tasks. The Object Group means were near chance level (50%) on all tasks, and the Picture Group means were higher than the Object Group means on all tasks. Levene's test of equality of variances was significant ($p < .05$) for two tasks (Generalized Symmetric Picture-Object and Generalized Identity Picture-Picture). Therefore, between-group comparisons were completed using independent samples *t*-tests for unequal

variances for these two tasks, and for equal variances for the other three tasks. Overall alpha was set at .05 with Bonferroni correction; thus each comparison was evaluated using $p = .01$, 1-tailed. The Picture Group means were significantly higher than the Object Group means on 4 tasks (Symmetric Object-Picture, Symmetric Picture-Object, Generalized Symmetric Object-Picture, and Generalized Identity Picture-Picture).

Table 3

Group means and standard deviations of percent correct responses.

	Object Group (n = 11) Mean (SD)	Picture Group (n = 9) Mean (SD)	t	p
Symmetric Object-Picture	51.82 (11.68)	75.56 (8.82)	5.028	.000
Symmetric Picture-Object	50.91 (8.31)	62.22 (10.93)	2.632	.009
Generalized Symmetric Object-Picture	51.82 (17.79)	74.44 (10.14)	3.383	.002
Generalized Symmetric Picture-Object	54.55 (10.36)	67.78 (17.87)	1.967	.036
Generalized Identity Picture-Picture	54.55 (8.20)	76.67 (16.58)	3.653	.002

The number and percentage of participants in each group who met the pass criterion on each task are shown in Table 4. Only 1 Object Group participant met the criterion on one task, whereas at least 5 of the 9 Picture Group participants met the criterion on 4 tasks.

For the Picture Group, examination of the individual data (Table 2) revealed that 8 of the 9 participants passed at least one task (P14 was the exception). Of these 8 participants, 2 passed 4 tasks (P17 and P19), 3 passed 3 tasks (P16, P18, and P20), 2 passed 2 tasks (P12 and P15), and 1 passed 1 task (P13).

Table 4

Number (%) of participants who met the pass criterion (≥ 80%) on each task

	Symmetric		Generalized Symmetric		Generalized Identify Picture-Picture
	Object-Picture	Picture-Object	Object-Picture	Picture-Object	
Object Group n = 11	0 (0)	0 (0)	1 (9.1)	0 (0)	0 (0)
Picture Group n = 9	5 (55.6)	1 (11.1)	5 (55.6)	5 (55.6)	6 (66.7)

Discussion

The Picture Group performed significantly better than the Object Group on 4 of the 5 matching tasks (Table 3). Individual data showed that only 1 Object Group participant (P11) met the pass criterion on one task, whereas 8 of 9 Picture Group participants met the pass criterion on at least one task with 7 participants passing two or more tasks (Table 2). Despite the between-group differences on the 4 discrimination tasks, the relative importance of each task is less clear. Moreover, examination of individual data revealed some unexpected performance patterns that warrant discussion.

First, for the symmetric object-picture and picture-object tasks, although 5 participants passed the object-picture task, 4 of these 5 participants did not pass the

reverse picture-object relation involving the same stimuli (Table 2). No participant showed the opposite pattern (passing picture-object and failing object-picture). This suggests that the picture-object relation (pictures as samples and objects as comparisons) may be more difficult than the object-picture relation when the roles of the stimuli are reversed. Of the 4 participants who did not pass the picture-object relation, 3 were tested on the symmetric picture-object task before the symmetric object-picture task, which may suggest a possible practice effect. However, this was not observed with other participants. For participants who were tested on the symmetric object-picture task before the symmetric picture-object task, mean accuracy was 60.9% and 53.6% for the two tasks, respectively. For participants who were tested on the two tasks in the reverse order, mean accuracy was 58.9% and 64.4% for the two tasks, respectively. In other words, participants performed slightly better on the symmetric object-picture task than on the picture-object task regardless of order.

Second, all 4 participants who had passed the symmetric object-picture task and failed the symmetric picture-object task also passed the generalized symmetric picture-object task. This was unexpected in that generalized matching is usually considered more difficult (Pear, 2001). This pass/fail pattern was also observed to a lesser extent in object-picture matching. Of the 5 Picture Group participants (P12, P15, P17, P18, and P19) and 1 Object Group participant (P11) who had passed the generalized symmetric object-picture task, 3 (P11, P12, P15) did not pass the symmetric object-picture task. What might have caused this performance pattern? One possibility is that the observed results may have been due to a task order effect (e.g., symmetric tasks preceded generalized symmetric tasks). For those participants who failed the symmetric object-picture

matching task, but passed the generalized symmetric object-picture matching task, 2 of the 3 participants received the symmetric task prior to the generalized symmetric tasks. Of the 4 participants who had failed the symmetric picture-object task, but passed the generalized symmetric picture-object task, 2 were presented with the symmetric task prior to the generalized symmetric task. Therefore, the results could not be attributed to order of testing. Examination of other participants also showed no apparent practice effect (i.e., performance improvement as testing progressed). When examining all participants as a group, the mean percent correct was 60.5, 60.0, 61.5, 61.0, and 62.5 from the first to the last task, respectively.

Participant 14 also provided an unexpected finding. He was the exception in the Picture Group who did not pass any of the 5 tasks. His highest accuracy was 70% correct on the symmetric and generalized symmetric object-picture tasks. It is possible that this participant would have met the pass criterion if more test trials were administered. However, Participant 14's performance does raise the possibility that skills other than those tested may be important to the effectiveness of picture preference assessment.

Except for P14, however, all other Picture Group participants passed at least 1 of the 3 generalized matching tasks (P13 and P20 passed 1 each; P12, P15, P16, and P18 passed 2 each; and P17 and P19 passed 3 each). Perhaps the critical skill is to be able to perform some form of generalized matching, be it object-to-picture, picture-to-object, or picture-to-picture. The discriminations investigated are clearly not exhaustive, future research might consider including identity object-object matching and generalized identity object-object matching.

Further research is needed to determine if generalized conditional discriminations are functionally related to the effectiveness of picture preference assessments. Similar to the approach taken by Nguyen et al. (2009), this can be evaluated by teaching individuals who are unable to perform picture preference assessments (i.e., individuals in the Object Group) the failed matching tasks. If, after learning the matching tasks, the individual shows improved concordance between object and picture preference assessments, it can be said that the matching task is a prerequisite to perform preference assessments using pictures. Research is also needed to evaluate whether one or a combination of discrimination skills should be targeted for training such that effective picture preference assessment can be achieved most efficiently.

A secondary finding of this study is the relation between ABLA performance and picture preference assessments. First, previous studies have shown that passing ABLA Levels 2 and 3 (ability to learn to perform simple discriminations with position and/or visual cues) are associated with object preference assessments (Thomson et al., 2007; Conyers et al., 2002; de Vries et al., 2005). Consistent with this finding, the present study shows that all 11 Object Group participants passed up to ABLA Levels 2 or 3 and none passed Level 4. Second, previous research also showed that passing ABLA Level 4 (ability to learn to perform partial-identity visual matching) is positively correlated with picture preference assessment effectiveness (Conyers et al.; de Vries et al.). Consistent with this finding, all 4 participants (P12, P16, P17, and P19) who passed up to ABLA Level 4 were able to respond to picture preference assessments. However, of the 16 participants who did not pass ABLA Level 4, 5 (P13, P14, P15, P18, and P20) or 31% selected their HP items in the picture mode at 80% or higher. This suggests that failing

ABLA Level 4 may underestimate a client's ability to respond to picture preference assessments. Future research should examine the association between the ABLA and picture preference assessment effectiveness.

Overall, the results of this study extend previous research on relations between discrimination skills and picture preference assessments in several ways. First, this study extends previous research by Clevenger and Graff (2005) and Nguyen et al. (2009) by using a larger sample in a group design. Second, the participant's ability to respond to vocal preference assessments was unknown in the Clevenger and Graff study. In this study, the participant's ability to respond to preference assessments in the vocal mode was controlled through direct preference assessment prior to testing the discrimination tasks. Third, this study expanded on previous research by examining the relation between generalized matching and picture preference assessment effectiveness. Lastly, this study added to the findings on the relation between ABLA performance and preference assessments. As our understanding of the relation between various discriminations and picture preference assessments increase, we are in a better position to design an effective training program for clinical patients by incorporating critical discriminations as target behaviors.

References

Bojak, S. L., & Carr, J. E. (1999). On the displacement of leisure items by food during multiple stimulus preference assessments. *Journal of Applied Behavior Analysis, 32*, 515-518.

Cannella, H. I., O'Reilly, M. F., & Lancioni, G. E. (2005). Choice and preference assessment research with people with severe to profound developmental disabilities: A review of the literature. *Research in Developmental Disabilities, 26,* 1-15.

Carr, J. E., Nicolson, A. C., & Higbee, T. S. (2000). Evaluation of a brief multiple-stimulus preference assessment in a naturalistic context. *Journal of Applied Behavior Analysis, 33*, 353-357.

Clevenger, T. M., & Graff, R. B. (2005). Assessing object-to-picture and picture-to-object matching as prerequisite skills for pictorial preference assessments. *Journal of Applied Behavior Analysis, 38*, 543-547.

Conyers, C., Doole, A., Vause, T., Harapiak, S., Yu, D.C.T., & Martin, G.L. (2002). Predicting the relative efficacy of three presentation methods for assessing preferences of persons with developmental disabilities. *Journal of Applied Behavior Analysis, 35,* 49-58.

Daly, E. J., III, Wells, N. J., Swanger-Gagne, M. S., Carr, J. E., Kunz, G. M., & Taylor, A. M. (2009). Evaluation of the multiple-stimulus without replacement preference assessment method using activities as stimuli. *Journal of Applied Behavior Analysis, 42*, 563-574.

DeLeon, I. G., & Iwata, B. A. (1996). Evaluation of a multiple-stimulus presentation format for assessing reinforcer preferences. *Journal of Applied Behavior Analysis, 29,* 519-533.

de Vries, C., Yu, C. T., Sakko, G., Wirth, K. M., Walters, K. L., Marion, C., & Martin, G. L. (2005). Predicting the relative efficacy of verbal, pictorial, and tangible stimuli for assessing preferences of leisure activities. *American Journal on Mental Retardation, 110,* 145-154.

Fisher, W., Piazza, C. C., Bowman, L. G., Hagopian, L. P., Owens, J. C., & Slevin, I. (1992). A comparison of two approaches for identifying reinforcers for persons with severe and profound disabilities. *Journal of Applied Behavior Analysis, 25,* 491-498.

Green, C. W., Gardner, S. M., & Reid, D. H. (1997). Increasing indices of happiness among people with profound multiple disabilities: A program replication and component analysis. *Journal of Applied Behavior Analysis, 30,* 217-228.

Green, C. W., Reid, D. H., White, L. K., Halford, R. C., Brittain, D. P., & Gardner, S. M. (1988). Identifying reinforcers for persons with profound handicaps: Staff opinion versus systematic assessment of preferences. *Journal of Applied Behavior Analysis, 21,* 31-43.

Hagopian, L. P., Long, E. S., & Rush, K. S. (2004). Preference assessment procedures for individuals with developmental disabilities. *Behavior Modification, 28(5),* 668-677.

Higbee, T. S., Carr, J. E., & Harrison, C. D. (1999). The effects of pictorial versus tangible stimuli in stimulus preference assessments. *Research in Developmental Disabilities, 20*, 63-72.

Higbee, T. S., Carr, J. E., & Harrison, C. D. (2000). Further evaluation of the multiple-stimulus preference assessment. *Research in Developmental Disabilities, 21,* 61-73.

Horrocks, E., & Higbee, T.S. (2008). An evaluation of a stimulus preference assessment of auditory stimuli for adolescents with developmental disabilities. *Research in Developmental Disabilities, 29,* 11-20.

Kerr, N., Meyerson, L., & Flora, J. (1977). The measurement of motor, visual, and auditory discrimination skills. *Rehabilitation Psychology, 24,*156-170.

Lee, M. S. H., Yu, C. T., Martin, T., & Martin, G. L. (2010). On the relation between preference and reinforcer efficacy. *Journal of Applied Behavior Analysis, 43,* 95-100.

Martin, G. L., & Pear, J. J. (2007). *Behavior modification: What it is and how to do it. 8th Edition.* Upper Saddle River, NJ: Prentice Hall.

Martin, G. L., & Yu, D. C. T. (2000). Overview on research of the Assessment of Basic Learning Abilities test. *Journal on Developmental Disabilities*, *7*(2), 10-36.

Martin, G. L., Yu, D., Quinn, G., & Patterson, S. (1983). Measurement and training of AVC discrimination skills: Independent confirmation and extension. *Rehabilitation Psychology, 28,* 231-237.

Martin, T. L., Yu, C. T., Martin, G. L., & Fazzio, D. (2006). On choice, preference, and preference for choice. *The Behavior Analyst Today, 7(2),* 234-241.

Nguyen, D. M., Yu, C. T., Martin, T. L., Fregeau, P., Pogorzelec, C., & Martin, G. L. (2009). Teaching object-picture matching to improve concordance between object and picture preferences for individuals with developmental disabilities: Pilot study. *Journal on Developmental Disabilities, 15(1),* 53-64.

Pace, G. M., Ivancic, M. T., Edwards, G. L., Iwata, B. A., & Page, T. J. (1985). Assessment of stimulus preference and reinforcer value with profoundly retarded individuals. *Journal of Applied Behavior Analysis, 18,* 249-255.

Paramore, N.W., & Higbee, T.S. (2005). An evaluation of a brief multiple-stimulus preference assessment with adolescents with emotional/behavioral disorders (E/BD) in an educational setting. *Journal of Applied Behavior Analysis, 38*, 399-404.

Pear, J. J. (2001). *The science of learning.* Philadelphia, PA: Psychology Press.

Reyer, H. S., & Sturmey, P. (2006). The Assessment of Basic Learning Abilities (ABLA) test predicts the relative efficacy of task preferences for persons with developmental disabilities. *Journal of Intellectual Disability Research, 50(6),* 404-409.

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

Thomson, K. M., Czarnecki, D., Martin, T. L., Yu, C. T., & Martin, G. L. (2007). Predicting optimal preference assessment methods for individuals with developmental disabilities. *Education and Training in Developmental Disabilities, 42(1)*, 107-114.

Vollmer, T. R., Marcus, B. A., & LeBlanc, L. (1994). Treatment of self-injury and hand mouthing following inconclusive functional analysis. *Journal of Applied Behavior Analysis, 27,* 331-344.

Wehmeyer, M., & Schwartz, M. (1997). Self-determination and positive adult outcomes: A follow-up study of youth with mental retardation or learning disabilities. *Exceptional Children, 63,* 245-255.

Windsor, J., Piche, L. M., & Locke, P. A. (1994). Preference testing: A comparison of two presentation methods. *Research in Developmental Disabilities, 15,* 439-455.

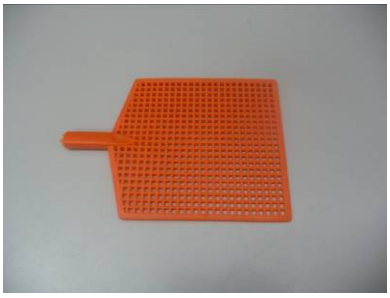
Appendix A

Stimuli for Symmetric Object-Picture and Picture-Object Matching Task

Appendix B

Stimuli for Generalized Symmetric Object-Picture and Picture-Object Matching Task



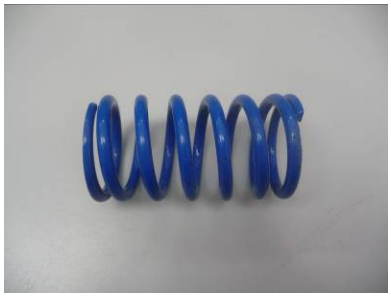







Appendix C

Stimuli for Generalized Identity Picture-Picture Matching Task










Appendix D

Procedural Reliability Checklist for Symmetric Object-Picture and Picture-Object Matching

Participant:_____ Tester: _____
Date: _____ Observer: _____

A = Blue Ball
B = Green / Yellow Sponge

Object-Picture Matching Task						
Trial	Sample (object)	Comparisons (pictures)		Procedural Reliability		
				Correct stimulus	Correct verbal prompt	Correct Consequence
1	Sponge	Sponge	Ball			
2	Ball	Sponge	Ball			
3	Ball	Ball	Sponge			
4	Sponge	Sponge	Ball			
5	Ball	Sponge	Ball			
6	Sponge	Ball	Sponge			
7	Ball	Ball	Sponge			
8	Sponge	Sponge	Ball			
9	Ball	Sponge	Ball			
10	Sponge	Ball	Sponge			

% correct = _____ % responses (left) = _____
% responses (right) = _____

Picture-Object Matching Task						
Trial	Sample (picture)	Comparisons (objects)		Procedural Reliability		
				Correct stimulus	Correct verbal prompt	Correct Consequence
10	Sponge	Ball	Sponge			
9	Ball	Sponge	Ball			
8	Sponge	Sponge	Ball			
7	Ball	Ball	Sponge			
6	Sponge	Ball	Sponge			
5	Ball	Sponge	Ball			
4	Sponge	Sponge	Ball			
3	Ball	Ball	Sponge			
2	Ball	Sponge	Ball			
1	Sponge	Sponge	Ball			

% correct = _____ % responses (left) = _____
% responses (right) = _____

Appendix E

Procedural Reliability Checklist for Generalized Symmetric Object-Picture and Picture-Object Matching

Participant:_____ Tester: _____
Date: _____ Observer: _____

Generalized Symmetric Object-Picture Matching Tasks						
Task ID	Sample (object)	Comparisons (pictures)		Procedural Reliability		
				Correct stimulus	Correct verbal prompt	Correct Consequence
1	A	A	B			
2	A	A	B			
3	B	A	B			
4	A	A	B			
5	B	A	B			
6	A	A	B			
7	B	A	B			
8	B	A	B			
9	A	A	B			
10	B	A	B			

% correct = _____ % responses (left) = _____
% responses (right) = _____

Generalized Symmetric Picture-Object Matching Tasks						
Task ID	Sample (object)	Comparisons (pictures)		Procedural Reliability		
				Correct stimulus	Correct verbal prompt	Correct Consequence
10	B	A	B			
9	A	A	B			
8	B	A	B			
7	B	A	B			
6	A	A	B			
5	B	A	B			
4	A	A	B			
3	B	A	B			
2	A	A	B			
1	A	A	B			

% correct = _____ % responses (left) = _____
% responses (right) = _____

Task Descriptions

ID #1
A = red/white stick
B = pink circle

ID #2
A = blue square
B = white hexagon

ID #3
A = pink stick
B = black circle

ID #4
A = gold half-circle
B = red knob

ID #5
A = red spoon
B = green wheels

ID #6
A = green hook
B = purple tube

ID #7
A = blue cone
B = orange net

ID #8
A = blue square
B = white tube

ID #9
A = silver dome
B = foam

ID #10
A = blue circle
B = red person

Appendix F

Procedural Reliability Checklist for Generalized Identity Picture-Picture Matching

Participant:_____ Tester: _____
Date: _____ Observer: _____

Generalized Identity Matching with Pictures Tasks						
Task ID	Sample (picture)	Comparisons (pictures)		Procedural Reliability		
				Correct stimulus	Correct verbal prompt	Correct Consequence
1	A	A	B			
2	A	A	B			
3	B	A	B			
4	B	A	B			
5	A	A	B			
6	B	A	B			
7	B	A	B			
8	A	A	B			
9	A	A	B			
10	B	A	B			

% correct = _____

% responses (left) = _____
% responses (right) = _____

Task Descriptions

ID #1
A = yellow foot
B = green wheel

ID #2
A = blue tube
B = black star

ID #3
A = blue scoop
B = pink hand

ID #4
A = orange hoop
B = pink sparkle

ID #5
A = blue flower
B = red circle

ID #6
A = blue coil
B = yellow spout

ID #7
A = green fan
B = pink clip

ID #8
A = gold dish
B = black brush

ID #9
A = blue circle
B = orange tube

ID #10
A = black funnel
B = green stick