MATCHING VISUAL STIMULI: DOES SIMILARITY MATTER?

Kelsey F. Burren [1,2], Chata A. Dickson[1,2]

[1] Western New England University
[2] New England Center for Children

Three young men with autism spectrum disorder participated in this study investigating whether the form of visual stimuli affects the acquisition of object-to-picture matching. In Experiment 1, object-to-picture matching was assessed with photos and line drawings. In Experiment 2, relations were taught between the objects and arbitrary stimuli and then probes were conducted for the emergence of untrained relations between the arbitrary stimuli and the photos and line drawings. A multiple probe design was used to compare performances based on photos vs. line drawings. For two participants, there was no difference in trials to mastery between photos and line drawings. For the third participant, mastery criteria were more readily achieved with photos than line drawings in eight of eleven comparisons; no differences were observed within the remaining three comparisons. Equivalence relations emerged between arbitrary symbols and both photos and line drawings for the first two participants, but object-to-arbitrary symbol relations were not demonstrated by the third participant even after direct training. The use of stimuli with greater visual similarity to the target object may yield higher accuracy with object-to-picture relations for some individuals with autism.

Keywords: augmentative and alternative communication; equivalence relations; matching-to-sample; visual similarity

Since the first applications of behavior analysis to the teaching of children with autism spectrum disorder, researchers and practitioners have sought to improve strategies for teaching stimulus relations to children with autism (McIlvane et al. 2016; Sidman & Stoddard, 1966). Matching related stimuli is a foundational skill upon which various important repertoires are built. Perhaps most importantly, for some children with autism who have difficulty communicating through speech, learned relations between pictorial stimuli and their referents are the basis for communication (Hurlbut et al., 1982; Shafer, 1993). Many of these children learn to communicate using augmentative and alternative communication (AAC) systems (McIlvane et al., 2016; Romski & Sevcik, 1997). Common aided AAC systems primarily rely on two-dimensional images, hereafter, *pictures*, which the learner uses to communicate with others. For example, the pictures used in aided AAC often take a variety of forms, including full-color photographs of real-world objects and people in the learner's environment, generic hand- or computer-drawn graphics (e.g., clip art), and black-and-white line drawings. Although some pictures share physical attributes with their real-world referents (e.g., photographs), allowing for the possibility of feature-based stimulus classes, arbitrary stimuli by contrast share no identifiable features and form arbitrary stimulus classes (e.g., written letters; McIlvane et al., 1993). The degree to which these pictures are similar refers to a quantifiable relation between the structure or

Author Note: Some of the data were presented at the annual convention of the Berkshire Association for Behavior Analysis and Therapy in 2019 and the annual convention of the Association for Behavior Analysis International in 2020. The data included in this manuscript were submitted to the Department of Psychology in the School of Arts and Sciences at Western New England University by the first author under the supervision of the second author in partial fulfillment of the requirements for the master's degree in behavior analysis. Address correspondence to Department of Chata A. Dickson, New England Center for Children, 33 Turnpike Rd, Southborough, MA - 01772. Email: cdickson@necc.org

Acknowledgments: We thank Dr. Rebecca MacDonald, Dr. Jonathan Pinkston, and Dr. Eileen Roscoe for their feedback on earlier versions of this manuscript, and Joseph Morency for preparing the line drawing stimuli used in the study. We also thank the students, staff, families, and administration of NECC whose cooperation and support made this study possible.

overlapping features of the stimuli. Practical use of these pictures requires discrimination between various stimuli, learning relations between these stimuli and their real-world referents, and using a communication exchange system with communication partners in their environments to access real-world objects, persons, or activities (Lionello-DeNolf & McIlvane, 2016).

There is evidence that some individuals with autism do not learn relations between objects and pictures in the same way as children without disabilities (e.g., Carr & Felce, 2008; Green et al., 1990; Higbee et al., 1999; McIlvane et al., 1990; Nguyen et al., 2009), and this can be a barrier to the effective use of aided AAC (Lionello-DeNolf & McIlvane, 2016; Shafer, 1993; Stromer et al., 1996). Improvements to the technology for teaching relations between pictures and their referents could contribute to establishing functional use of AAC systems for these children. Little is known about the influence that the form or features of these pictures have on learning outcomes for children with autism spectrum disorder, but it is possible that the use of a particular form may help children with autism learn to communicate more efficiently using aided AAC (Lionello-DeNolf & McIlvane, 2016).

A few studies have evaluated methods for remediating problems of learning relations between two- and three-dimensional stimuli. Dixon (1981) found that, when matching objects and pictures, some children with autism matched on the basis of physical properties, and paired objects with objects and pictures with pictures regardless of their content. When modified three-dimensional picture cutouts were presented as samples with corresponding pictures and objects as comparisons, the majority of participants selected the objects, suggesting control by the dimensional properties of the stimuli. Dixon developed an effective strategy for establishing object-to-picture matching by gradually fading in a background to the cutout and then gradually reducing its three-dimensional properties; essentially, gradually transforming comparison objects into pictures. Lionello-DeNolf and McIlvane (2016) extended Dixon's findings by investigating strategies to teach object-to-picture matching to a nine-year-old boy with autism spectrum disorder and limited language skills. In initial testing, the participant demonstrated proficiency with identity matching of both objects and photos, but undifferentiated responding with objects as the sample and photos as the comparison. In this case, although features of the stimuli controlled responding on identity-matching trials, there was no control by these features when the task was to match pictures to objects. Unlike in Dixon's study, accuracy remained low when two-dimensional photos were replaced with size-matched three-dimensional photo cutouts, and when the cutout was presented as a sample with the corresponding photo and object as comparisons, there was no clear tendency to match the cutout with either the photo or the object. A series of classification tests and stimulus-fading training steps were unsuccessful in isolating the variables that controlled comparison selection.

In another study of the influence of physical properties of stimuli on children's selections in a matching task, Hartley and Allen (2015) conducted tests with different types of stimuli. Children with autism were taught to select a target picture of an unfamiliar object upon hearing the name of the object. Target picture stimuli included black-and-white line drawings, color line drawings, greyscale photos, and color photos. Next, both the target picture and the previously unseen target object were presented as comparisons, and the participant could select one or both given the same spoken name. Finally, participants were presented with the target picture along with a version of the object in a different color and could again select one or both given the spoken name. Unlike children without disabilities, children with autism generally selected the picture given the spoken name rather than the corresponding object, but when the target picture was a color line drawing or photo, the object was twice as likely to be selected compared to when the picture was black and white or greyscale. The authors concluded that for children with autism, color was a critical feature for enhancing generalization of spoken names from pictures to three-dimensional objects.

The current study is a translational investigation of the effects of stimulus type on relational learning in children with autism. Educational or therapeutic procedures were not directly evaluated; however, the identification of differences in acquisition or emergence of relations by stimulus type may lead to improvements in preparing children to communicate using visual stimuli. For example, more efficient mastery of matching objects to photographs than to line drawings could suggest that greater visual similarity to the target is important in learning to match visual stimuli. The purpose of Experiment 1 was to evaluate performance of object-to-picture matching with different types of visual stimuli prior to training to determine whether the form of the picture stimulus produced differential accuracy.

EXPERIMENT 1

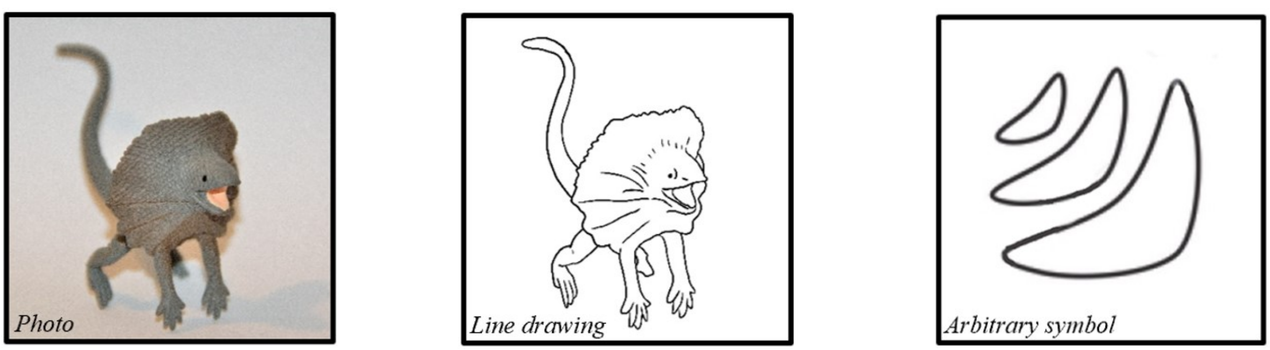
Method

Participants

Participants were three students at a behaviorally based residential special education school for children with autism spectrum disorder and developmental disabilities. Participants were recruited from the primary investigator's caseload to ensure familiarity with their communication needs and educational programming and availability to conduct research sessions. Inclusion criteria were (a) receipt of consultative services from a speech-language pathologist, (b) appropriate session behavior for tabletop discrimination tasks (e.g., remaining seated for at least nine consecutive trials, scanning between at least three visual stimuli, attending to prompts), and (c) a score of 5/5 on both picture-to-picture conditional identity matching and object-to-picture conditional matching on the NECC-CSA (New England Center for Children- Core Skills Assessment; Dickson et al., 2014). The last criterion was implemented to verify object-to-picture conditional matching skills such that the effects of manipulating different stimulus conditions could be assessed. Consent was obtained from each participant's legal guardian. Participants provided daily assent by indicating willingness to participate in sessions and could refuse to participate or request to terminate sessions at any time by using any form of communication or simply moving away from the task.

Rowan was a 20-year-old man with a dual diagnosis of Landau-Kleffner syndrome and autism spectrum disorder. Rowan primarily communicated using an iPad[1] with the AAC app Proloquo2Go[2]. He used AAC to make requests and to answer conversational questions. In some cases, he typed his responses to these questions. Rowan also used gestures and some manual signs. The Peabody Picture Vocabulary Test- 4th Edition (PPVT-4) was administered to assess his responses to single spoken words, and he obtained an age-equivalence score of 2;3 (years;months). Carter was a 21-year-old man diagnosed with autism spectrum disorder. Carter communicated vocally using single words and short phrases, as well as complete sentences when prompted. His age-equivalence score from the PPVT-4 was 4;10. Francis was a 16-year-old boy with a diagnosis of autism spectrum disorder. Francis communicated primarily with an iPad using the AAC app TouchChat[3]. He used AAC to make requests by touching one or two icons on the screen. Francis also used gestures and some manual signs. The PPVT-4 was administered, and Francis did not obtain the minimum number of correct responses in the first set of this test to achieve a basal score. He

Figure 1. Sample Picture Stimuli



Note. From left to right, photo, line drawing, and arbitrary symbol.

[1] iPad is a product of Apple Computers Inc., Cupertino, CA, www.apple.com

[2] Proloquo2go is a product from AssistiveWare, www.assistiveware.com/product/proloquo2go

[3] TouchChat is a product from Saltillo, www.touchchatapp.com

Table 1. Stimuli Included in Each Experimental Set

Set	Stimulus 1	Stimulus 2	Stimulus 3
1	Cockatoo	Crocodile	Dingo
2	Frilled lizard	Elephant	Eagle
3	Armadillo	Gorilla	Hippo
4	Platypus	Polar Bear	Rhino
5	Turtle	Wombat	Tiger
6	Prairie dog	Shark	Blue-footed booby
7	Crab	Water iguana	Big horn ram
8	Coyote	Red-footed booby	Horned lizard
9	Scorpion	Penguin	Mountain Lion
10	Zebra	Tasmanian devil	Road runner
11	Seal	Land iguana	Bobcat
12	Kangaroo	Koala	Panda

Note. Sets 1-4 were tested for Rowan and Carter. Sets 1-12 were tested for Francis.

identified some common nouns (e.g., ball, banana, shoe), but errors were noted for verbs as well as concepts that may have been unfamiliar (e.g., eating, duck, dog).

Setting and Materials

Sessions took place in a quiet space at the participants' school or residential home. Sessions were conducted in each participant's typical learning environment, which varied based on each individual's learning needs but minimally included a desk and a chair. In addition to the participant and experimenter, up to two other students and one other teacher were present in the room and engaging in educational activities while sessions took place.

The materials included three-dimensional animals (objects) made of solid plastic and ranging in size from 3.2 to 9.5 cm tall, each with a corresponding photo, line drawing, and arbitrary stimulus. The two-dimensional stimuli were printed on white paper and laminated and measured approximately 12.5 by 8.5 cm. Figure 1 shows the stimuli that correspond with one of the objects, Figure 2 shows the stimulus relations to be trained and tested for emergence, and Table 1 lists the stimuli in each of the sets used for this study.

Research Design

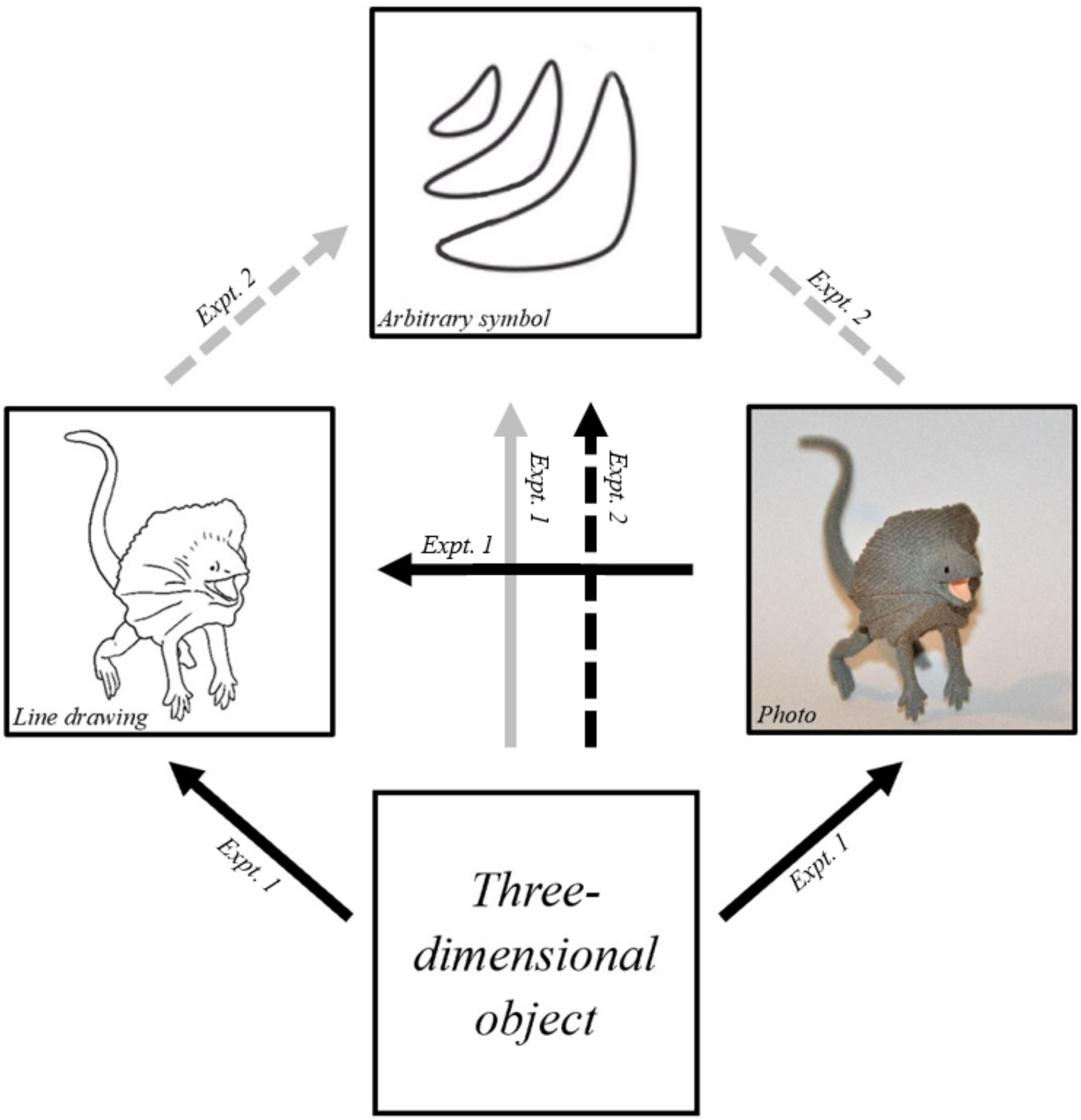
Prior to recruiting participants, all aspects of the study were reviewed and approved by the Institutional Review Board of the university with which the authors were affiliated. Matching-to-sample performances were assessed with photos and line drawings and compared within each participant. For Experiment 1, the dependent variable was accuracy of untrained responses and the independent variables were the various pairings of sample and comparison stimulus types in the matching-to-sample arrangement: object-to-photo (O-P), object-to-line drawing (O-LD), object-to-arbitrary symbol (O-AS), line drawing identity matching (LD-LD), and photo-to-line drawing matching (P-LD). Although Experiment 1 consisted of assessment only (no training was conducted), criteria were applied to determine whether a relation was mastered: 89% or greater accuracy across two consecutive sessions. A visual schematic of the trained and potential emergent relations is provided in Figure 2.

Trained observers with experience in data collection and reliability measures viewed 33% of session videos for each participant across Experiments 1 and 2 and collected trial-by-trial data on accuracy of responses and whether reinforcement was delivered following a response. Responses scored by both the researcher and the second observer as correct or incorrect were rated as agreements. Procedural integrity was evaluated through measurement of the delivery of the programmed consequences.

Procedure

Matching-to-sample assessments were conducted with each stimulus set. Prior to each session, the researcher conducted a brief paired stimulus preference assessment with potential reinforcers for each participant. In general, social reinforcers were used with Rowan and Carter and preferred snacks were used with Francis. Preferred items or activities were delivered following every third or fourth trial along with praise related to session behavior (e.g., *nice* *sitting*). Sets of three objects were selected for each participant, along with corresponding photos, line drawings, and arbitrary symbols (see Figure 1 for an example). Each session consisted of nine trials with one set of three samples and the corresponding pictures, with each sample stimulus presented on three of the nine trials. At the start of each trial, the researcher handed the sample (either objects or pictures) to the participant and then revealed an array of three comparison stimuli (either photos, line drawings, or arbitrary stimuli). The participant was expected to point to or touch a comparison or place the sample stimulus with a comparison. No programmed differential consequences were provided for correct or incorrect responses. After each trial, the researcher recorded whether the response was correct or incorrect, the stimuli were removed from the table, and the next trial began. Each of

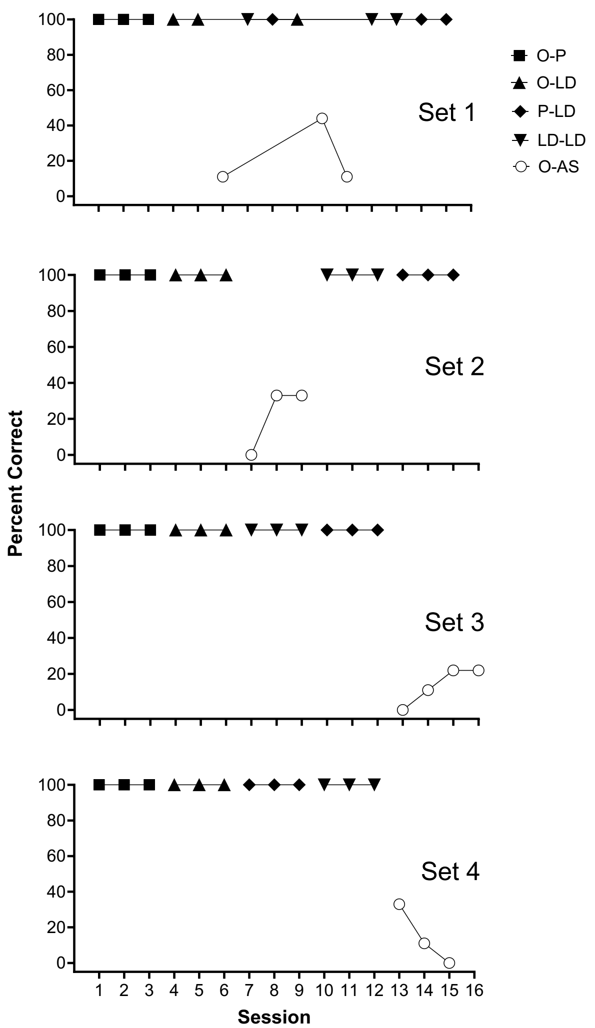
Figure 2. Networks of Matching-to-Sample Performances



Note. Solid black and grey arrows represent relations assessed during Experiment 1. Broken black arrows represent relations established during training in Experiment 2, and broken grey arrows represent potential emergent relations assessed during Experiment 2.

the following tests were conducted for all sets: O-P, O-LD, O-AS, LD-LD, and P-LD. Sessions continued with each set of stimuli for a minimum of three sessions, until the mastery criterion of 89% or greater accuracy across two consecutive session was met, or until stable accuracy scores were observed. A maximum of eight sessions were conducted for each test.

Figure 3. Accuracy of Responses During Experiment 1 for Rowan



Replication
The above procedures were repeated with each stimulus set, for a total of four stimulus sets for Rowan and Carter. For Francis, additional replications were conducted for a total of 12 stimulus sets following observed differences in outcomes between O-P and O-LD tests with the initial four sets.

Results

Reliability
Interobserver agreement (IOA) for correct responses was calculated for each session by dividing the number of responses scored in agreement by the total number of trials and converting the result to a percentage. For Rowan, mean agreement for 24 sessions across all conditions was 100%. For Carter, mean agreement for 15 sessions across all conditions was 100%. For Francis, mean agreement for 97 sessions across all conditions was 99% (range: 67% – 100%).

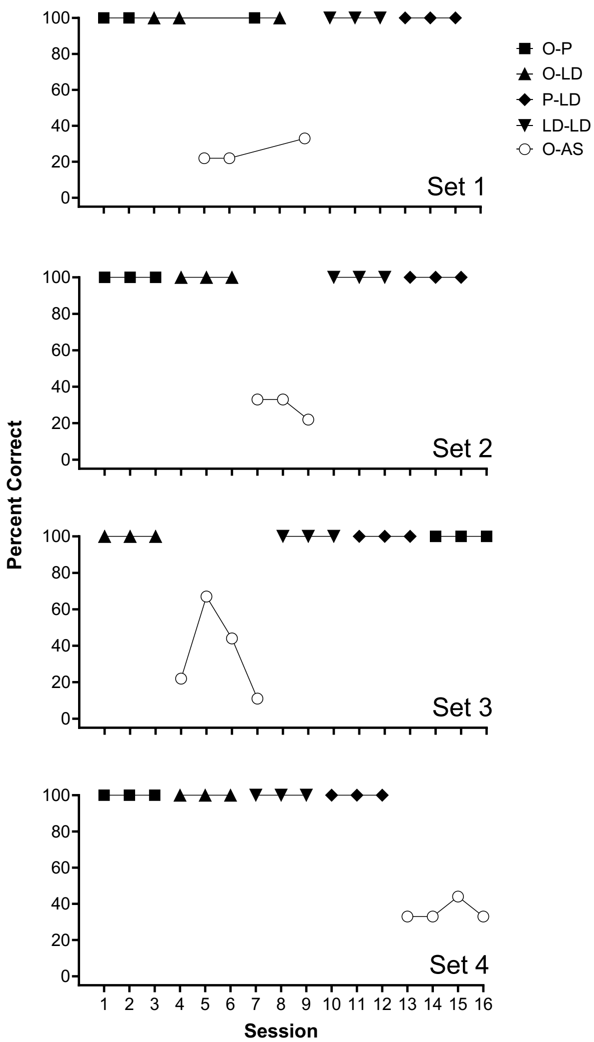
Procedural integrity was calculated for each session by dividing the total number of responses with the programmed consequence by the total number of responses and converting the result to a percentage. For Rowan, mean procedural integrity for 24 sessions across all conditions was 99% (range: 89% – 100%). For Carter, mean procedural integrity for 15 sessions across all conditions was 99% (range: 89% – 100%). For Francis, procedural integrity for 97 sessions across all conditions was 100%.

Rowan and Carter
Figures 3 and 4 display the accuracy scores across Sets 1 through 4 for Rowan and Carter, respectively. For all four sets, accuracy was 100% with O-P, O-LD, P-LD, and LD-LD and the mastery criterion of 89% or greater accuracy across two consecutive sets was met with each relation. For O-AS relations, there was no evidence of conditional control with any of the four sets.

Francis
Figure 5 displays performance on Sets 1 through 12 with Francis. For O-P relations, the mastery criterion was met for 11 out of 12 sets (mean sessions to mastery = 2.5) and was not met for one out of 12 sets (Set 5). For O-LD relations, the mastery criterion was met for nine out of 12 sets (mean sessions to mastery= 3.8) and was not met for three out of 12 sets (Sets 4, 5, and 12). For P-LD relations, the mastery criterion was met for 11 out of 12 sets (mean sessions to mastery= 2.3) and was not met for one out of 12 sets (Set 4). For LD-LD relations, the mastery criterion was met for 11 out of 12 sets (mean sessions to mastery= 2.0) and was not met for one out of 12 sets (Set 4). For O-AS relations, there was no evidence of conditional control with any of the 12 sets.

The mastery criterion for the O-P relation was met in fewer sessions than the O-LD relation in eight of eleven comparisons. For the other three comparisons, the mastery criterion was met

Figure 4. Accuracy of Responses During Experiment 1 for Carter



in two sessions for both relations. Data from Set 5 were excluded from this total, as Francis did not meet the mastery criterion for either O-P or O-LD.

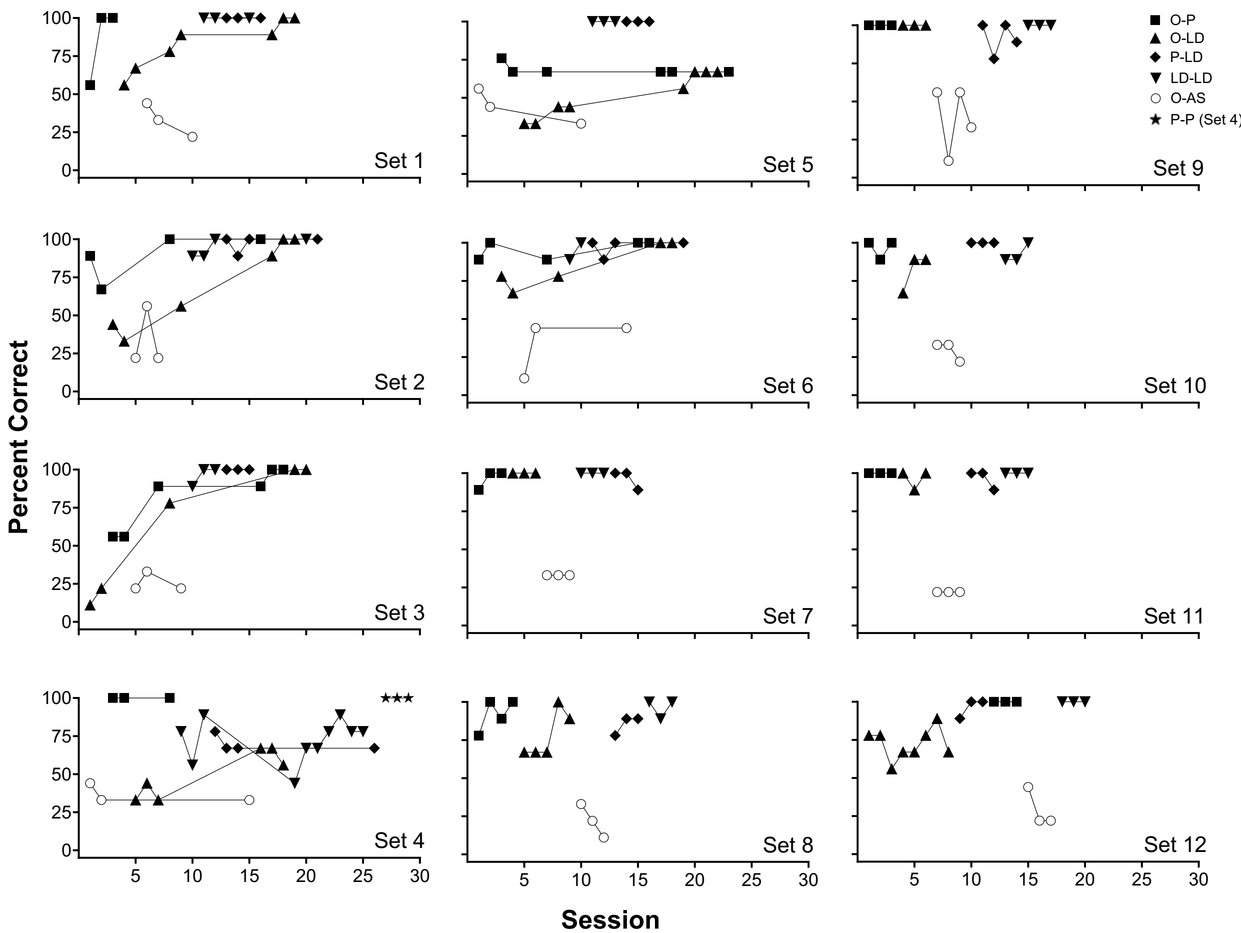
For those sets where mastery of relations was not demonstrated, there were persistent errors with specific stimuli and this pattern of differences in accuracy between photos and line drawings was also seen in other cases where line drawings served as comparison. For example, with Set 4, Francis did not meet the mastery criterion for P-LD, O-LD, or LD-LD and in each of these relations, the comparison stimuli were line drawings. Francis performed with high accuracy when the line drawing of the platypus was the S+, but accuracy was low during the trials in which the correct response was the line drawing of the polar bear or rhino. Additional testing was conducted to evaluate photo identity matching (P-P) with Set 4 to verify identity matching by photos and confirm simultaneous discrimination between the stimuli in the comparison array, and accuracy was 100% in these sessions. These examples highlight irregular error patterns that demonstrated poorer performance with some individual stimuli than others.

EXPERIMENT 2

In Experiment 1, Rowan and Carter readily matched objects to both line drawings and color photos with high accuracy. Francis, however, reliably acquired object-to-photo relations more quickly than object-to-line drawing relations. Although no effect of type of form could be detected with Rowan and Carter in the acquisition of object-to-picture matching, Francis's performance suggests that stimulus form could influence performance under certain conditions. Experiment 2 was designed to examine whether differences in stimulus form affected the emergence of untrained relations between objects and pictures.

Understanding variables that influence performance in matching-to-sample is important in special education contexts, where this procedure commonly is used to teach stimulus relations that are the basis for conceptual learning and communication (Carr & Felce, 2008). In some cases, this teaching can be structured in a way that produces the emergence of untrained relations between dissimilar stimuli. An example of this phenomenon is known as stimulus equivalence (Sidman, 2009). Stimuli form an *equivalence class* when, after learning a subset of possible relations within a socially defined class, untrained relations emerge between all members. Within such a class, stimuli become interchangeable, and can be thought of as sharing the same meaning. For example, an equivalence class could consist of a written word, a black-and-white picture, and an object. After being taught to choose the correct word and the correct picture when shown the object, the student may also respond correctly to tests for relations that were not directly trained. If an equivalence class has emerged, that student will select the object when shown the corresponding word or picture (*a symmetry relation*) and will select the correct comparison during tests for untrained relations between the word and the picture (*a transitivity relation*).

Figure 5. Accuracy of Responses During Experiment 1 for Francis



The emergence of equivalence relations has been extensively examined with a variety of stimulus types and individual learning histories (e.g., Arntzen et al., 2015; Hollis et al., 1986; Mackay, 1985; Osborne & Gatch, 1989; Sidman, 1971; Sidman & Tailby, 1982; Sidman et al., 1985; Stromer et al., 1996). In almost all of these studies, the stimuli within an equivalence class share no formal similarity. For example, Sidman (1971) reported on a teaching program used with an individual who had existing vocal imitation, oral naming, and picture-to-dictated-name matching repertoires. After conditional relations between dictated names and printed words were taught, untrained relations between printed words and pictures emerged, and the participant also read the words aloud. There were no overlapping stimulus features between the stimuli that could have resulted in the positive outcomes. Likewise, in a study with a different population and different stimulus types than Sidman (1971), Osborne and Gatch (1989) taught preschool children with profound hearing impairment to respond to relations between manually signed words, pictures, and printed words. Regardless of the order in which these relations were taught and tested, the untrained relations emerged.

In contrast to the studies that show the emergence of untrained relations across dissimilar stimuli, one more recent study provides evidence that overlapping stimulus features can increase the likelihood of emergent (untrained) relations between arbitrary stimuli in the same class. Arntzen et al. (2015) assessed the effect of visual similarity on the formation of equivalence classes with arbitrary stimuli. When all five stimuli were dissimilar, equivalence classes were not established, but when one picture shared features with a three-dimensional object (e.g., a house), equivalence class formation occurred in 80% of participants. This difference in outcomes between classes of stimuli with no overlapping features and classes with one relation that contains a shared feature shows that the form of visual stimuli can affect the emergence of untaught relations; in particular, the use of stimuli with more shared features can enhance the formation of equivalence classes. Building upon these findings, Experiment 2 investigated whether the form of visual stimuli influences the emergence of untrained equivalence relations with arbitrary symbols following the establishment of object-to-arbitrary symbol conditional discrimination. In other words, were equivalence relations more likely to emerge between photos and arbitrary symbols than between line drawings and arbitrary symbols?

Method

All participants, settings, stimuli, materials, and reliability procedures were the same as those described above in Experiment 1.

Research Design

In Experiment 2, a multiple probe design was used to evaluate differences between emergent relations with photos and line drawings following training and subsequent mastery of O-AS relations. The dependent variable was the accuracy on tests for emergent relations between each type of picture and arbitrary symbols, and the independent variables were the two different sample stimulus types assessed in the matching-to-sample arrangement with arbitrary symbols: photos (P-AS) and line drawings (LD-AS). A visual schematic of the trained and potential emergent relations for both Experiment 1 and Experiment 2 is provided in Figure 2.

Procedure

Preliminary Training. Mastery of all O-P, O-LD, P-LD, and LD-LD relations within each stimulus set was required prior to advancing to training with O-AS for that set. In most cases, these baseline relations were verified in Experiment 1. Exceptions are described below.

Training. Training was initiated for selected sets that did not meet the mastery criterion during Experiment 1 as well as with O-AS for sets that had met all prerequisites for Experiment 2. Training sessions consisted of nine matching-to-sample trials. Prior to each session, the researcher conducted a brief preference assessment with potential reinforcers for each participant. In general, social reinforcers were used with Rowan and preferred snacks were used with Francis. For Carter, a variety of modifications were made to the available reinforcers throughout the training phase (described below). A progressive delay prompt fading procedure was used to train arbitrary conditional matching with the following hierarchy: 0-s delay gestural prompt, 2-s delay gestural prompt, and no prompt. In all training phases, if the participant touched the correct comparison, the experimenter provided a participant-specific reinforcer, removed the stimuli from the table, and began the next trial. If the participant touched an incorrect comparison, the experimenter implemented an error-correction procedure: the stimuli were briefly removed, the trial was re-presented with a 0-s delay gestural prompt, the stimuli were removed from the table, and the next trial began. Correct responding resulted in neutral vocal affirmation of the response (e.g., *that's the polar bear*), but not tangible preferred items. With a few exceptions, described below, training sessions continued in this way until the mastery criterion of 89% or greater accuracy across two consecutive sessions was met.

Modifications. During the training phase, modifications were made to the training procedures for Carter and Francis following no observed progress under the procedures described above.

Differential Observing Response (DOR). A DOR procedure was used to increase the likelihood of attending to the comparison stimuli: prior to every O-AS training trial, the comparisons were arrayed on the table, but instead of presenting the object as the sample, a copy of each of the arbitrary stimuli was presented as the samples. After the participant completed this identity-matching task for all three stimuli, the object was presented as the sample. The comparison stimuli remained in the same position throughout each of these trials.

Differential Reinforcement of Independent Responses. In this phase, the DOR requirement was removed and differential reinforcement of independent responses was introduced. In addition to social praise, a frequently requested snack was provided following each correct, unprompted trial. This procedure was used in an attempt to decrease the likelihood of prompt dependence (Cividini-Motta & Ahearn, 2013).

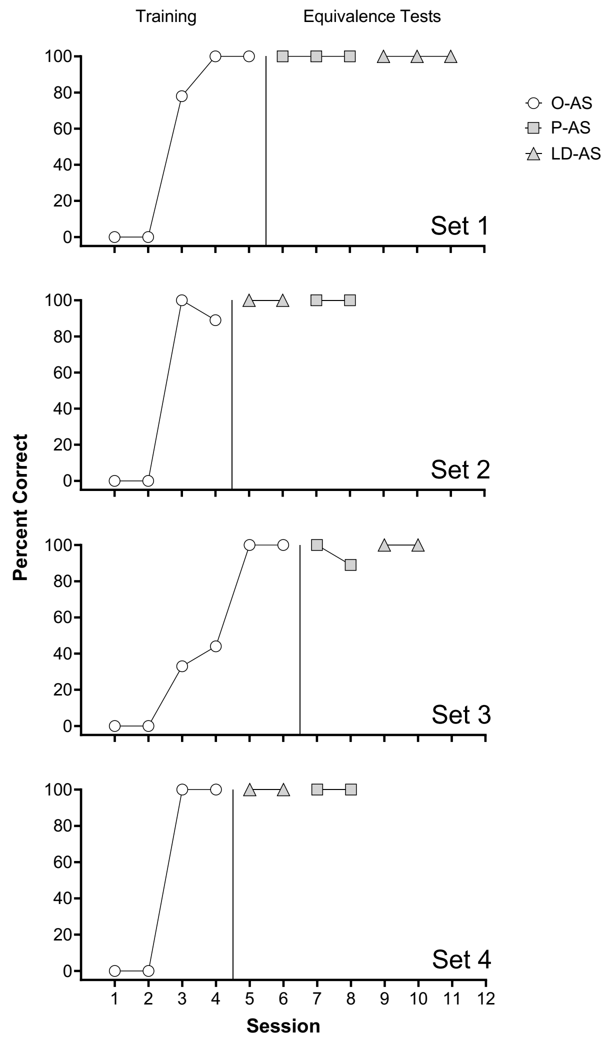
Removal of the 0-s Delay Prompt Training Step. To provide more opportunity for independent responding, the 0-s delay prompt was eliminated from the prompting hierarchy.

Differential Access to the Sample Object (Carter Only). The experimenter noticed that Carter tended to spend time manipulating the objects following each trial. The format of the task was changed so that rather than holding the object and placing it on a comparison stimulus, the experimenter held the object and required Carter to point to a comparison. Following correct responses, the participant was permitted to engage with the object for a few seconds.

Learning by Exclusion Procedure. An exclusion training procedure based on the methods used by Dixon (1977) was introduced in place of the progressive delay prompting procedure. For all exclusion sessions, the mastery criterion was 89% correct responding across two consecutive sessions. Reinforcement contingencies and error correction procedures were the same as in the training phase. In the first level of the exclusion training, two known stimuli (photos from previously established O-P relations) were presented as comparisons along with one unknown stimulus (arbitrary symbol). Each session consisted of six trials in which a known stimulus was the S+ (control trials), and three trials in which the unknown stimulus was the S+ (exclusion trials). Following mastery with the first unknown stimulus, a second unknown stimulus was introduced and trained with two known stimuli until the mastery criterion for this second comparison was met. Next, on the second level of the exclusion training, the comparison stimuli included two known stimuli and one of the two newly mastered stimuli, which alternated across trials. Training sessions again included three exclusion trials and six control trials. Following mastery at the second level, the comparison stimuli on the third level included one known stimulus and both of the newly mastered stimuli. Training sessions at this level included six trials in which the sample object corresponded with one of the two presented newly mastered stimuli (discrimination trials) and three control trials. At the fourth level, the third unknown stimulus was presented as a comparison alongside the two newly mastered stimuli. Training sessions included three exclusion trials and six discrimination trials.

Equivalence Tests. The equivalence test phase began once the O-AS relation had been mastered. The relations tested included photo-to-arbitrary symbol (P-AS) and line drawing-to-arbitrary symbol (LD-AS). During the equivalence test phase, no programmed differential consequences were provided for correct or incorrect responses. After each trial, the researcher recorded whether the response was correct or incorrect, the stimuli were removed from the table, and the next trial began. Preferred items or activities were delivered following every third or fourth trial along with praise related to session behavior (e.g., *nice sitting*). Two equivalence test sessions were conducted for each potential emergent relation.

Figure 6. Accuracy of Responses During Experiment 2 for Rowan



Replication. Following the equivalence test phase, the training and equivalence test phases were repeated with each remaining set using the procedures outlined above. The order of tests for the two types of equivalence relations was balanced across sets.

Results

Reliability

Interobserver agreement (IOA) for correct responses was calculated for each session by dividing the number of responses scored in agreement by the total number of trials and converting the result to a percentage. For Rowan, mean agreement for nine sessions across all conditions was 99% (range: 89% – 100%). For Carter, mean agreement for 42 sessions across all conditions was 99% (range: 78% – 100%). For Francis, mean agreement for 15 sessions across all conditions was 97% (range: 78% – 100%).

Procedural integrity was calculated for each session by dividing the total number of responses with the programmed consequence by the total number of responses and converting the result to a percentage. For Rowan, mean procedural integrity for nine sessions across all conditions was 99% (range: 89% – 100%). For Carter, mean procedural integrity for 42 sessions across all conditions was 99% (range: 78% – 100%). For Francis, procedural integrity for 15 sessions across all conditions was 100%.

Rowan

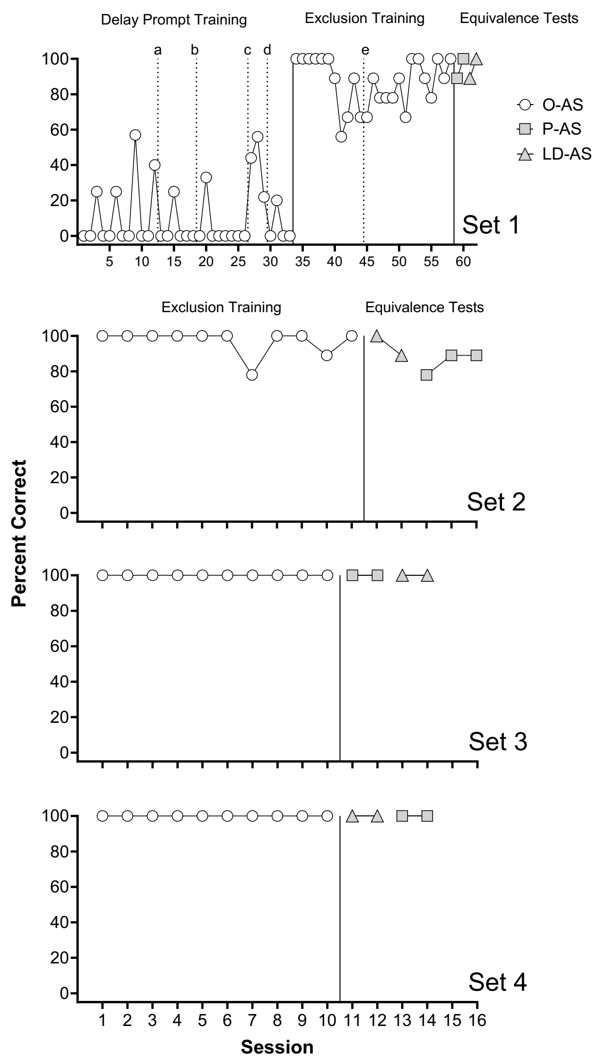
Figure 6 displays the accuracy scores for Rowan across Sets 1 through 4. During the training phase with O-AS for Set 1, the mastery criterion was met after five sessions of training. Next, the equivalence test phase for Set 1 was conducted, and 100% correct responding was observed with both the P-AS and LD-AS relations. With Set 2, the mastery criterion was met during the training phase for O-AS after four sessions. During the subsequent equivalence test phase, 100% correct responding was observed with both the P-AS and LD-AS relations. During training with Set 3, the mastery criterion for O-AS was met after six sessions. In the equivalence test phase, 89% or greater accuracy was observed with both the P-AS and LD-AS relations. For Set 4, the mastery criterion for O-AS was met after four sessions during the training phase. During the equivalence test phase, 100% accuracy was observed with P-AS and LD-AS relations.

Carter

Figure 7 displays the accuracy scores for Carter across Sets 1 through 4. During the training phase with O-AS for Set 1, no progress was observed in 12 training sessions, so the differential observing response procedure was introduced. No progress was observed in an additional six sessions. The differential observing response was removed, and differential reinforcement was introduced for independent responses. No progress was observed in an additional eight sessions. Next, the 0-s delay prompt step was eliminated, and access to an iPad was programmed. No progress was observed in three additional sessions. After observing that Carter expressed interest in the sample objects, the response was changed from putting the object on the picture to a point response, and access to the object was provided contingent on correct responding. After four additional sessions with no progress observed, an exclusion teaching procedure was introduced.

Accuracy on the first two levels of the exclusion procedure was 100%. Following five sessions with no progress on the third level,

Figure 7. Accuracy of Responses During Experiment 2 for Carter



Note. a) Introduced DOR. b) Removed DOR, introduced differential reinforcement of independent responses. c) Removed 0-s delay prompt step, access to iPad following independent responses. d) Differential access to the sample object. e) Added edible rewards contingent on correct responding for exclusion and discrimination trials.

additional differential reinforcement (a highly preferred edible) contingent on correct responding on exclusion and discrimination trials was introduced. Performance gradually improved until the mastery criterion was met in nine sessions. On the fourth level with Set 1, the mastery criterion was met in four sessions. Following the training phase, the equivalence test phase was conducted with Set 1. The P-AS relation was tested first, and 89% or greater accuracy was observed across a two-session block. Next, the LD-AS relation was tested, and again 89% or greater accuracy was observed.

With Set 2 for Carter, O-AS training was initiated with the exclusion procedures described above. The first two levels of the exclusion training were mastered in the minimum number of sessions with 100% accuracy. During the third level, the mastery criterion was met in three training sessions. On the fourth level, the mastery criterion was met in two sessions. During the equivalence test phase, the LD-AS relation was tested first, and accuracy across a two-session block was 89% or greater. The P-AS relation was tested next, and three sessions were required to meet the mastery criterion. For Sets 3 and 4, O-AS training was again conducted using the exclusion procedures. All four levels of the exclusion training were mastered in the minimum number of sessions, with 100% accuracy. During the equivalence test phase, accuracy for the P-AS and LD-AS tests was 100%.

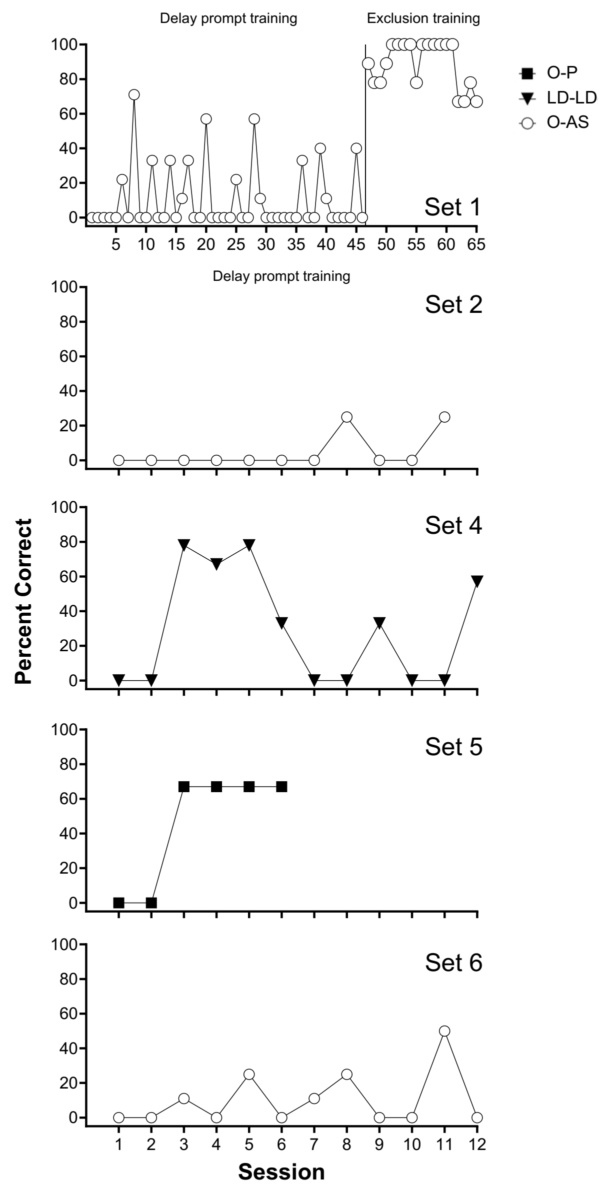
Francis

Figure 8 displays the accuracy scores for Francis for each of the sets that were exposed to training procedures during Experiment 2. Progressive delay prompting was used to teach LD-LD for Set 4, O-P for Set 5, and O-AS for Sets 1, 2, and 6. For Set 4, no progress was observed in 12 sessions of LD-LD. For O-P with Set 5, no progress was observed in six sessions. With O-AS training for Set 1, no progress was observed across 46 sessions. Each of the modifications described above were implemented, with the exception of restricting access to the sample object. With O-AS for Sets 2, and 6, no progress was observed in 11 and 12 sessions, respectively. Throughout training with each of these sets, Francis exhibited prompt dependency, often waiting up to two minutes before responding on trials on which there was no prompt. Additional training was conducted using exclusion training procedures with O-AS for Set 1. Francis progressed through the first two levels of the exclusion training. On the third level he exhibited prolonged response latencies and no evidence of conditional control was observed across discrimination trials. Given Francis's prompt dependency and the lack of progress across all training sessions under various procedural modifications, Francis's participation in Experiment 2 was discontinued during the training phase.

DISCUSSION

The purpose of the present investigation was to assess whether the form of visual stimuli

Figure 8. Accuracy of Responses During Experiment 2 for Francis



produced differential responding with object-to-picture matching and the subsequent emergence of untrained equivalence relations in young men with autism spectrum disorder. In Experiment 1, participants were tested on matching-to-sample relations between objects and photos, objects and line drawings, objects and arbitrary symbols, and photos and line drawings, as well as line drawing identity matching. In Experiment 2, participants learned object-to-arbitrary symbol relations and then equivalence tests were conducted to assess the emergence of relations between photos and line drawings and corresponding arbitrary symbols.

For Rowan, no effect of the form of the visual stimuli on the acquisition of these relations or emergence of untrained equivalence relations was observed. Each relation was mastered with 100% accuracy in Experiment 1, except for the O-AS relation. Training with the progressive delay prompting procedure in Experiment 2 resulted in rapid mastery of the O-AS relation with all four sets. Following mastery of the O-AS relation, Rowan passed tests for emergence of untrained P-AS and LD-AS relations (see Figure 6).

Carter's results, like Rowan's, also do not indicate an effect of the form of visual stimuli on acquisition of these relations or emergence of untrained equivalence relations. Each relation was mastered with 100% accuracy in Experiment 1, except for the O-AS relation. Although training with various procedural modifications did not produce mastery of the O-AS relation with the first set of stimuli in Experiment 2, an exclusion training procedure ultimately resulted in mastery. With each subsequent set, mastery of the O-AS relation under exclusion training conditions was demonstrated more efficiently, resulting in errorless learning with the third and fourth sets. Following mastery of the O-AS relation, the emergence of equivalence relations was demonstrated with the P-AS and LD-AS relations (see Figure 7). With three of the sets, 100% accuracy was observed in two sessions with both the P-AS and LD-AS relations. In one case, emergence of the LD-AS relation was seen in fewer sessions than the P-AS relation (two and three sessions, respectively).

With Francis, it appears that the form of visual stimuli had an effect on accuracy of responding with object-to-picture relations. Mastery of several relations emerged following multiple exposures in the absence of differential reinforcement for correct responding in Experiment 1. As differences were observed in response accuracy between O-P and O-LD relations with the initial four sets, Experiment 1 was replicated with several additional sets in order to verify these findings. In eight out of eleven opportunities for Francis, the mastery criterion was met after fewer assessment exposures with the O-P relation than the O-LD relation (see Figure 5). With the remaining three stimulus sets, mastery of both relations was demonstrated in the minimum number of sessions. With the 12th tested stimulus set, neither relation was mastered following eight sessions under these conditions.

During the training phase in Experiment 2 for Francis, no progress was observed despite the use of a variety of remediation procedures. Francis also exhibited prompt dependency under the gesture prompt training conditions; when these prompts were faded and later absent under the exclusion training conditions, Francis exhibited long response latencies of up to two minutes. Subsequently, Francis's participation in Experiment 2 was discontinued and the emergence of untrained equivalence relations was not evaluated.

Several idiosyncrasies were noted in Francis's responses to both pictures and objects (i.e., the line drawings of the polar bear and the rhino, the turtle object). Sidman (2009) suggests that, even when participants perform a discrimination accurately, "the stimulus aspects that control their behavior may not be the same as those specified by the experimental contingencies." Under the present experimental conditions, it could not be determined which stimulus features controlled responding, and it is possible that Francis's responding to these particular stimuli was under the control of features not accounted for in the manipulation of the photos to create the black-and-white line drawings.

Overall, the results of this investigation suggest that, for some individuals with autism, the form of visual stimuli affects accuracy of responding with object-to-picture relations. The pictures used in the present study included photos, line drawings, and arbitrary symbols. While the photos shared many stimulus features with their corresponding three-dimensional objects, including color, shape, and shading, the line drawings shared only the shape defined by a black outline and the arbitrary symbols did not share any definable features with the other corresponding stimuli in the set. For Francis, the relations involving stimuli with the highest degree of similarity to the corresponding object (photos) were mastered the most efficiently. Relations involving stimuli with reduced similarity (line drawings) were mastered less efficiently, and relations involving stimuli with no similarity to the corresponding object were not mastered even with continued remedial training.

A 1982 study by Hurlbut and colleagues described similar findings with three individuals with cerebral palsy. Colored line drawings and an ideographic writing system called Blisssymbols were compared in terms of mastery of picture selection given a vocal cue, generalization to untrained stimuli, and spontaneous usage in the natural environment. The colored line drawings produced mastery four times faster, occasioned more correct responses to untrained stimuli, and accounted for nearly all of the spontaneous usage in the natural environment. These results similarly suggest that relations involving stimuli with some similarity to the corresponding object produce more efficient learning for some individuals than relations involving no similarity, and the present study supports the interpretation that the degree of similarity might play an important role in this effect.

Implications

The current study extends the research literature on teaching stimulus relations to children with autism spectrum disorder. For one individual, the visual similarity between the picture and the corresponding object affected accuracy of responding for matching-to-sample relations. Several researchers have argued that children with autism spectrum disorder do not learn stimulus relations in the same way as children without disabilities (Carr & Felce, 2008; Dixon, 1981; Hartley & Allen, 2015; Higbee et al. 1999; Lionello-DeNolf & McIlvane, 2016; and Nguyen et al., 2009), and the current study described differences in learning stimulus relations involving pictures with varying similarity to the corresponding object among a small sample of individuals with autism spectrum disorder. As learning relations between pictures and corresponding objects is critical to the use of picture exchange systems in aided AAC systems, these results have implications for improving our technologies for teaching these critical relations. For some individuals, such as Francis, the use of full-color photographs of real-world objects and people in place of generic hand- or computer-drawn graphics and black-and-white line drawings may contribute to more efficient communication.

Fagot et al. (2000) proposed three main modes of picture perception to explain formation of stimulus relations with pictorial stimuli. The *independence* mode refers to a situation in which the individual responds to pictures as combinations of features or patterns, regardless of the depicted image (e.g., selecting a red picture when presented with a picture of an apple instead of the apple itself). The *confusion* mode refers to a situation in which the individual responds to the picture as the depicted object (e.g., attempting to eat a picture of an apple). The *equivalence* mode refers to a situation in which the individual discriminates the photo from the object and perceives the image as a referent to the object (e.g., indicating a picture of an apple to request one from a communication partner). Sidman (1994) suggested that this framework proposed by Fagot et al. usefully categorizes human performance on discrimination tasks involving pictures depicting objects. In the present investigation, photos and line drawings both shared identifiable visual features with the depicted objects, and participants could match objects to pictures by features without attending to the entire depicted image. Francis's responding was likely under the control of particular visual features of the pictures, and as photos have more features in common with the objects than the line drawings, photos produced more efficient mastery than line drawings. This would be consistent with the independence mode of picture discrimination. However, the arbitrary stimuli shared no discriminable features with the objects, and performing the discrimination accurately required responding in the equivalence mode. Following training, both Carter and Rowan responded to the arbitrary stimuli as referents to the objects, while Francis did not accurately perform this discrimination. Carter and Rowan also demonstrated mastery of the P-AS and LD-AS relations with minimal differences, including the emergence of untrained class-consistent stimulus relations. For individuals who have difficulties performing conditional discriminations, increasing the number of visual features in common with the depicted object can improve the accuracy of discriminations that depend on such features. When applied to communication tools, the use of stimuli with greater visual similarity to the target might allow individuals with minimal verbal repertoires to interact verbally with others more effectively.

Limitations and Future Directions

Two features of the present study that may limit generality warrant discussion. First, this study was translational and did not directly evaluate performance in an educationally or therapeutically relevant context. Although the results carry implications for stimulus selection in aided AAC systems, there was no systematic manipulation of these stimuli in the context of AAC use in this study. Furthermore, aided AAC systems serve a wide array of communicative purposes, not all of which have an identifiable three-dimensional correlate in the natural environment. However, for many individuals with severely limited communication skills, it is possible that a smaller vocabulary built on photos of targets in their natural environment could more effectively permit communicative exchanges than a more extensive vocabulary built on a mix of photos and abstract stimuli requiring precise individualized training with each target.

Second, Francis's responses to several of the experimental stimuli suggest faulty stimulus control, but the particular features controlling this responding could not be identified within the parameters of the present study. It is possible that more precise identification of these critical discriminative features could inform a more detailed analysis of the aspects of each type of stimulus which resulted in more efficient mastery of discrimination tasks with photos than line drawings (e.g., color, shading, shape). Such an analysis could also direct steps to be taken in the design of an individualized stimulus fading curriculum to gradually transfer stimulus control to pictures with minimal similarities to the target by systematically fading the critical features. Sidman and Stoddard (1966) recommended the use of such fading procedures to teach visual discriminations, suggesting that program failures occur when steps are too large. However, McIlvane et al. (2016) cautioned that stimulus fading procedures with small, incremental steps may actually shape "increasingly overselective attending" to common features that are preserved during shaping. Future research will need to evaluate the best methods for teaching such conditional discriminations when it is necessary to establish functional relations involving stimuli with minimal visual similarity to the other class members.

Several avenues for future research are invited by the findings of the current study. An evaluation of the effects of visual similarity on efficiency of mastering new learning targets on aided AAC systems in an educational or therapeutic intervention would inform the significance of the present findings in an applied context. Such an evaluation could also be extended to assess the extent to which any such effects can be applied to learning targets without an identifiable physical correlate in the natural environment. Future research could also explore the particular features contributing to more efficient mastery with photos than with line drawings. The results of Hartley and Allen's (2015) investigation suggested that color is a critical feature that enhances stimulus discrimination for children with autism spectrum disorder. In the current study, however, there was an instance of a failure to match full-color photos to corresponding objects (i.e., the turtle photo and the wombat photo). It is possible that Francis's responding was controlled by other stimulus features. Identification of these features could lead to research on the best methods for teaching conditional discriminations involving minimal visual similarity without increasing stimulus overselectivity. Finally, future research could help to develop more efficient analytical tools to identify learners who might benefit from the use of stimuli with greater visual similarity to the target. Although adaptation of visual stimuli as a remedial step when learners do not demonstrate efficient mastery of new targets is a possible recommendation for applied practice, the use of communicative tools better suited to the particular needs of each individual at the onset of instruction may help to improve outcomes related to functional communication and quality of life. Such research could be a valuable step towards identifying best practices for instruction of children with autism spectrum disorder who have difficulty learning to communicate effectively with the people around them.

CONCLUSION

The results of this investigation suggest that the form of visual stimuli influences the establishment of stimulus relations involving pictures with varying similarity to the corresponding object for some individuals with autism. For one participant, accuracy of responding was highest with pictures that had the highest degree of similarity to the corresponding object (i.e., color photos). For some individuals, the use of photos of real-world objects and people in place of line drawings may contribute to more effective communication

REFERENCES

Arntzen, E., Nartey, R. K., & Fields, L. (2015). Enhanced equivalence class formation by the delay and relational functions of meaningful stimuli. *Journal of the Experimental Analysis of Behavior*, *103*(3), 524–541. https://doi.org/10.1002/jeab.152

Carr, D., & Felce, J. (2008). Teaching picture-to-object relations in picture-based requesting by children with autism: A comparison between error prevention and error correction teaching procedures. *Journal of Intellectual Disability Research*, *52*(4), 309-317. http://doi.org/10.1111/j.1365-2788.2007.01021.x

Cividini-Motta, C., & Ahearn, W. H. (2013). Effects of two variations of differential reinforcement on prompt dependency. *Journal of Applied Behavior Analysis*, *46*(3), 640–650. http://doi.org/10.1002/jaba.67

Dickson, C. A., MacDonald, R. P. F., Mansfield, R., Guilhardi, P., Johnson, C., & Ahearn, W.H. (2014). Social validation of the New England Center for Children-Core Skills Assessment. *Journal of Autism and Developmental Disorders*, *44*, 65–74. http://doi.org/10.1007/s10803-013-1852-5

Dixon, L. S. (1977). The nature of control by spoken words over visual stimulus selection. *Journal of the Experimental Analysis of Behavior*, *27*(3), 433–442. https://doi.org/10.1901/jeab.1977.27-433

Dixon, L. S. (1981). A functional analysis of photo-object matching skills of severely retarded adolescents. *Journal of Applied Behavior Analysis*, *14*(4), 465–478. https://doi.org/10.1901/jaba.1981.14-465

Fagot, J., Martin-Malivel, J., & Depy, D. (2000). What is the evidence for an equivalence between objects and pictures in birds and nonhuman primates? In *Picture Perception in Animals* (pp. 295–320). Psychology Press Ltd.

Green, G., Mackay, H. A., McIlvane, W. J., & Saunders, R. R. (1990). Perspectives on relational learning in mental retardation. *American Journal on Mental Retardation*, *95*(3), 249–259. Retrieved from: https://psycnet.apa.org/record/1991-10210-001

Hartley, C., & Allen, M. L. (2015). Symbolic understanding of pictures in low-functioning children with autism: The effects of iconicity and naming. *Journal of Autism and Developmental Disorders*, *45*(1), 15–30. https://doi.org/10.1007/s10803-013-2007-4

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*(1), 63–72. https://doi.org/10.1016/s0891-4222(98)00032-8

Hollis, J., Fulton, R., & Larson, A. (1986). An equivalence model for vocabulary acquisition in profoundly hearing-impaired children. *Analysis and Intervention in Developmental Disabilities*, *6*(4), 331–348. https://doi.org/https://doi.org/10.1016/S0270-4684(86)80013-1

Hurlbut, B. I., Iwata, B. A., & Green, J. D. (1982). Nonvocal language acquisition in adolescents with severe physical disabilities: Blissymbol versus iconic stimulus formats. *Journal of Applied Behavior Analysis*, *15*(2), 241–258. https://doi.org/10.1901/jaba.1982.15-241

Lionello-DeNolf, K. M., & McIlvane, W. J. (2016). Developing 3D-2D functional equivalence in a child with autism and severe intellectual disability: An exploratory study. *Experimental Analysis of Human Behavior Bulletin*, *30*, 6–16.

Mackay, H. A. (1985). Stimulus equivalence in rudimentary reading and spelling. *Analysis and Intervention in Developmental Disabilities*, *5*(4), 373–387. https://doi.org/10.1016/0270-4684(85)90006-0

McIlvane, W., Dube, W., Green, G., & Serna, R. (1993). Programming conceptual and communication skill development: A methodological stimulus-class analysis. In *Communication and language intervention series*, Vol. 2. Enhancing children's communication: Research foundations for intervention (pp. 243–285). Paul H. Brookes Publishing.

McIlvane, W. J., Dube, W. V, Kledaras, J. B., Iennaco, F. M., & Stoddard, L. T. (1990). Teaching relational discrimination to individuals with mental retardation: some problems and possible solutions. *American Journal of Mental Retardation*, *95*(3), 283–296. Retrieved from: https://pubmed.ncbi.nlm.nih.gov/2261161/

McIlvane, W. J., Gerard, C. J., Kledaras, J. B., Mackay, H. A., & Lionello-DeNolf, K. M. (2016). Teaching stimulus-stimulus relations to minimally verbal individuals: reflections on technology and future directions. *European Journal of Behavior Analysis*, *17*(1), 49–68. https://doi.org/10.1080/15021149.2016.1139363

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 of Developmental Disabilities*, *15*(1), 53–64. Retrieved from: https://pmc.ncbi.nlm.nih.gov/articles/PMC3607582/

Osborne, J. G., & Gatch, M. B. (1989). Stimulus equivalence and receptive reading by hearing-impaired preschool children. *Language, Speech, and Hearing Services in Schools*, *20*(1), 63–75. https://doi.org/10.1044/0161-1461.2001.63

Romski, M. A., & Sevcik, R. A. (1997). Augmentative and alternative communication for children with developmental disabilities. *Mental Retardation and*

Developmental Disabilities Research Reviews, *3*(4), 363–368. https://doi.org/10.1002/(SICI)1098-2779(1997)3:4<363::AID-MRDD12>3.0.CO;2-T
Shafer, E. (1993). Teaching topography-based and selection-based verbal behavior to developmentally disabled individuals: Some considerations. *The Analysis of Verbal Behavior*, 11, 117–133. https://doi.org/10.1007/BF03392892
Sidman, M. (1971). Reading and auditory-visual equivalences. *Journal of Speech and Hearing Research*, *14*(1), 5–13. https://doi.org/10.1044/jshr.1401.05
Sidman, M. (1994). Equivalence relations and behavior: A research story. Authors Cooperative.
Sidman, M. (2009). Equivalence relations and behavior: An introductory tutorial. *The Analysis of Verbal Behavior*, *25*, 5–17. https://doi.org/10.1007/BF03393066
Sidman, M., Kirk, B., & Willson-Morris, M. (1985). Six-member stimulus classes generated by conditional-discrimination procedures. *Journal of the Experimental Analysis of Behavior*, *43*(1), 21–42. https://doi.org/10.1901/jeab.1985.43-21
Sidman, M., & Stoddard, L. T. (1966). Programming perception and learning for retarded children. *International Review of Research in Mental Retardation*, *2*, 151–208. http://doi.org/10.1016/S0074-7750(08)60206-2
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*(1), 5–22. https://doi.org/10.1901/jeab.1982.37-5
Sidman, M., Wynne, C. K., Maguire, R. W., & Barnes, T. (1989). Functional classes and equivalence relations. *Journal of the Experimental Analysis of Behavior*, *52*(3), 261–274. https://doi.org/10.1901/jeab.1989.52-261