Published in final edited form as:

Psychol Rec. 2014 December 1; 64(4): 731–742. doi:10.1007/s40732-014-0086-z.

Rapid Teaching of Arbitrary Matching in Individuals with Intellectual Disabilities

Greg Morro,

Eunice Kennedy Shriver Center

Harry A. Mackay, and

Eunice Kennedy Shriver Center

Michael T. Carlin

Rider University

Abstract

This research extended to arbitrary matching-to-sample procedures a method that was successful in rapidly establishing identity matching in children with and without intellectual disabilities (Mackay et al., 2002). The method involves increasing the number of identical comparison stimuli in a choice array in order to create a homogenous background that makes the target more salient, thus likely to prompt selection. The number of comparison stimuli then is faded systematically contingent on accurate responding. This method unites cognitive research on visual search and behavior analytic research on conditional stimulus control. Two experiments examined use of the method to teach arbitrary relations between visual stimuli (numerals and colors and their printed names) and between visual and auditory stimuli (e.g., numerals and colors and their dictated names). Results demonstrated the generality of the method to symbolic matching. This finding is important for conceptual reasons and for its relevance to special education.

The conditional discriminations of symbolic matching-to-sample procedures (e.g., selecting from sets of pictures (numerals, forms) the ones corresponding to their dictated or printed names) are used in a broad variety of contexts, including analyses of language and cognitive development (e.g., Burgoyne et al., 2012; Hayes, Barnes-Holmes, & Roche, 2001; Lovaas, 1987; Mackay, 2012; Mackay & Fields, 2009; Mackay, Wilkinson, Farrell, & Serna, 2011; Wilkinson & McIlvane, 1997), and the assessment of intellectual functioning (e.g., Dunn & Dunn, 1981; Kaufman & Kaufman, 2004; Leiter, 1979). Further, they provide critical aspects of pre-academic and later instruction in reading and number skills (e.g., Byrne, MacDonald & Buckley, 2002; de Souza et al., 2009; Mackay, 1991; Mackay, Kotlarchyk & Stromer, 1997; Maurice, Green & Luce, 1996; Maydack, Stromer, Mackay, & Stoddard, 1995; Stromer, Mackay & Stoddard, 1992). The current research extends to arbitrary or symbolic matching-to-sample procedures a method that Mackay, Soraci, Carlin, Dennis and Strawbridge (2002) used to establish identity matching in children with and without

intellectual disabilities. The general method brings together cognitive research on visual search and behavior analytic research on stimulus control and is important for conceptual and practical reasons.

Previous studies have shown that when searching multi-element arrays for a predefined target, individuals with intellectual disabilities are sometimes less efficient than typically developing peers (Carlin, Soraci, Goldman, & McIlvane, 1995; Munir, Cornish, & Wilding, 2000; Scerif, Cornish, Wilding, Driver, & Karmiloff-Smith, 2004). For highly salient targets (e.g., triangle among circles), individuals with and without intellectual disabilities search very rapidly and without additional time being necessary as number of comparisons is increased (i.e., participants demonstrate “parallel search”). Group differences commonly emerge when target-comparison disparities are reduced (e.g., circle among pentagons). As target-comparison disparity is reduced those with intellectual disabilities often resort to serial search (i.e., one stimulus at a time), sometimes in conditions that those without disabilities still search highly efficiently. However, some experiments have identified conditions that facilitate performance or that can be used to guide attention to particular subsets of elements in complex visual arrays (e.g., Carlin, Soraci, Dennis, Strawbridge, & Chechile, 2002; Soraci et al., 1991). Several studies have demonstrated the use of color to speed target selection in visual search (Wilkinson, Carlin & Thistle, 2008; Wilkinson, Carlin, & Jagaroo, 2006). Research reported by Soraci et al. (1987) is particularly important for the research described here. Soraci and his colleagues manipulated the visual arrays used in oddity tasks and produced rapid acquisition of oddity performance in children with mental ages below five years who often experience difficulty with three-stimulus oddity tasks (e.g., selection of a square displayed together with two circles). To facilitate detection of the odd target stimulus (i.e., the positive stimulus or S+) it was presented among eight other stimuli that were identical to one another rather than only one S-. The intent was to enhance the similarity-difference relation so the odd stimulus would “popout” and be more likely to be selected. The young participants, who previously had failed the three-stimulus oddity task, rapidly demonstrated oddity performances when the nine-stimulus displays were introduced. After this intervention, oddity performance then continued after the three-stimulus display was presented again.

Mackay et al.’s (2002) novel procedure for establishing identity matching of forms used the same display manipulation on initial training trials to direct attention to and prompt selection of the target comparison stimulus on training trials. The prompt then was faded out gradually across trials to allow development of conditional control by the sample stimuli that changed from trial to trial. Participants were children who did not perform identity matching of form stimuli. Sixteen were typically developing, 3–5 year-olds (mean estimated mental age = 5.20 years as assessed using the Peabody Picture Vocabulary Test-Revised (PPVT-R; Dunn & Dunn, 1981) and 12 had various intellectual disabilities (aged 11–20 years, mean PPVT-R = 5.52 years). The method is outlined in Figure 1. In early trials, the correct comparison on each trial matched the sample physically and differed from the other eight comparisons, which were identical to one another. The oddity condition in the comparison display was used to increase the probability that the target would be attended and selected even on the first trial. All selections of the odd forms in the comparison displays then were reinforced, thus establishing the first instances (Ray & Sidman, 1970; Skinner 1957) of the

necessary simultaneous discriminations between the forms presented concurrently as comparison stimuli. Next, across correct trials, the number of incorrect comparisons was reduced systematically until only one remained. This fading procedure was designed to reduce gradually the effectiveness of the oddity cue in controlling responding in these simultaneous discriminations and instead allow development of the successive discriminations among samples required in matching-to-sample tasks (see Mackay, 1991).

The method succeeded with 75% of the participants, often yielding rapid almost errorless identity matching-to-sample performances. This outcome improved substantially upon the success rate of 32% reported previously (Serna, Dube, & McIlvane, 1997) after training with different procedures. Other teaching procedures have established the critical complex discriminations in many difficult-to-teach children. However, they require many sessions (Dube & Serna, 1998; Pilgrim, Jackson, & Galizio, 2000). It is important then to continue to research new procedures that may help to establish such performances quickly and efficiently, with few errors.

Mackay et al.'s (2002) procedure may be viewed as an addition to the existing set of prompting and fading methods used to teach a variety of tasks to children with intellectual disabilities. Commonly used methods include adding and then gradually removing verbal cues, manual guidance, gestural and imitative prompts, and enhancement of features of the correct stimuli (e.g., brightness and size) in tasks to be trained. However, the added cues, which are intended to increase the likelihood that appropriate responses will occur and be reinforced, may prove difficult to remove (see MacDuff, Krantz, & McClannahan, 2001, for a review focused on teaching those with autism.). Notably, Mackay et al.'s (2002) procedure established the critical form discriminations without requiring the addition of cues (manual guidance, gestures etc.) or their enhancement. Furthermore, the research by Soraci et al. (1987) suggests a method that appears to establish the prerequisite oddity performance if necessary.

The experiments reported here extend Mackay et al.'s (2002) research by examining the use of their method to establish arbitrary relations between sample and comparison stimuli (i.e., symbolic matching) rather than identity relations. In particular, the research, which has strong educational relevance, examined training of arbitrary relations between visual stimuli (e.g., numerals and their printed names) and between visual and auditory stimuli (e.g., numerals and their dictated names). The work is closely related to extensive research concerning development of relations of equivalence between such stimuli (e.g., Mackay et al., 1997; Mackay, 1991; Mackay & Sidman, 1984; Maydack, et al., 1995; see Mackay, Stromer & Serna, 1998 and Sidman, 1994 for broad overviews). In this respect, the interest lies in whether the new procedure quickly teaches participants with lengthy histories of failure new symbolic matching performances that may integrate readily with other performances to yield a more extensive repertoire without additional training. Furthermore, Experiment 1 uses visual training to establish equivalence classes of numerals and their printed names rather than the auditory-visual training used by Mackay et al. (1997) to do so. Experiment 2 was a single-subject follow-up study that used a variant of the "popout" procedure but with auditory samples and visual comparisons.

Experiment 1

This experiment had two parts. Initially, a pretest-posttest design was used to examine the use of the ‘popout’ training method to establish arbitrary visual relations between three numerals (1, 2, 6) and their printed names. Relations between these visual stimuli and their oral names also were assessed before and after the training. Then, in a systematic replication, the experiment was repeated with another three numerals (4, 5, 9) and their printed and oral names.

Method

Participants—Four experimentally naive participants (two males, two females) who already had diagnoses (consistent with the Diagnostic and Statistical Manual of Mental Disorders 1994) of autism or Pervasive Developmental Disorder (PDD) were recruited from a school serving individuals with developmental disabilities. Their chronological ages ranged from 16 years to 18.2 (years.months) and their estimated mental ages as assessed by the Peabody Picture Vocabulary Test-Revised (PPVT-R; Dunn & Dunn, 1981) ranged from 3.0 to 6.1, characteristics resembling those of participants in Mackay et al., (2002). Recall, however, that participants in that earlier research did not perform identity matching before they received training. Discussions with the participants’ teachers indicated that the current participants matched identical stimuli, identified fewer than 10 sight words, had learned to name numerals, and that learning the relations between numbers and their names (printed and spoken), while not currently an educational goal, was an appropriate, positive activity. Pretests that evaluated relevant specific performances with experimental stimuli are described later. Participation was voluntary. Informed consent was obtained from guardians.

Apparatus and setting—An IMac (Apple ©) computer fitted with a MicroTouch GoldStar touch-sensitive screen (MicroTouch Systems Inc.®) was used to control stimulus presentations and other experimental events and to record responses automatically. The software was that described by Dube and Hiris (1997).

Single sessions took place typically every weekday in a quiet room at the participant’s school. The participant sat facing the computer, which was placed on a table at eye level. The experimenter sat beside the participant who could not observe the experimenter and the apparatus simultaneously. Prior to the first session the experimenter demonstrated the touchscreen operation and the matching-to-sample procedure with stimuli not used in later experiments. Minimal (e.g., “Touch it” in reference to a sample) or no verbal instructions then were given at the start of sessions.

Stimuli—Stimuli for Experiment 1 were printed numerals (1, 2 and 6 initially; 4, 5, and 9 in the replication; Times, 120 point) and the corresponding printed (Times, 48 point) and dictated names. The printed names in each set were the same length. The dictated names were recorded and presented as samples via the computer.

The stimuli displayed as comparisons were presented within an invisible 3×3 grid that measured 13×13 cm, just as illustrated in Figure 1. Visual samples (printed numerals and

words) were centered above this comparison array. On trials with dictated samples, a black square was displayed in the sample area.

Procedure

General: Each match-to-sample trial began with the presentation of a sample stimulus. For visual tasks, touching the sample resulted in immediate presentation of an array of comparison stimuli. From two to nine comparison stimuli were displayed at different training and test stages. The positions of the comparison stimuli changed unsystematically from trial to trial and those where no comparison was displayed were white. The sample and all comparisons remained visible until a response occurred to a comparison. On trials with dictated samples, the comparison array was added after a touch to the black square. The dictated sample repeated every 1 s throughout the trial.

During training, praise accompanied by a display of colored stars and a brief melody (duration 1 s) was presented after selection (a touch) of the correct comparison. The subsequent trial began following an additional inter-trial interval (ITI) of 1.5 s during which the screen was white and screen touches had no effect. An incorrect selection produced a 1 s timeout with a blank screen. The ITI then followed. Some sessions were terminated because the occurrence of frequent stereotypic movements (e.g., hand flapping) interfered with performance and endangered the participant and equipment.

Pretests: The experimental procedures are listed in Figure 2. All pretests used a two-choice matching task. On each trial, a touch to the sample produced two comparison stimuli. The positions of the comparisons changed unsystematically from trial to trial. For the pretest involving matching numerals 1, 2, and 6 to their printed names, each of the three words served as the sample on six trials in each 18-trial block. The same word was never the sample on more than two consecutive trials. Because two stimuli were available as incorrect comparisons in the presence of each sample (e.g., 2 and 6 for the sample ONE) each of these stimuli served as the incorrect comparison on three trials.

In addition, two-choice matching-to-sample assessed each participant's pre- and post-training performances on matching printed words to numerals, numerals to their dictated names, and printed words to the same dictated names. Each of these tests contained 18 trials and was constructed as described above with six trials for each numeral (1, 2, 6 initially; 4, 5, 9 in the replication) and the corresponding dictated and printed words as samples and comparisons.

Additionally, participants were asked to name each numeral and printed word orally (9 trials, 3 of each numeral and word). All the testing was conducted in extinction conditions. Comparison selections, regardless of their accuracy, produced the ITI. The experimenter recorded naming responses as correct or incorrect via the keyboard. An independent observer also evaluated the naming performances on one third of the test sessions to provide for reliability assessments. The reliability assessments all showed 100% agreement.

“Popout” training program: The first step, a block of up to 18 trials, was designed to facilitate detection of the correct comparison stimulus just as illustrated in Figure 1 for the

earlier research on identity matching (Mackay et al., 2002). After the participant touched the sample for a particular trial an array of nine comparisons appeared, one that matched the sample and eight others that differed from the sample but were identical to one another. Touching the correct comparison produced the auditory–visual starburst display. Any other selection was an error and produced the blank screen and ITI.

Within all training trial blocks, the sample and the particular stimulus serving as the incorrect comparison changed unsystematically across trials with the restriction that the same stimulus was not the sample on more than two successive trials. For trials with a particular printed word as the sample (e.g., ONE), the incorrect comparisons were 2 on about half the trials scheduled and 6 on the remaining trials. (Note that the number of trials that occurred with each incorrect stimulus in a given block could deviate slightly from one half because the block of trials ended when criterion performance was met. Any trial block thus could even contain an odd number of trials.)

In the first block of trials, nine-stimulus displays were presented until the criterion of three consecutive correct trials was met. The number of incorrect comparison stimuli then was reduced by one, to seven, for the next block of trials. After criterion was met at this stage, one additional comparison was eliminated for trials in the next step. This systematic and gradual reduction in the number of comparisons continued until only the correct and one incorrect comparison remained (i.e., a two-choice matching-to-sample task). If no errors occurred, the training program took 24 trials to complete. If the accuracy criterion was not met within 18 trials or errors occurred on 3 consecutive trials at a particular stage of the training sequence the program backed up to the preceding stage, thus adding one incorrect comparison to the displays. Failure to meet criterion after two additional repetitions of a particular stage ended the session. The next session then began at the nine-stimulus stage, typically on the next day. After criterion performance occurred at the final stage of the training program, the session continued with a block of two-comparison trials in which no back-up could occur following an error. This step lasted for 18 trials or until 10 consecutive correct trials occurred.

Posttests: Posttests were conducted for all pretested relations. Posttests employed two-choice matching only. Posttests ended after 10 consecutive correct trials or a total of 18 trials.

Results

Trained performance: Matching numerals to printed names—In pretests, the four participants (Sara, Mark, Ann, and Chris) did not match the numerals 1, 2 and 6 to their printed names. Their accuracy scores were 50%, 50%, 56% and 45%, respectively. This difficulty was not due to failure to discriminate among the numerals; all four individuals already had learned to match the numerals to their dictated names.

The training established arbitrary relations between the printed numerals (1, 2, and 6) and printed names for three of the participants. The training was not successful for Chris. Table 1 shows detailed data for Ann, the first participant. Ann progressed relatively smoothly through the training sequence in Session 1. However, the posttest lasted just four trials

because Ann demonstrated stereotyped behavior (hand flapping). In Session 2, the session again ended prematurely due to the occurrence of the stereotyped behavior. However, this was a lengthy session comprising 72 trials. Ann's performance was nearly perfect in Session 3 and remained so in a fourth session, thus confirming the reliability of her earlier performance.

The other two participants who met criterion (detailed data not shown) did so in fewer trials than Ann. Mark took two sessions and a total of 78 trials and Sara took a single session of only 44 trials. Sara's performance during training was perfect until presentation of the two-choice trials. She made errors on all three of these trials. Her earlier accurate performance then was reinstated after the back-up procedure reintroduced the three-comparison displays. The fourth participant, Chris, received only a single training session. He made few errors until the displays contained only the sample and two comparison stimuli. Just as for Sara, errors then occurred. Backup to the three-comparison level then produced perfect performance, resulting in return to the two-comparison level where frequent errors again occurred.

Systematic replication—Sara and Mark received training also for the numerals 4, 5, and 9 and their printed names. Each completed the training in a single session (48 trials and 49 trials, respectively). Sara's performance was again almost perfect. Mark's performance improved in his second exposure to the procedure, making fewer errors than with the original set of numerals.

Untrained performances: Stimulus equivalence and naming—Though not central to this experiment, test data showed that the training yielded immediate positive effects beyond those produced directly. Figure 3 shows accuracy scores for Sara, Mark, and Ann for the matching-to-sample and naming tests given before and after training. Data for Chris are not shown because his performance did not satisfy the training criteria for the numerals 1, 2, and 6. All three participants made many errors on pretests of matching printed words to numeral samples and to their dictated names and they did not name the printed words.

The post-training test scores for Sara, Mark, and Ann showed emergence of accurate matching of printed words to numeral samples, a performance demonstrating symmetry of the trained relations (i.e., the functions of the sample and comparison were reversed without disrupting performance). In addition, matching printed words to their dictated names improved thus demonstrating transitivity of the previously learned and newly trained relations between the dictated names, numerals and printed words. Emergence of these performances allows the inference that these stimuli had become equivalent. In addition, oral naming of numerals improved substantially and naming of the printed words, a rudimentary form of reading, emerged. Despite these gains, however, most of these untrained performances remained imperfect.

Figure 4 shows the pre- and post-test performances of Sara and Mark who also were trained with the numerals 4, 5, and 9. Their pre-training patterns of performance resembled their earlier pre-training performances with the numerals 1, 2, and 6. Both showed imperfect performances with the task to be trained, matching numerals to printed words (78 and 33

percent correct, respectively). As with Set 1, 2, 6, earlier, both accurately matched numerals to their dictated names but did poorly in matching printed words to numerals and made errors in matching the printed to the dictated names. They showed inaccuracy in naming numerals and did not name any of the printed words orally. In the post training assessments all but one of the scores (Mark: matching printed words to the dictated names) increased although performances often remained imperfect.

Discussion

These results confirm the effectiveness of the training method for establishing matching-to-sample performances and demonstrate that the procedure can yield arbitrary or symbolic relations among stimuli as well as the identity relations examined by Mackay et al. (2002). The replication data suggest that the training becomes more efficient with increased exposure to the method, thus replicating systematically earlier experiments that allowed assessments of this form of “learning set” (e.g., Buffington, Fields & Adams, 1997; Saunders & Spradlin, 1993). Participants may become more sensitive to the contingencies programmed in the sense that they come to discriminate the consistent relations that occur between stimuli that are presented. Nevertheless, that success was not uniform across participants. More research thus is needed to identify prerequisites that are critical for success and clarify the conditions under which the training may fail. Relevant error analyses are discussed later.

The results of the assessments of untrained performances are consistent with previous studies of stimulus equivalence (e.g., Mackay & Sidman, 1984; Sidman 1971; Mackay et al. 1997), which show emergence of new matching-to-sample performances after prerequisite performances have been established. (Recall too that matching of identical words was accurate.) Here, the participants already matched the numerals to their dictated names and were taught to match numerals to the corresponding printed words. After that limited training, improved performances were found for matching the printed words to the numerals and to their dictated names. These new performances suggest the developing interchangeability (equivalence) of the stimuli. In addition, oral naming of the numerals and printed words improved. Frequently, such imperfect, although improved, performances emerge on tests that follow training immediately and then improve if testing is repeated (e.g., Lazar, 1977; Spradlin, Cotter & Baxley, 1973) or if reinforcement is scheduled to follow only correct selections (e.g., McDonagh, McIlvane, & Stoddard, 1984). Neither of these procedures was used in this experiment but the follow-up research conducted with Chris involved a modified test procedure in which selections of correct comparison stimuli were reinforced.

Experiment 2

This single subject research provided follow-up after the failure to teach Chris the relations between the visual stimuli, numerals and their printed names. Visual color patches and their dictated and printed names were used instead of numerals and their names, thus extending the range of stimuli presented in “popout” training procedures. The color patches were used in order to make less difficult the visual discriminations required by the training tasks, a possibility suggested by results of visual search experiments showing that colors yielded

more effective search than forms or letter-like stimuli (e.g., Carlin et al., 2002; Carlin et al., 1995). The use of dictated names as sample stimuli during training also introduced the potential for facilitation of the stimulus class formation and emergent performances that may derive from training of arbitrary relations among stimuli. Such results would add to those from a number of experiments (e.g., Green, 1990; Sidman et al. 1986; see also Stromer, Mackay, & Remington, 1996) demonstrating that training of baselines that include auditory-visual stimulus control can foster stimulus class formation in comparison with entirely visual baselines like those of Experiment 1.

Two multiple-baseline analyses, each across three sets of auditory and visual color stimuli, were conducted as shown in Figure 5, Phases A and C. The use of entirely visual training was re-examined using one set of the new stimuli (Phase B). The potential stimulus class outcomes were assessed. Additional specific aspects of rationale will be given together with procedural arrangements. Only differences in procedure from Experiment 1 are detailed below.

Method

Stimuli—The stimuli were three sets of colored squares (each 3.25 cm × 3.25cm), and the corresponding dictated and printed (Times, 48 point; whole and part word) color names (For Phase A: Set 1: teal, rust, lime; Set 2: shale, indigo, salmon; Set 3: aqua, plum, olive). The dictated names were presented via the computer. A picture of a “talking teacher” replaced the black square used in Experiment 1 to prompt observing responses to the samples.

Procedure

Baseline and post-training assessments: Chris’s pre-training performances (for all 3 sets of stimuli) for matching of the identical printed color names and matching colors to their dictated names were evaluated using the two-choice task described earlier (36 trials, 4 trials per sample stimulus). Oral naming of all colored patches was evaluated at baseline (36 trials, 4 trials per stimulus). After training concluded for a particular stimulus set, naming tests were given with only the colors used during that training. During each of these evaluations, in a change from Experiment 1, correct responses were reinforced by presentation of praise and the starburst and melody. Incorrect responses resulted in trial termination and the 1s time-out.

Two sets of trials were used to evaluate training outcomes and acquisition criteria were changed from Experiment 1 in the attempt to enhance the maintenance of trained performance. One set of 18 test trials followed training trials without a break in the session. Accuracy of at least 85% was required in order to proceed to another set of trials given a day after training (longer at a weekend). The same level of accuracy was required before training began with new stimuli. Failure to meet criterion resulted in repetition of the preceding training step. Long-term maintenance of trained and untrained performances also was assessed repeatedly 10, 14, 20, and 45 days after the training ended.

Training: Phase A. Matching colors to dictated names: This auditory-visual training resembled the training with visual stimuli (Figure 1) except the samples were dictated by the

computer and the comparisons were colors. Throughout each trial, the computer-generated sample was presented repeatedly at 1 s intervals. The colored patches were presented after the participant touched the talking teacher. One patch corresponded to the dictated word and was designated correct. The other (incorrect) comparisons in the array for a given trial were all the same color. The particular color changed from trial to trial but was from the same stimulus set as the correct stimulus. The number of incorrect comparisons was changed from trial to trial as a function of performance just as in Experiment 1.

The training was carried out in the context of a multiple baseline analysis designed to examine the effectiveness of this auditory-visual adaptation of the “popout” training used in Experiment 1. Training began with the stimuli in Set 1 and was followed by use of Set 2, and then Set 3.

Training: Phase B. Matching colors to printed words: This visual training was designed to allow systematic replication of Experiment 1 but with the color stimuli, thus examining whether the difficulty that Chris had in learning the numeral-word relations extended to relations among other visual stimuli. Such observations would run counter to the common expectation that learning in individuals with autism proceeds best with visual stimuli.

The samples were the printed color names and the comparisons were the patches of color. A multiple baseline was planned using all the same color and printed word stimuli as in Phase A but was not completed with any stimulus Set (see below).

Training: Phase C. Matching printed words to dictated names: For this training, the three sets of color stimuli remained the same but were used in a different order. Here, teal, rust, lime constituted Set 1 as before. However, aqua, plum, olive were used as Set 2, and shale, indigo, salmon, the longest words, were Set 3. The dictated color names were samples and presentation of the comparisons followed a touch to the talking teacher. The correct comparison (i.e., the odd stimulus for each training display) was the printed word corresponding to the dictated sample.

To facilitate the visual oddity discriminations for Set 1 training the non-odd, incorrect printed words were introduced gradually across three sessions as follows. On each trial of the first session, the non-odd comparisons, which changed from trial to trial, consisted of only the first two letters of a word from Set1 that did not correspond to the sample. For the second session, only the final letter was omitted from non-odd comparisons. These same incorrect comparisons then were presented for the initial three steps (8, 7, and 6 incorrect stimuli) of a third session followed by complete words for the remaining steps of training. The criteria for completing each training step were the same as for earlier training. For back-up following an error, the shortened words of the stage preceding the error were presented as non-odd comparisons on the trial following the error.

Post-training assessments: After completion of the Phase C training, tests assessed the matching of colors to printed words, the task that received limited exposure to training trials in Phase B and printed words to colors.

Results

Phase A—Table 2 shows accuracy of performance with the three sets of stimuli presented in tests given before training (columns at left of thick line) and after training was completed for each stimulus set during this multiple baseline experiment. The pretest accuracy levels demonstrate the imperfect, unreliable aspects of Chris's performance in matching the colors to their dictated names. Error analyses confirmed that none of the color-name relations were established reliably. Some improvement occurred on successive, pre-training assessments for Sets 2 and 3 perhaps reflecting the use of differential reinforcement in this experiment and effects of training on prior sets. In two additional tests not shown in Table 2, Chris showed perfect matching of identical printed words but named only one of the nine colors. Following training, performance was virtually perfect and well maintained. Chris also named eight of the nine colors (not shown).

Training—Training for Set 1 took three sessions. In the first two sessions, Chris completed the training trials virtually without error, but performance did not meet the accuracy criterion (85%) on the test trials given in the same session as training (61% and 72% respectively). However, his performance met criterion in the third training session.

For Set 2, Chris performed almost perfectly during the training, and all test trials. However, on a repetition of the posttest, which was inserted because of the passage of four days between sessions, accuracy fell to 17%. A second training session therefore was conducted. Perfect performance occurred during that session and on all later test sessions. After training concluded, he named at least two of the three Set 2 colors: His pronunciation of the word "teal" was difficult to distinguish from "shale".

Training with Set 3 occurred without error and performance remained accurate during later tests. He named all Set 3 colors. Additionally, without further training on matching colors to their dictated names, or naming the colors, assessments also were administered 10, 14, 20, and 45 days after the training ended. Chris performed perfectly on all matching trials, and maintained the naming for eight of nine colors.

Phase B—The entirely visual training given in this part of the Experiment failed to establish the visual matching of the Set 1 colors (teal, rust, lime) to their printed names. As in Experiment 1, Chris completed the training trials but his performance did not meet criterion on the posttests. No training was given for Sets 2 and 3 to avoid building an extensive error history with the stimuli involved.

Phase C—Table 3 shows the data from assessments of performance before and after training of the relations between the printed words and their dictated names. The baseline data confirmed the need for training with all three sets of stimuli and, contrasting with Phase A, showed no effects of repeated testing. After training with each individual set, performance was highly accurate on the posttests and also well maintained.

With respect to oral names of printed words (not shown in Table 3), Chris supplied none before training. Then, as each set was trained, he named the three words involved. These performances were maintained perfectly.

Training—Three training sessions were necessary for Set 1. For the first session (incorrect comparisons shortened by two letters), Chris's performance was errorless through training and the posttest that followed immediately. However, accuracy was only 50% in a nine-trial, two-choice probe that used full-length words as comparisons. Additional training then was given using words shortened by one letter as comparisons and he again performed perfectly. However, probe accuracy was only 66%. After he completed an additional training session that involved the third set of training stimuli (i.e., shortened and complete words as described above) with only four errors, he performed perfectly on all remaining tests, which used complete words as comparisons. After the training concluded, Chris named all Set 1 color words.

For Set 2, Chris's performance met criterion in only one session, an outcome repeated with Set 3. It should be noted that after the relations among the stimuli in Set 1 were acquired, the programmed exaggeration of oddity by shortening S- words was not required for Sets 2 and 3. For these training trials, on which all stimuli were complete words, the oddity in the display sufficed to direct attention to the S+.

Stimulus class assessments—The training given in Phases A and C of this experiment established the matching of colors to their dictated names and also the printed color names to the same dictated names. These two kinds of auditory-visual relations have provided the prerequisite performances that permitted the formation of classes of equivalent stimuli in previous studies (e.g., Sidman, 1971). On that basis it might be expected that the training given to Chris would suffice to yield the entirely visual performances, matching of colors to printed words and vice versa, that were not established via direct training in Phase B. Stimuli from all sets were used in these assessments. Chris's scored 92% for matching colors to printed words, the task used in the limited training with Set 1 (teal, rust, lime) in Phase B and 95% for matching words to colors.

Discussion

This multi-component experiment involved Chris, the student with whom the "popout" training failed to establish visual matching of numerals to their printed names. The outcome of the new training that was provided suggested first that the "popout" procedure, perhaps in combination with the delivery of praise-starburst-melody presentations delivered following correct responses in the baseline condition of this experiment, quickly produced reliable arbitrary relations between auditory stimuli (the dictated color names) and visual stimuli (colors). The positive effects of the popout procedure then were confirmed later using the same dictated names and the printed color names, thus demonstrating the generality of the training procedure. Further, the emergence of visual matching (colors to printed words and vice versa) confirmed that symbolic relations (equivalences) had been established among the stimuli.

Recall, however, that one set of stimuli (teal, rust, lime) was used in the attempt to train color to printed word matching directly. The training did not establish reliable relations between these stimuli. It is possible, nevertheless, that the exposure to these trials facilitated the appearance of the visual matching performances. However, it may be more likely that

the auditory component of the training was important for development of the target performances. The demonstration of these performances is consistent with research (e.g., Green, 1990; Sidman et al. 1986; see also Stromer, Mackay, & Remington, 1996) suggesting that use of dictated names as stimuli during training facilitates the formation of equivalences among stimuli. Particular significance may attach to appearance of the visual matching outcomes (color to word and vice versa) because direct training failed to establish such visual matching performance. Such results have not been reported previously. Notably too, repeated use of the training procedure for additional sets of stimuli produced acquisition of the successive new auditory-visual relations in fewer trials with fewer errors, thus replicating systematically the improved performances observed with training of arbitrary visual relations in Experiment 1.

Finally, much needs to be learned specifically about the training that includes printed words. What is the value of presenting shortened words as S- (non-odd) stimuli? In the present studies, use of the method may not have been necessary. Until Set 3 was introduced, the required discriminations were possible on the basis of initial letters of the words presented.

General Discussion

The emergence of the untrained, visual, matching performances are of considerable interest because of the earlier failure to establish similar visual matching performance by direct training. Such results have not been reported previously. The failure of training confirmed that the difficulty that Chris had in acquiring the numeral-word relations (Experiment 1) occurred also with relations among other visual stimuli. The observations suggest that the training of the arbitrary, auditory-visual relations did not enhance the effectiveness of the entirely visual training procedure. While these results run counter to the common expectation that learning in individuals with autism proceeds best with visual stimuli, the emergence of the visual matching performances is consistent with research suggesting that use of dictated names as stimuli during training facilitates the formation of equivalences between stimuli (e.g., Green, 1990; Sidman et al. 1986; see also Stromer et al., 1996).

Additional research clearly is needed to examine potential effects of continued training with more stimuli. Replication and extension with additional participants would have both practical and theoretical implications, particularly about the role of spoken names. Also, much needs to be learned specifically about the training that includes printed words. What is the value of presenting shortened words as S- (non-odd) stimuli? In the present studies, use of the method may not have been necessary. Until Set 3 was introduced, the required discriminations were possible on the basis of initial letters of the words presented.

Acknowledgments

The authors thank the participants, their parents and the staff of the Evergreen Center, Milford, MA for their cooperation and acknowledge the support of grant PO1 HD 25995 from the National Institute of Child Health and Human Development.

References

- American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 4th ed. Washington, DC: Author; 1994.
- Buffington DM, Fields L, Adams BJ. Enhancing equivalence class formation by pretraining of other equivalence classes. *The Psychological Record*. 1997; 47:69–96.
- Burgoyne K, Duff FJ, Clarke PJ, Buckley S, Snowling MJ, Hulme C. Efficacy of a reading and language intervention for children with Down syndrome: a randomized controlled trial. *Journal of Child Psychology and Psychiatry*. 2012; 53:1044–1053. [PubMed: 22533801]
- Byrne A, MacDonald J, Buckley SJ. Reading, language and memory skills: A comparative longitudinal study of children with Down syndrome and their mainstream peers. *British Journal of Educational Psychology*. 2002; 72(4):513–529. [PubMed: 12495564]
- Carlin MT, Soraci SA, Dennis NA, Strawbridge C, Chechile NA. Guided visual search in individuals with mental retardation. *American Journal on Mental Retardation*. 2002; 107:237–252. [PubMed: 12069643]
- Carlin MT, Soraci S, Goldman AL, McIlvane WJ. Visual search in unidimensional arrays: A comparison between subjects with and without mental retardation. *Intelligence*. 1995; 21:175–196.
- de Souza DG, de Rose JC, Faleiros TC, Bortoloti R, Hann ES, William J, McIlvane WJ. Teaching generative reading via recombination of minimal textual units: A legacy of verbal behavior to children in Brazil. *International Journal of Psychology and Psychological Therapy*. 2009; 9:19–44. [PubMed: 19960112]
- Dube, WJ.; Hiris, J. Match to Sample Program (Version 11.1.3) [Computer software]. Waltham, MA: E. K. Shriver Center for Mental Retardation; 1997.
- Dube WV, Serna RW. Reevaluation of a programmed method to teach generalized identity matching to sample. *Research in Developmental Disabilities*. 1998; 19:347–379. [PubMed: 9690280]
- Dunn, LM.; Dunn, LM. Peabody Picture Vocabulary Test-Revised. Circle Pines, MN: American Guidance Service; 1981.
- Green G. Differences in development of visual and auditory-visual equivalence relations. *American Journal on Mental Retardation*. 1990; 95:260–270. [PubMed: 2261159]
- Hayes, SC.; Barnes-Holmes, D.; Roche, B. Relational frame theory: A post-Skinnerian account of human language and cognition. New York, NY: Plenum Press; 2001.
- Kaufman, AS.; Kaufman, NL. Kaufman Assessment Battery for Children Second Edition. Circle Pines, MN: American Guidance Service; 2004.
- Lazar R. Extending sequence-class membership with matching to sample. (1977). *Journal of the Experimental Analysis of Behavior*. 1977; 27:381–392. [PubMed: 16812000]
- Leiter, RG. Leiter International Performance Scale: Instruction manual. Chicago: Stoelting; 1979.
- Lovaas OI. Behavioral treatment and normal educational and intellectual functioning in young autistic children. *Journal of Consulting and Clinical Psychology*. 1987; 55:3–9. [PubMed: 3571656]
- MacDuff, GS.; Krantz, PJ.; McClannahan, LE. Prompts and prompt-fading strategies for people with autism. In: Maurice, C.; Green, G.; Foxx, RM., editors. Making a difference: Behavioral intervention for autism. Austin, TX: Pro-ed; 2001. p. 37-50.
- Mackay, HA. Conditional stimulus control. In: Iversen, IH.; Lattal, KA., editors. *Experimental Analysis of Behavior*. Vol. 1. Elsevier: Amsterdam; 1991. p. 301-350.
- Mackay, HA. Stimulus equivalence: Implications for the development of adaptive behavior. In: Remington, R., editor. *The challenge of severe mental handicap: An applied behaviour analytic approach*. London, England: John Wiley & Sons; 1991. p. 235-259.
- Mackay HA. Categories of relations between stimuli: Foundations for transitive inference and other emergent behavior. *European Journal of Behavior Analysis*. 2012; 13:111–121.
- Mackay, HA.; Fields, L. Syntax, grammatical transformation, and productivity: A synthesis of stimulus sequences, equivalence classes and contextual control. In: Rehfeldt, RA.; Barnes-Holmes, Y., editors. *Derived relational responding: Applications for learners with autism and other developmental disabilities*. Oakland, CA: New Harbinger Publications; 2009.

- Mackay, HA.; Kotlarchyk, BJ.; Stromer, R. Stimulus classes, stimulus sequences, and generative behavior. In: Pinkston, EM.; Baer, DM., editors. *Environment and behavior*. Boulder, CO: Westview Press; 1997. p. 124-137.
- Mackay, HA.; Sidman, M. Teaching new behavior via equivalence relations. In: Sperber, B.; MacCauley, C.; Brooks, PH., editors. *Learning and cognition in the mentally retarded*. Hillsdale, NJ: L. Erlbaum Associates; 1984. p. 493-513.
- Mackay HA, Soraci SA, Carlin MT, Dennis NA, Strawbridge CP. Guiding visual attention during acquisition of matching to sample. *American Journal on Mental Retardation*. 2002; 107:445-454. [PubMed: 12323069]
- Mackay, HA.; Stromer, R.; Serna, RW. Emergent behavior and intellectual functioning: Stimulus classes, generalization, and transfer. In: Soraci, S.; McIlvane, WJ., editors. *Perspectives on fundamental processes in intellectual functioning*. Norwood, NJ: Ablex; 1998. p. 287-310.
- Mackay HA, Wilkinson K, Farrell C, Serna R. Evaluating merger and intersection of equivalence classes with one member in common. *Journal of the Experimental Analysis of Behavior*. 2011; 96:87-105. [PubMed: 21765547]
- Maurice, C.; Green, G.; Luce, S., editors. *Behavioral intervention for young children with autism: A manual for parents and professionals*. Austin, TX: PRO-ED; 1996.
- Maydak M, Stromer R, Mackay HA, Stoddard LT. Stimulus classes in matching to sample and sequence production: The emergence of numeric relations. *Research in Developmental Disabilities*. 1995; 16:179-204. [PubMed: 7652201]
- McDonagh EC, McIlvane WJ, Stoddard LT. Teaching coin equivalences via matching to sample. *Applied Research in Mental Retardation*. 1984; 5:177-197. [PubMed: 6465880]
- Munir F, Cornish KM, Wilding J. A neuropsychological profile of attention deficits in young males with fragile X syndrome. *Neuropsychologia*. 2000; 38:1261-1270. [PubMed: 10865102]
- Pilgrim C, Jackson J, Galizio M. Acquisition of arbitrary conditional discriminations by young normally developing children. *Journal of the Experimental Analysis of Behavior*. 2000; 73(2): 177-193. [PubMed: 10784008]
- Ray, BA.; Sidman, M. Reinforcement schedules and stimulus control. In: Schoenfeld, WN., editor. *The theory of reinforcement schedules*. New York: Appleton-Century-Crofts; 1970. p. 187-214.
- Saunders KJ, Spradlin JE. Conditional discrimination in mentally retarded subjects: Programming acquisition and learning set. *Journal of the Experimental Analysis of Behavior*. 1993; 60:571-586. [PubMed: 8283149]
- Scerif G, Cornish K, Wilding J, Driver J, Karmiloff-Smith A. Visual search in typically developing toddlers and toddlers with fragile X or Williams syndrome. *Developmental Science*. 2004; 7:116-130. [PubMed: 15323123]
- Serna RW, Dube WV, McIlvane WJ. Assessing same/different judgments in individuals with severe intellectual disabilities: A status report. *Research in Developmental Disabilities*. 1997; 18:343-368. [PubMed: 9292929]
- Sidman M. Reading and auditory-visual equivalences. *Journal of Speech and Hearing Research*. 1971; 14:5-13. [PubMed: 5550631]
- Sidman, M. *Equivalence relations and behavior: A research story*. Boston, MA: Authors Cooperatives; 1994.
- Sidman M, Willson-Morris M, Kirk B. Matching-to-sample procedures and the development of equivalence relations: The role of naming. *Analysis and Intervention in Developmental Disabilities*. 1986; 6:1-19.
- Skinner, BF. *Verbal behavior*. New York: Appleton-Century Crofts; 1957.
- Soraci SA Jr, Deckner CW, Haenlein M, Baumeister AA, Murata-Soraci K, Blanton RL. Oddity performance in pre-school children at risk for mental retardation: Transfer and maintenance. *Research in Developmental Disabilities*. 1987; 8:137-152. [PubMed: 3659440]
- Soraci SA Jr, Deckner CW, Bryant JT, Baumeister AA, Mackay HA, Stoddard LT, McIlvane WJ. Generalized oddity performance in preschool children: A bimodal training procedure. *Journal of Experimental Child Psychology*. 1991; 51:280-295. [PubMed: 2033364]

- Spradlin JE, Cotter VW, Baxley N. Establishing a conditional discrimination without direct training: A study of transfer with retarded adolescents. *American Journal of Mental Deficiency*. 1973; 77:556–566. [PubMed: 4122227]
- Stromer R, Mackay HA, Remington B. Naming, the formation of stimulus classes, and applied behavior analysis. *Journal of Applied Behavior Analysis*. 1996; 29:409–431. [PubMed: 8810064]
- Stromer R, Mackay HA, Stoddard LT. Classroom applications of stimulus equivalence technology. *Journal of Behavioral Education*. 1992; 2:225–256.
- Wilkinson KM, Carlin MT, Jagaroo V. Preschoolers' speed of locating a target symbol under different color conditions. *Augmentative and Alternative Communication*. 2006; 22:123–133. [PubMed: 17114170]
- Wilkinson KM, Carlin MT, Thistle J. The role of color cues in facilitating accurate and rapid location of aided symbols by children with and without Down syndrome. *American Journal of Speech Language Pathology*. 2008; 17:179–193. [PubMed: 18448605]
- Wilkinson, KM.; McIlvane, WJ. Contributions of stimulus control perspectives to psycholinguistic theories of vocabulary development and delay. In: Adamson, LB.; Ronski, MA., editors. *Communication and language acquisition: Discoveries from atypical development*. Baltimore: Brookes; 1997. p. 25-48.

Figure 1.

Outline of training program used by Mackay et al. (2002) to establish identity matching to sample.

Figure 2.

Flowchart showing the sequence of procedures used in Experiment 1. Words refers to printed labels, whereas Names refers to dictated labels.

Figure 3.

Scores (presented as pairs) showing test performances before (upper score of pair) and after (lower score) training with Set 1, 2, 6 for Sara (left pair), Mark (middle pair), and Ann (right pair). Arrows point from comparison to sample stimuli of matching tasks and from samples to oral naming responses. The large arrow indicates the training task. Broken arrows show untrained tasks. See text for details.

Figure 4.

Scores (presented as pairs) showing test performances before (upper score of pair) and after (lower score) training with Set 4, 5, 9 for Sara (left pair) and Mark (right pair). Arrows point from comparison to sample stimuli of matching tasks and from samples to oral naming responses. The large arrow indicates the training task. Broken arrows show untrained tasks. See text for details.

Figure 5.

Flowchart showing the sequence of procedures used in Experiment 2. Words refers to printed labels. Names refers to dictated labels. See text for additional details.

Table 1

Accuracy at Each Training Stage (Set 1, 2, 6) for Ann

Number of Comparisons	Session 1	Session 2	Session 3	Session 4
9	3/4	3/5 12/17	7/9 3/3	3/3
8	5/6	1/5 0/3	3/3 3/3	4/6
7	3/3	3/3	3/3	3/3
6	3/3	5/6 3/3	3/3 3/3	3/3
5	3/3	9/17 0/1	3/3 3/3	3/3
4	5/6 4/5		3/3 3/3	3/3
3	3/8 6/8		3/3 3/3	3/3
2	5/6		4/5 3/3	3/3
Post	0/4		17/18 10/10	

Note. Additional columns of data within a session designate occurrence of backup in training program.

Table 2
Percent Correct Trials for Chris in Phase A: Training Relations Between Colors and Dictated Names

	BASELINE	POSTTEST			MAINTENANCE
Set 1	67	92	100	100	100
Set 2	42	67	100	100	100
Set 3	67	75	75	100	100

The line shown in bold marks the place of training.

Table 3

Percent Correct Trials for Chris in Phase A: Training Relations Between Printed Words and Dictated Names

	BASELINE	POSTTEST	MAINTENANCE
Set 1	50	100 100 100	100
Set 2	42	33 100 100	100
Set 3	50	50 58 100	92

The line shown in bold marks the place of training.