2013

On the utility of matrix training for teaching action-object relations to preschoolers : a comparative analysis

Meghan E. Martineau
Western New England University

Running Head: Matrix Training

On the Utility of Matrix Training for Teaching Action-Object Relations to Preschoolers: A Comparative Analysis

Meghan E. Martineau, MS, BCBA

Western New England University &

The New England Center for Children

A dissertation proposal submitted to the Department of Psychology in the School of Arts and Sciences at

Western New England University

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Behavior Analysis

Supervised by:

Gregory P. Hanley, Ph.D., BCBA

Abstract

Matrix training is the orderly arrangement of learning opportunities, structured in a manner to promote the development of untrained relations. Previous research has demonstrated the effectiveness of matrix training for promoting recombinative generalization across a variety of skills, but no study has compared the relative effectiveness of matrix training to other teaching procedures. The purpose of the current study is to compare acquisition and recombinative generalization of novel relations across non-overlapping matrix training (NOMT), traditional discrimination training (TDT), and overlapping matrix training (OMT). The three students who participated all demonstrated similar acquisition performance across teaching procedures. A difference in performance across the teaching procedures was only evident when tests for recombinative generalization were conducted. Two of the students demonstrated stronger performance on recombinative generalization tests in the TDT condition relative to the NOMT condition. The third participant demonstrated stronger performance on recombinative generalization tests in the OMT condition relative to the NOMT. Secondary analyses suggested that NOMT promoted the development of stimulus overselectivity. Critical features of matrix training for promoting recombinative generalization are discussed.

Keywords: autism, discrete-trial teaching, recombinative generalization, stimulus generalization

On the Utility of Matrix Training for Teaching Action-Object Relations to Preschoolers: A Comparative Analysis

Matrix training is the orderly arrangement of learning opportunities, structured in a manner to promote the development of untrained relations. Previous research has demonstrated the effectiveness of matrix training for teaching a variety of skills including preposition usage (Mineo & Goldstein, 1990), sociodramatic play (Dauphin, Kinney, & Stromer, 2004), object and location identification (Nigam, Olmi, & Saunders, 2006), use of syntactic rules (Goldstein, Angelo, & Mousetis, 1987), spelling (Kinney, Vedora, & Stromer, 2003) and American Sign Language (ASL; Light, Watson, Remington, 1990; Remington, Watson, & Light, 1990). When properly designed, matrix training appears to be an efficient tactic for teachers and behavior analytic practitioners because it promotes acquisition of untaught relations following the training of some relations.

Typically matrices are organized 2-dimensionally with a row and column designated for two target components. For instance, experimenters may develop a matrix with prepositions (e.g., on, off, in) designated for each row, and locations (e.g., boat, glass, cup) for each column. The objects and prepositions are then combined by teaching relations diagonally down the matrix (Dauphin et al., 2004; Goldstein & Mousetis, 1989). This design exposes individuals to each stimulus component at least once. For example, if a matrix is composed of the 3 prepositions and locations noted above, the combined preposition-location relations selected for teaching would be *on-boat*, *off-glass*, and *in-cup*. By teaching those 3 relations, it is possible for 6 untaught relations to emerge (e.g., *off-boat, in-boat, in-glass, on-glass, on-cup, and off-cup*).

Selected relations are taught to a mastery criterion typically set at 80% or above in matrix training (Goldstein & Mousetis, 1989; Kinney et al., 2003; Remington et al., 1990). After taught relations are mastered, a generalization test of the untaught relations is conducted. The emergence of any untaught skills in matrix training is referred to as *recombinative generalization* (RG). Goldstein (1983) defines RG as “differential responding to novel combinations of stimulus components that have been included previously in other stimulus contexts” (p. 281). The possibility of RG makes matrix training an appealing teaching strategy because, when RG occurs, it becomes unnecessary to teach students to behave differentially with each novel stimulus combination.

Although matrix training appears efficient due to possible RG, the extent of its efficiency remains unclear. There has been no research comparing the acquisition and RG produced by matrix training to that produced by other teaching strategies. Other commonly implemented teaching procedures, such as multiple exemplar training, are also selected because they promote stimulus generalization of learned relations (Ducharme & Holborn, 1997). Therefore, the efficiency of matrix training relative to other commonly implemented teaching procedures remains unknown.

Previous research conducted on matrix training provides examples of procedures common to matrix-based teaching. For example, Goldstein and Mousetis (1989) used matrix training to teach object-location or object-location-preposition relations to six individuals diagnosed with mental retardation. A pre-test was conducted in order to determine whether students showed accurate use of the prepositions and locations to be taught. Matrices were then developed with a combination of known and unknown prepositions and locations derived from the pre-test. The size of the matrices varied by student; some were 2-dimensional (e.g., 7 x 7) and some were 3-dimensional (e.g., 3 x 5 x 6). Relations selected for teaching, referred to as target relations, were selected diagonally down the matrix. All correct responses were reinforced with edibles, tokens, or social praise, and there were no programmed consequences for incorrect responses. The teaching of target relations and tests for RG occurred in each session of the experiment. For each student, high percentages of relations were acquired through RG (range, 79-95%). High performance on transfer tests conducted with different teachers and in different contexts illustrated additional benefits of matrix training. In sum, Goldstein and Mousetis (1989) demonstrated that an orderly arrangement of teaching opportunities can yield both recombinative and stimulus generalization.

Although the work of Goldstein and Mousetis (1989) is a benchmark demonstration of matrix training, there are several limitations to their procedures. These limitations are also pervasive in the matrix training literature (Ezell, & Goldstein, 1989; Goldstein et al., 1987; Light et al., 1990; Mineo, & Goldstein, 1990; Nigam et al., 2006; Remington et al., 1990). First, Goldstein and Mousetis arranged matrices by combining known and unknown components, which is a common characteristic within matrix research (Goldstein et al., 1987; Kinney et al., 2003). Goldstein et al. (1987) developed their matrices based upon known and unknown words with the justification that doing so has "been required to facilitate RG" (Goldstein et al., 1987, pp. 550-551). Therefore, the extent to which RG is dependent on the blending of known and unknown relations during matrix training is not yet understood. The extent to which RG is dependent on known relations is important because if components of compound relations must be taught before matrix training occurs, then the purported efficiency of matrix training is questionable.

Second, the conditions under which generalization occurred in Goldstein and Mousetis (1989), and in many matrix training studies (Axe, & Sainato, 2010; Dauphin et al., 2004; Ezell, & Goldstein, 1989; Goldstein et al., 1987; Karlan, et. al, 1983; Mineo, & Goldstein, 1990;

Striefel, Wetherby, & Karlan, 1978) also raises some important questions. Stimulus generalization is by definition the spread of the effects of reinforcement (or other operations such as punishment) during one stimulus to other stimuli differing from the original along one or more dimensions (Catania, 2007). It seems important then that reinforcement does not occur during tests of RG; however, reinforcement was scheduled for all correct responses in Goldstein and Mousetis (as well as in Axe, & Sainato, 2010; Dauphin et al., 2004; Ezell, & Goldstein, 1989; Goldstein et al., 1987; Karlan, et. al, 1983; Mineo, & Goldstein, 1990; Striefel et al., 1978). If the novel composite relation comes into contact with a reinforcer, then the behavior is likely under the control of a contingency, and continued instances are not a function of generalization. This is a critical issue in the matrix literature because it complicates the interpretation of previously reported demonstrations of RG (Karlan et al., 1982; Goldstein et al., 1987). In these studies, the first response can be considered an instance of RG, but, once reinforcement occurs, the response can no longer be considered generalized responding. Remington et al. (1990) addressed this methodological issue somewhat by assessing untaught relations under an intermittent schedule of reinforcement. Nevertheless, future research is needed that removes any reinforcement contingency from RG tests so that questions regarding the generalized effects of matrix training can be assessed under more stringent conditions.

Goldstein and Mousetis (1989) used a matrix in which target relations were selected diagonally down the matrix in a non-overlapping fashion (see Figure 1). This procedure has been replicated by other experimenters who have demonstrated the emergence of untaught relations through the implementation of non-overlapping matrix training (e.g., Dauphin et al., 2004; Kinney et al., 2003; Striefel et al., 1978). A procedural difference between these experiments and Goldstein and Mouseties is that no pre-test was conducted in these experiments, and it is, therefore, unclear whether components in these matrices were known to the participants. As discussed above, it is unclear whether the advantageous outcomes regarding RG depend on a blending of known and unknown components. A study evaluating the efficacy of matrix training for promoting RG when all component relations are initially unknown is therefore necessary.

The issue regarding the inclusion of known relations in matrix training becomes even more pertinent when we consider that a number of experimenters who have implemented the same procedure (i.e., a non-overlapping matrix) have not demonstrated the emergence of untaught relations. Furthermore, several of these experimenters have demonstrated that the training of additional relations was then successful at promoting RG (Axe & Santiano 2010; Ezell & Goldstein 1989; Goldstein et al., 1987; Mineo & Goldstein, 1990; Striefel et al., 1978). In other words, the training of an additional overlapping component appears to influence RG performance for some participants.

Instead of relying on known relations while teaching non-overlapping matrices (e.g., Goldstein & Mousetis, 1989), or arranging additional training of some relations in the matrix (e.g., Mineo & Goldstein, 1990), there are a number of experimenters who have selected a procedure with overlapping matrices (Karlan et. al, 1983; Kinney et al., 2003; Light et al., 1990; McCuller & Salzberg 1984; Nigam et al., 2006; Remington et al., 1990; Striefel et al., 1978). An overlapping matrix is a procedure where the selected target relations have overlapping stimulus components. That is, one stimulus component is presented across multiple target relations. For example, the selected target relations for an overlapping matrix composed of the actions (*lift, twist, shake*) and objects (*cat, dog, fish*) would be *lift-cat, lift-dog, twist-dog, twist-fish, and shake-fish*. In this example, there are two objects (*dog* and *fish*) and two actions (*lift* and *twist*) that were presented across more than one target relation. A non-overlapping matrix developed from the same example would include the following target relations; *lift-cat, twist-dog,* and *shake-fish*. Each action is presented with only one object, and each object is only presented with one action. Previous matrix training research has demonstrated both successful implementation of overlapping matrices and non-overlapping matrices. Therefore, it remains unclear as to whether to teach target relations via overlapping or non-overlapping matrices.

Comparing matrix training to other ways of arranging target relations is also necessary because RG to all stimuli does not always occur, and RG may also be evident in more traditional teaching arrangement that are not consistent with matrix training. Karlan et al. (1982) found that only 33% of potential ASL relations emerged for one participant following matrix training. In addition, some of the participants in the Remington et al. (1990) and Striefel et al. (1978) studies did not perform any potential emergent relations. On the other hand, other studies have successfully demonstrated RG to untaught relations without the use of matrix training (Mahon, Lyddy, Barnes-Holmes, 2010; Mueller, Olmi, Saunders, 2000; Saunders, O’Donnell, Vaidya, Williams, 2003). For example, Mueller et al. (2000) used a match-to-sample procedure to teach children to discriminate within-syllable units. Mueller et al. taught students to select a written word upon hearing the spoken word from the experimenter. Words were taught in sets of six, and each set of words consisted of overlapping components (e.g; *sat, mat, sop, sug*). Students were presented with RG tests after correctly performing the four target words. The RG tests consisted of two novel words with components from the trained words (e.g; *mop, mug*). Two participants demonstrated strong RG performance to untaught relations following the teaching of one set of words, and the remaining participant demonstrated similar RG performance after learning two sets of target words. The demonstration of RG with other forms of teaching and the discrepancies in the percentage of emergent relations observed in matrix research underscores the need for clarification of variables affecting RG performance. It is not clear what variables influence RG performance in matrix training, and, more importantly, why RG does not emerge at all in some applications. A comparative analysis of matrix training to different teaching procedures may allow for the critical variables affecting RG performance to be examined. It is possible that procedural differences across teaching procedures may uncover variables influencing RG. It is also quite possible that other teaching procedures may result in similar levels of RG, but, without empirical research, conclusions regarding the superiority of matrix training for producing RG remain speculative.

A final concern is with the timing and conditions under which RG has been assessed and demonstrated in matrix training research. A number of studies on matrix training have conducted tests for RG either within the series of teaching trials, or immediately following teaching trials (Goldstein & Mousetis, 1989; McCuller, & Salzberg, 1984; Mineo & Goldstein, 1990). These types of tests may be considered relatively weak demonstrations of RG because participants had a temporally proximate history of reinforcement relative to the RG assessment. The only study to conduct an RG test that was temporally removed from the teaching conditions was done by Nigam, Schlosser, and Lloyd (2006), and their results demonstrated RG for two participants. The RG tests in Nigam et al. (2006) were conducted a day after participants were exposed to teaching. This was considered a more rigorous test of RG because the RG assessment conditions were further removed from the teaching conditions. The current study will replicate the procedures of Nigam et al. (2006) by conducting tests for RG that are temporally removed from the teaching conditions.

The primary purpose of the current study was to compare acquisition and RG of novel relations across non-overlapping matrix training (NOMT) and traditional discrimination training (TDT). The relative effectiveness of both teaching strategies was examined with two children. A participant diagnosed with autism participated because of the relevance of direct instruction for learning socially-constructed relations to this population (Lovass, 1987). A typically developing child also participated to extend matrix training research to this population. There are a number of additional methodological features that were incorporated in the current study based on the limits of previous research. First, matrices were developed with only unknown components. This was accomplished by conducting a pre-test of all components in the matrices prior to teaching, and only including those to which the children did not respond correctly. Second, reinforcement and corrective feedback were not provided when testing for the emergence of untaught relations. Third, tests for RG were conducted at least a day after participants were exposed to teaching. This was done to conduct a more rigorous test of RG. Fourth, a test for stimulus generalization was conducted with a different experimenter and a different setting. This stimulus generalization test assessed the extent to which relations acquired through RG persisted outside of the teaching conditions. Fifth, the influence of overlapping versus non-overlapping target relations on RG performance was directly evaluated in a third participant, who was a typically developing child.

Experiment 1

Method

Participants and setting. Two students participated in Experiment 1. Ernie was 7-year 9-month old boy, diagnosed with an autism spectrum disorder (ASD), with an age equivalency score of 3-years and 11-months on the Peabody Picture Vocabulary Test (PPVT). Shelly was a typically developing 3-year 8-month old girl who scored an age equivalency of 5-years and 5-months on the PPVT. Ernie attended an early intensive behavioral intervention program for individuals with autism, and spent a majority of his day included in a general education classroom with a 1:1 aid. Shelly was a preschool student in the same general education classroom. Ernie and Shelly met criteria for inclusion in the experiment because they demonstrated compliance with simple one-step directions, mastered gross motor imitation, and did not require a break prior to completing 5 min of academic work. These skills were determined through record review and teacher reports, and confirmed during the pre-tests described below.

The setting for primary teaching sessions was a 7 m by 5 m room located next to the student's everyday classroom. The room contained tables, chairs, a shelf with toys, materials needed to conduct sessions, and a video camera. The student's everyday classroom served as the second setting in a test of RG. Their classroom consisted of a circle area, a break area, and a central table with small chairs. In addition, peers of the same age were present in the classroom. The experimenter and students participating in the study sat a central table away from the others students in the classroom.

Preference assessment. Prior to any teaching, a paired-item preference assessment (Fisher et al. 1992) was conducted twice with each student to identify five highly preferred edibles for each participant. The items selected for use in the preference assessment were determined through a brief interview with each student's primary teacher. The five selected edibles were distinct and complementary to each other, and were isolated to the experimental conditions of this study. This was done in order to avoid satiation with the same reinforcer type during experimental sessions.

Component pre-test. The goal of the pre-test was to determine actions and objects that were unknown to the students. Students were presented with opportunities to perform actions described by the experimenter. In these trials, the students were presented with a single item (e.g., a bear) and given one of several directives (e.g., twist, turn, lift) across trials. Students were also presented with opportunities to touch various objects when they were nominated by experimenter. In these trials, seven target objects and one distracter object were placed on the table in front of the student and the students were given directive to “touch the (object).” A minimum of 30 objects and 25 actions were assessed in baseline. Each action and object was assessed twice. Some students were tested on more than the minimum number of actions and objects in order to identify unknown components to assign to each of the teaching conditions.

There were no programmed consequences for correct or incorrect responding during the pre-test (any response was simply followed by the presentation of the next trial). Simple one-step gross motor imitations were embedded, however, in each session in order for the student to experience some reward during pre-test sessions. For instance, the experimenter said “do this,” modeled some simple action (e.g., tap head), and provided the student with social praise and a reinforcer when the student correctly imitated the action. Across all conditions, a continuous schedule of reinforcement was in place for these gross motor imitations. Upon completion of a gross motor imitation, the student was presented with the three to five highly preferred edibles from the preference assessments and instructed to pick one.

There were a minimum of 84 trials in the pre-test for objects and 70 trials for actions (see Table 1). The 84 object trials (30 object relations assessed twice, and 24 gross motor imitations) were randomized and occurred in 14-trial sessions. The 70 action trials (25 action relations assessed twice, and 20 gross motor imitations) were randomized and occurred in 14-trial sessions. Randomness was achieved through the use of the Microsoft Office Excel 2007 program. If student performance was 100% for any given object or action, it was excluded from the subsequent teaching arrangements.

Component assignment. Twelve actions and twelve objects were randomly selected from all pre-test stimuli in which performance was 0%. If, after the pre-test, less than 12 actions or 12 objects were performed at 0%, then a maximum number of two objects or two actions with performance of 50% were assigned to each teaching condition. The stimuli with performance at 50% were randomly assigned, but in such a way, that one stimulus was assigned to each teaching condition. The matrices for each teaching condition were then arranged by randomly assigning six actions vertically and six objects horizontally in two 3 x 3 grids (see Figures 1, 2, 3 and 4).

NOMT arrangement. Six unknown actions and six unknown objects were randomly assigned to NOMT. The target relations were selected diagonally down the matrix in such a way that one stimulus component was paired with one only other stimulus component (see Figure 1). In other words, the relations selected for teaching were non-overlapping.

TDT arrangement. The actions and objects randomly selected for TDT were separate and distinct from the actions and objects selected for NOMT. In order to make the comparison between the two conditions as equal as possible, the selected target relations included three actions and three objects because it was comparable to the stimuli exposure in NOMT. The target relations were formed by paring one object with two actions, and two objects with one action (see Figure 2), a similar procedure to one utilized by Bunce, Ruder, and Ruder (1985).

Baseline. When testing action-object relations, students were presented with seven objects located on a table in front of them. They were then provided with the instruction "(action) the (object)." There were no programmed consequences for correct or incorrect action-object responses in baseline. Consistent with the pre-test, simple gross motor imitations were embedded in order for the student to experience intermittent rewards. The reward system was the same as the pre-test in that students received an opportunity to choose from three to five highly preferred items upon successfully completing gross motor imitations.

The experimenter assessed all relations in each matrix (i.e., each potential emergent relation and each target relation) twice. The remaining trials consisted of simple gross motor imitations. There were 24 trials in each condition (see Table 1). NOMT sessions consisted of three NOMT target relations assessed twice, six potential emergent relations assessed twice, and six gross motor imitations. TDT sessions consisted of 24 total trials with four TDT target relations assessed twice, five potential emergent relations assessed twice, and six gross motor imitations.

One session occurred in the morning, and one session occurred in the afternoon with a minimum of 120 min separating sessions. At the end of a session (i.e., 24 trials), the student was brought back to his or her classroom and continued with daily programming. The same experimenter conducted all baseline and teaching sessions.

Teaching. When teaching action-object relations, the experimenter placed seven objects in an array directly in front of the student. Six of the objects were from the corresponding NOMT or TDT grid, and one object was a distracter stimulus that was never correct. For each trial, the student was instructed "(action) the (object)." For all sessions, the order of the instructions was randomized, and the placements of the objects randomly varied as well. This was done in order to avoid responding influenced by object positioning or instruction order. Students were provided with brief descriptive praise (i.e., "good job, lifting the bear") and the opportunity to choose from three to five preferred edibles was contingent on correct responses.

Across all teaching trials, if a student responded incorrectly, the student was represented with the same trial, and a model prompt was provided by the experimenter. The student was then represented with the same discriminative stimulus and given the opportunity to respond independently. If they responded correctly, brief descriptive praise was provided (e.g., “that’s lifting the bear”), but the opportunity to choose from among preferred items was withheld. If students did not respond correctly following a model prompt from the experimenter, then the trial was represented, and the student was manually guided by the experimenter to perform the correct response. The student was then provided with another opportunity to perform the task correctly. If they responded correctly, brief praise was provided by the experimenter. If the student responded incorrectly, no programmed consequences occurred and the experimenter presented the next trial.

Each target relation was presented six times in a teaching block. This was done to make the number of reinforcement opportunities for each relation equal across both teaching conditions. The only difference between NOMT and TDT conditions was the selection of relations to teach, and, as a result, the number of total teaching trials in each session was unequal. A teaching block consisted of 18 trials for NOMT (three target relations assessed six times), and 24 trials for TDT (four target relation assessed six times). The blocks of trials were broken into sessions of nine trials for NOMT, and 12 trials for TDT (see Table 1).

All other procedures including presentation of materials, instructions to participants, and consequences for correct and incorrect responses remained the same. The order of teaching conditions was randomized with a minimum of 120 min separating each condition.

Recombinative generalization test 1 (same experimenter). The setting and experimenter remained the same as in the teaching conditions. RG tests were implemented only after students demonstrated correct performance of at least 72% (13/18 opportunities) in NOMT or 79% (19/24 opportunities) in TDT teaching condition. Once RG tests were initiated, they occurred regularly after every four days of teaching (120 trials of each condition). RG tests occurred prior to four days of teaching when students demonstrated correct performance of at least 94% (17/18 opportunities) in NOMT, or 95% (23/24 opportunities) in TDT RG tests.

RG tests were conducted similarly to baseline sessions and occurred at least 120 min after a teaching session. For each trial, the student was presented with seven objects in an array, and instructed to "(action) the (object)." Each potential emergent relation was assessed twice (see Table 1). There were 16 trials in NOMT RG tests (six potential emergent relations assessed twice, and four gross motor imitations). There were 14 trials in TDT RG tests (five potential emergent relations assessed twice, and four gross motor imitations). No reinforcement was provided by the experimenter when a student correctly performed an action-object relation (and no correction was provided following incorrect responses). Reinforcement and social praise were provided by the experimenter only when the student correctly performed a gross motor imitation.

Recombinative generalization test 2 (different experimenter, different setting). RG Test 2 was implemented only after 50% correct was demonstrated in RG 1 for at least one teaching condition. All procedures in RG Test 2 remained the same as RG Test 1 except the experimenter and setting differed. The test occurred in the student's everyday classroom, and was conducted by a trained graduate student in behavior analysis who had not been involved in any teaching sessions.

Dependent measures. The dependent measure was the percentage of correct action-object relations demonstrated by each participant. This was defined as the student independently completing the correct action with the correct object as described by the experimenter within 5 s of the experimenter presenting the instruction. The percentage correct was calculated by taking the total number of correct action-object relations performed divided by the total number of opportunities.

Data were also collected on whether each component of a response was performed correctly. Specifically, were collected data on the extent to which the child correctly performed the action instructed by the experimenter, and the extent to which the child correctly touched the object described by the experimenter.

Interobserver agreement. Interobserver agreement (IOA) was assessed by having an independent second observer record the occurrence and non-occurrence of correct responses. At least 20% of baseline, training, and RG test sessions were scored for IOA for both participants. An agreement was defined as the same measurement of behavior across both observers within that interval. The total percentage of agreement was calculated by dividing the number of agreements by the number of agreements plus disagreements and multiplying by 100.

During baseline, IOA was assessed for 29% of sessions, and the mean agreement across participants was 100% for target responses. For the teaching conditions, IOA was assessed for 20% of sessions, and the mean agreement was 100% for target responses. In the RG tests, IOA was assessed for 28% of sessions, and the mean agreement was 96% (session range, 93-100%) for target responses.

Procedural integrity. The extent to which the experimenter implemented the described procedures was assessed by having the second observer record the primary experimenter's behavior. The second observer scored procedural integrity from videotaped sessions. Data were collected on whether the experimenter instructed the student to perform the prescribed action-object relation, and whether she provided the appropriate consequence to a student's response. A correct instruction occurred when the primary experimenter stated the action-object relation in the sequence prescribed by the data sheet. A correct consequence was scored when, for instance, the primary experimenter provided social praise and a reinforcer contingent on a student's correct response or modeled the response contingent on a student's incorrect response during teaching sessions.

During baseline, procedural integrity was assessed for 29% of sessions, and the mean agreement across participants was 100% for prescribed relations and 100% for consequences. For the teaching conditions, procedural integrity was assessed for 20% of sessions, and the integrity across participants was 100% for prescribed relations and 99% (range 99-100%) for consequences. In the RG tests, procedural integrity was assessed for 28% of sessions, and integrity was 96% (range 93-97%) for prescribed relations and 99% (range 99-100%) for consequences.

An IOA measure was calculated on the procedural integrity data collected. The IOA measure was calculated by a trained graduate student for at least 20% of trials scored for procedural integrity. An agreement was defined as the same measurement of a prescribed relation, and the same measurement of experimenter's consequences following a student response. The total percentage of agreement was calculated by dividing the number of agreements by the number of agreements plus disagreements and multiplying by 100. Agreement was scored for 22% of sessions scored for procedural integrity, and agreement was 100% for prescribed relations and 100% for consequences.

Experimental design. An alternating treatments design (Sindelar, Rosenberg, & Wilson, 1985) was used to compare acquisition and RG of composite relations across the NOMT and

TDT conditions. The order of conditions randomly alternated between NOMT and TDT each day. A multiple baseline design across grids was used to determine the effects of teaching on correct responding from baseline.

Results

Preference assessments. A Pearson correlation was calculated to analyze the extent to which the results of each of the two preference assessments were related. The Pearson correlation for Ernie and Shelly were 0.85 and 0.94 respectively. The obtained strong positive correlations provided evidence that the selections in the two preference assessment were consistent; highly preferred items were then selected from the established hierarchies. The five highly preferred edibles selected for Ernie were White Cheeze Itz ®, Cheetos®, Regular Cheeze Itz®, Goldfish®, and Oreos® (see Figure 5). The highly preferred edibles selected for Shelly were Fudge Cookies, Cheetos®, Starbursts®, Dots®, Cheeze Itz®, and Goldfish® (see Figure 6).

Component pre-tests. The pre-test was conducted to determine 12 unknown objects and 12 unknown actions to be randomly assigned into each of the two teaching conditions. Fifteen actions were never performed correctly by Ernie (see Figure 7), but only 10 objects were never identified correctly in the pre-test. This required that two more objects identified correctly on only half of the trials (i.e., performance of 50% correct) be selected for training. Twelve actions were never performed correctly by Shelly. She correctly identified more objects in the pre-test than Ernie. As a result, a greater number of objects needed to be assessed with Shelly (see Figure 8). Eventually, 12 objects were never identified correctly. These unknown actions and objects were then randomly assigned to each of the teaching conditions.

Acquisition of target relations during NOMT and TDT. It was evident that no relations were emitted correctly by Ernie in baseline (see Figure 9). This performance was anticipated because the grids were developed with unknown actions and objects from the pre-test. It was evident that Ernie started to learn the target relations shortly after teaching began. Within 3 trials of teaching for each target relation in NOMT grid one, Ernie demonstrated at least some correct responding to the target relations. Within 14 trials of teaching for each target relation, Ernie consistently demonstrated independent and correct responding for all taught target relations in NOMT grid one. Throughout the remaining teaching trials, Ernie consistently engaged in correct responses. The overall percentage of success was calculated by dividing the total number of correct responses by the total number of response opportunities. Ernie's overall percentage of success for acquisition of the target relations was 71% for NOMT grid one.

We observed similar acquisition results for Ernie under the conditions of TDT grid one. Within 7 trials of teaching, Ernie demonstrated at least some correct responding with all the target relations, and by 14 trials of teaching, Ernie was consistently demonstrating independent and correct responding for all the target relations in TDT grid one. Ernie then demonstrated correct performance of the target relations throughout training. His overall percentage of success was 80% for TDT grid one.

The results with grid two were similar to the results observed in grid one. Ernie demonstrated acquisition of target relations shortly after teaching began. He was consistently correct with target relations within 14 trials for both TDT and NOMT. Performance for the target relations was consistent throughout the remaining teaching trials. His overall grid two success rate was 75% for NOMT, and 73% for TDT.

Figure 10 shows that we saw very similar, albeit faster, acquisition with Shelly. Shelly correctly performed 0% of relations in the matrix during baseline. Performance with target relations was 0% for all target relations on the first teaching trials, but shortly after teaching began, Shelly demonstrated consistent correct responding. Within 3 trials, Shelly was consistently performing the grid one target relations for both NOMT and TDT. Her performance differed from Ernie's because her acquisition was faster and more complete. That is, she consistently demonstrated more correct target relations throughout the course of teaching and fewer errors occurred after the initial training trials. Her overall percentage of success for grid one acquisition was 92% for NOMT and 93% for TDT.

Acquisition performance for Shelly was replicated in grid two (see Figure 10). Shelly was exposed to only one teaching trial for each target relation before demonstrating acquisition of the target relations across both conditions. Her performance remained strong in that she was consistently correct with target relations throughout the remaining teaching trials. The overall percentage of success for grid two was 95% for NOMT and 95% for TDT.

Recombinative Generalization during NOMT and TDT. The comparison of performance in the RG tests provided information on the relative efficiency of each teaching condition. On the first test of RG grid one, Ernie performed 7 out of 10 RG trials correctly (70%) in the TDT condition, and none of the relations were performed correctly in the NOMT condition (see Figure 9). We observed similar results in the second RG test for grid one. Partial RG (60%) was demonstrated in the TDT condition, and no RG was demonstrated in the NOMT condition. The third and final RG test for grid one was assessed with a different experimenter and in a different setting (RG 2). In this test, we saw almost complete RG in the TDT condition (90%), and a small amount of RG in the NOMT condition (17%).

The same relative effects were observed for grid two with Ernie. Ernie never performed any emergent relations in the NOMT condition, but some emergent relations were always performed correctly in the TDT condition. His overall percentage of success in the NOMT condition was 0% for all three RG tests in grid two. His overall percentage of success in the TDT condition was 10% in the first test, 90% in the second test, and 60% in the final test.

Shelly's performance replicated the overall effects observed with Ernie (see Figure 10). Shelly correctly performed more RG relations in the TDT condition than the NOMT condition. Her overall percentage of success for each RG test in grid one was 60%, 60%, and 100% for TDT, and 25%, 42%, and 58% for NOMT (see Figure 10). The second grid demonstrated the more pronounced differences in RG performance. Shelly's RG performance in grid two was 90%, 90%, and 90% for TDT and 8%, 8%, and 0% for NOMT.

Data collected across the two participants were also depicted as a percentage of correct responding (see Figures 11 and 12). This analysis allowed for a more aggregated depiction of acquisition and RG performances. The rate of acquisition for Ernie and Shelly was comparable across the teaching procedures, whereas the amount of RG was clearly different with much higher amounts observed in the TDT condition. This analysis provides evidence that the difference in the rates of acquisition across the two teaching conditions was not responsible for the observed differences in RG performance.

Discussion

Both participants acquired the target relations across the two teaching procedures at the same rate. A difference between the two teaching procedures was not evident until tests for RG occurred. RG performance was consistently higher in the TDT condition than the NOMT condition; these differences were replicated within subjects as well as across learners with and without autism. We hypothesize that the difference observed across the two teaching conditions is likely due to the inclusion of overlapping relations in the TDT condition. The inclusion of overlapping relations exposed students to the same action twice, and each of these times it was correlated with a different object (each object was also correlated with two actions). For instance, with TDT, students were taught *rub-bowl* and *rub-snow mobile* as well as *squeeze-bed*, and *drop-bed* (see Figure 2). By contrast, in the NOMT procedures, each action was exclusively correlated with a particular object. For instance, with NOMT, students were taught *hand me-boat*, *turn-glass*, and *flip-cup* in NOMT. Students were essentially required to demonstrate discriminated responding to both components of the action-object directive in order to maximize reinforcement during teaching in the TDT condition. By contrast, control by either action or object would have been sufficient to maximize reinforcement in the NOMT condition.

We hypothesize that the critical difference between to the two teaching procedures was presence (TDT) or absence (NOMT) of overlapping components. The purpose of Experiment 2 was to determine whether overlapping relations were required for matrix training to be successful in producing RG. Overlapping relations can be arranged in a matrix in a stair-step manner, and is referred to as overlapping matrix training (OMT; e.g., Ezell, & Goldstein 1989; Karlan, et. al, 1983; Light, Watson & Remington, 1990; McCuller, & Salzberg, 1984; Nigam, Schlosser & Lloyd, 2006; Remington, Watson, & Light, 1990). Experiment 2 was designed to compare a non-overlapping matrix procedure (see Figure 13) to an overlapping procedure (see Figure 14).

Experiment 2

Method

The purpose of Experiment 2 was to compare two commonly implemented matrix training procedures; NOMT and OMT. All procedures remained the same as Experiment 1 except for the selection of target relations in the OMT condition. The preference assessments, pre-tests, RG tests, and the component assignments were arranged similarly to Experiment 1.

Participant and setting. One student participated in Experiment 2. John was a typically developing 4-year, 5-month old who scored an age equivalency of 6-years and 5-months on the PPVT. John was a preschool student in the same general education classroom as Ernie and Shelly, and met the same criteria as Ernie and Shelly for inclusion in the experiment. The settings for teaching and RG tests were the same as Experiment 1.

NOMT arrangement. Consistent with Experiment 1, six actions and six objects were randomly selected from all the pre-test stimuli in which performance was 0%. The matrices for NOMT were then arranged by randomly assigning six actions vertically and six objects horizontally in two 3 x 3 grids (see Figure 13). The target relations were selected diagonally down the matrix in such a way that one stimulus component was only paired with one other stimulus component. That is, the relations selected for teaching were non-overlapping as they were in Experiment 1.

OMT arrangement. The six actions and six objects assigned to OMT were also randomly selected from all pre-test stimuli in which performance was 0%. The selected actions and objects were separate and distinct from the actions and objects taught in NOMT. The target relations were selected diagonally down the matrix in an overlapping fashion (see Figure 14). The critical difference between the NOMT and OMT conditions was that stimulus components in the OMT condition overlapped. More specifically, the OMT condition contained target relations with the same component across multiple target relations. For example, the five target relations selected for John in grid one were *turn-tool, turn-soup, rub-soup, rub-jalapeno,* and *lift-jalapeno* (see Figure 14).

Baseline. The baseline procedures remained the same as Experiment 1.

Teaching. The teaching procedures remained the same as Experiment 1 with the exception of the number of target relations taught in the OMT condition. Five target relations were taught in the OMT condition due to the overlapping nature of matrix. A teaching block consisted of 30 trials for OMT (five target relation assessed six times), and 18 trials for NOMT (three target relations assessed six times). The blocks of trials were broken into sessions of 15 trials for OMT, and nine trials for NOMT (see Table 2).

Recombinative generalization test 1 (same experimenter). The procedures remained the same as Experiment 1 with the addition of the OMT criteria to test for RG. RG tests were implemented only after student's demonstrated independent performance of at 76% (23/30 of opportunities) in the OMT condition. The criteria to test for RG based on NOMT performance remained the same as Experiment 1 (13/18 opportunities). RG tests occurred prior to four days of teaching when the student demonstrated correct performance of at least 94% (17/18 opportunities) in NOMT, or 90% (27/30 opportunities) in OMT RG tests.

Recombinative generalization test 2 (different experimenter, different setting). The RG test 2 procedures remained the same as Experiment 1.

Interobserver agreement. During baseline, IOA was assessed for 25% of sessions, and the mean agreement was 100% for target responses. For the teaching conditions, IOA was assessed for 20% of sessions, and the mean agreement was 100%. In the RG tests, IOA was assessed for 20% of sessions, and the mean agreement was 88% (range 75-100%) for target responses.

Procedural integrity. During baseline, procedural integrity was assessed for 29% of sessions, and the mean agreement was 100% for prescribed relations and 100% for consequences. For the teaching conditions, procedural integrity was assessed for 20% of sessions, and the mean agreement was 100% for prescribed relations and 99% (range 99-100%) for consequences. In the RG tests, procedural integrity was assessed for 20% of sessions, and the mean agreement was 99% (range 98-100%) for prescribed relations and 100% for consequences. The IOA for procedural integrity was calculated for 26% of sessions that were scored for procedural integrity, and agreement was 100% for prescribed relations and 100 % for consequences.

Experimental design. An alternating treatments design was used to compare acquisition and RG of composite relations across the NOMT and OMT conditions. The order of conditions randomly alternated between NOMT and OMT each day. A multiple baseline design across grids was used to determine the effects of teaching on correct responding from baseline.

Results

Preference assessments. A Pearson correlation was calculated to analyze the extent to which the results of each preference assessment were related. The correlation for John was 0.94, demonstrating a strong positive correlation. The five edibles selected for John were Lays Chips ®, Fritos®, Cheetos®, Swedish Fish®, and Oreos® (see Figure 15).

Component pre-tests. Ten actions and 10 objects were never performed correctly by John (see Figure 16). As a result, 10 of these unknown actions and 10 unknown objects were randomly assigned to each of the teaching conditions, and two more objects and two more actions performed at 50% independence were randomly selected and arranged in such a way that one was placed in the OMT grid and one was place in the NOMT grid.

Acquisition of target relations during NOMT and TDT. John did not correctly perform any of the relations in baseline (see Figure 17), but he did learn the target relations shortly after teaching began. Within 3 trials, John was consistently performing all the target relations in NOMT grid one. Acquisition was similar to Shelly's in that minimal errors occurred after acquisition of the target relations was demonstrated. His overall percentage of success for grid one was 85% for NOMT.

We observed similar results with the OMT condition grid one. Acquisition of the target relations was achieved within 12 trials, and correct responding was consistent throughout the remaining teaching trials. John's overall percentage of success was 86% for OMT grid one.

Similar results were observed across both teaching conditions when grid two was implemented. John did not perform any of the target relations in baseline, and shortly after teaching, he demonstrated correct responding to the target relations. His overall percentage of grid two success was 93% for NOMT and 89% for OMT.

Recombinative Generalization during NOMT and TDT. Performance on the first test of RG conducted in grid one was 42% for NOMT and 63% for OMT (see Figure 17). The difference in performance across the two conditions was more pronounced in the second and third RG tests in grid one. John's RG performance was 67% for NOMT and 100% for OMT on the second test. The final test for RG was conducted with a different experimenter and in a different setting, and performance was 67% for NOMT and 100% for OMT. These data demonstrated that RG effects occurred at higher levels in the OMT condition relative to the NOMT condition for John.

These same effects were replicated in grid two with John. John's performance across all three tests in grid two for RG were 88%, 88%, and 100% for OMT and 17%, 67%, and 83% for NOMT. The performance on all RG tests was consistently higher in the OMT condition relative to the NOMT condition. In addition, the emergence of RG was more immediate in the OMT condition.

Data were also depicted as a percentage of correct responding (see Figure 18) over the course of teaching. The rate of acquisition across the two teaching conditions was comparable across the conditions for John, but consistently higher levels of RG were observed in OMT.

Discussion

John acquired all target relations at the same rate across both teaching conditions. A difference in performance across the two procedures was not evident until tests for RG were conducted. RG performance was more immediate in the OMT condition relative to the NOMT condition. That is, the implementation of matrices with overlapping target relations resulted in stronger RG performance with John. These results were replicated across both grids. These findings also replicated the data obtained from Experiment 1 because RG performance was stronger when matrices with overlapping target relations were implemented. Therefore the inclusion of overlapping relations appears to be a critical factor for promoting generalized responding.

Conclusion

The three participants acquired all target relations in all the teaching procedures. The main difference in learning was evident when RG probes were conducted. RG in matrices that did not include overlapping target relations was not observed with Ernie, it rarely occurred with Shelly, and it occurred in a delayed manner with John. By contrast in matrices with overlapping target relations, there was immediate RG with Ernie, strong RG with Shelly, and maintained and immediate RG with John. The difference between the selection of overlapping and non-overlapping target relations was critical to the success of matrix training. As a result, we recommend that if one is going to implement matrix training, it is best to arrange matrices with overlapping target relations.

Overlapping matrices were composed of target relations in which the pairing of stimulus components varied. As a result, participants exposed to overlapping matrices were taught discriminations that required attending to the first and second component of the relation. For example, in an overlapping matrix, the action *turn* was paired with the objects *tool* and *soup (*see Figure 14). The reinforcement for *turn* depended on discriminated responding to *tool* and *soup*. Both the action *turn*, and the objects (*tool* and *soup*) in the target relations must have been performed correctly in order for a participant to contact reinforcement. In other words, selecting overlapping target relations fostered the acquisition of a *conditional discrimination*. Discriminations are considered conditional when reinforcement of responding during a stimulus depends on other stimuli (Saunders & Spradlin, 1989; Sidman, 1986).

In contrast, the action-object pairings in the NOMT condition were invariant. Participants contacted reinforcement in the NOMT condition when the same object was paired with the same action. Thus, non-overlapping target relations promoted the development of simple discriminations where participants were not required to behave discriminatively toward both the first and second component of a target relation. Participants maximized reinforcement by performing the same action with the same object. For example, the action *hand me* was always paired with the object *drainer,* and was never paired with another stimulus component (see Figure 1). In other words, the stimulus components in a non-overlapping relation (*hand me-drainer*) were always a discriminative stimulus (S+). There were no trials in which either stimulus components (*hand me* or *drainer*) served as a non-discriminative stimulus (S-).

The function of stimulus components in matrices with overlapping components, by contrast, varied across teaching trials. For example, in some trials the action *turn* was considered a S+ (e.g., *turn-tool*), and in some trials it was considered an S- (*rub-tool*). The difference between the function of stimulus components in non-overlapping and overlapping matrices is important because it affected the type of responding that was reinforced. That is, the greatest amount of reinforcement occurred in overlapping matrices when participants behaved discriminatively to both the first *and* second component of a relation.

When participants were taught discriminations in non-overlapping matrices, performance was not under the control of both components of a relation. Under these teaching conditions, overselectivity can result. *Overselectivity* occurs when behavior comes under the control of a range of input that is atypically restricted (Dube, 2009). Overselectivity occurred at some point for all participants in the current study. This was evident during tests for RG when responding was controlled by either the first or second component of a relation. For instance, Ernie was taught *lift-mixer* in OMT grid one, but when tested for emergence of the relation *lift-drainer*, he demonstrated overselective responding by lifting the mixer. The response performed was not under the control of the both the action and object of the emergent relation. Instead, Ernie's responding was under the control of the taught relation (*lift-mixer)*.

Reynolds (1961) provided an early demonstration of overselectivity with non-human animals. Reynolds (1961) reinforced the pecking of pigeons in the presence of different colored triangle or circles. Tests occurred in extinction to see whether the pecking of pigeons was under the control of the colors or the shapes. One pigeon demonstrated stimulus control by color, and another pigeon demonstrated stimulus control by shape. The behavior of both pigeons was not under the proper stimulus control of both the color and the shape. Instead, stimulus control had become restrictive to one component. Though the research conducted on stimulus overselectivity has primarily focused on visual stimuli rather than auditory stimuli, this research has taught us that overselectivity is related to mental age on standardized tests (Schover &

Newsom, 1976; Wilhelm & Lovass, 1976), and that it frequently occurs with individuals diagnosed with an ASD (Dickson, Wang, Lombard, & Dube, 2006b). Overselectivity developed with all of the participants independent of diagnosis, and instead as a function of the particular manner in which relations were selected for teaching.

Research on this topic has taught us that overselectivitiy is less likely to develop when responding based on multiple stimuli is differentially reinforced (Dube, 2009). Data from the current analysis provides additional evidence to support this claim. Matrices that were designed to provide reinforcement for discriminated responding to both components of a target relation (i.e., overlapping matrices), promoted more accurate RG performance. That is, the combined overall percentage of correct RG in overlapping matrices was 63%, 82%, and 90% for Ernie, Shelly and John respectively. By contrast, the percentage of correct RG responding in non-overlapping was 3%, 24%, and 57% for Ernie, Shelly, and John. It is, therefore, best to arrange matrices with overlapping target relations because this arrangement requires participants to behave discriminatively to both components of a relation in order to maximize reinforcement.

We designed our study to have the same number of teaching trials per relation in order for each target relation to have the same number of opportunities to contact reinforcement. A limitation of the current analysis then is that the overall number of teaching trials was unequal across the conditions. It could be argued that the difference in RG observed across the teaching conditions was a result of more teaching trials occurring in the overlapping conditions (i.e., TDT and OMT). However, there are a number of reasons why it is unlikely that the number of teaching trials accounted for the difference in RG performance across the two conditions. First, Ernie did not demonstrate any RG in grid two of the NOMT condition despite being exposed to a higher number of teaching trials. For example, with OMT we saw initial RG with Ernie in grid one after 12 teaching trials for each relation. Despite 60 teaching trials with NOMT, RG was never observed with Ernie in TDT grid one. Similar findings were observed across all participants. Within 12 teaching trials, Shelly was demonstrating consistent RG performance for TDT across both grids. After exposure to 24 trials in the NOMT condition, she only demonstrated a percentage 58% success in grid one and 0% in grid two. Similarly, John demonstrated strong RG performance across both grids after exposure to 12 teaching trials in the OMT condition. After 24 teaching trials, John did demonstrate an increase in the percentage of correct RG responding in the NOMT condition, but his performance was never as consistent as it was in the OMT condition. The performance of all three participants indicated that exposure to a high number of teaching trials could not be responsible for producing RG. Therefore, another variable must be responsible for the differences observed in RG performance across the teaching procedures.

A post-hoc analysis of performance during the RG tests provided us with evidence that the difference between the two conditions was a result of the inclusion of overlapping relations and not the unequal number of teaching trials across the two conditions. Figure 19 represents average performance on RG tests for all participants in the study, and summarizes the stimulus components that were controlling responding. This analysis revealed that stimulus overselectivity occurred during NOMT teaching and was less likely during TDT or OMT teaching conditions. In this analysis, performances on RG tests were categorized in two ways. First, behavior was categorized as being under the control of both components of a relation during RG tests (e.g., the action and the object were both performed correctly). Second, behavior was categorized as being under the control only one component of a relation during RG tests. That is, the relation was performed incorrectly because only the action or the object component was performed accurately. For example, control by one component was scored if the student twisted the bear when they were asked to *squeeze-bear*. The aggregated results of this analysis, displayed in Figure 19, show that there were a greater number of responses controlled by both components for all participants when the target relations included overlapping components. A more important point, however, was that there was relatively more responding controlled by one component of a relation in grids with non-overlapping components. In sum, our analyses showed that stimulus overselectivity was more likely to develop when teaching relations with non-overlapping components.

Nonetheless, it would still be important for future research to involve comparative analyses in which the number of teaching trials across each condition is equal. This could be accomplished by conducting the same experiment, but re-arranging the number of times each relation is presented. The modification from the current experiment would be that the five target relations in OMT would be presented three times (15 total trials), and the three target relations in NOMT would be presented five times (15 total trials). An analysis such as this would extend the current results and more directly evaluate the influence of uneven distribution of total teaching trials.

Our analysis was also limited because we only tested for emergent relations within part of the complete matrix. Figure 14 depicts the target relations, and those relations assessed for emergence in the OMT condition. There are additional relations within the complete matrix (i.e., the grey relations in Figure 14) that were never assessed in the current analysis. For example in the OMT condition, we taught 10 relations and assessed emergence to 8 untaught relations. If all potential emergent relations (including the grey relations) were assessed in the OMT condition, the total number of potential emergent relations would increase to 26. This is a significant increase in the number of potential emergent relations. It is possible that if those additional relations were assessed, that at least some emergence would have occurred. Future lines of research should evaluate whether matrix training would promote the emergence of this even greater number of relations.

A point to consider is that we included only unknown components in our matrices. It is possible that if we used matrices composed of known components, the importance of overlapping relations may have been less evident. For example, Goldstein and Mousetis (1989) developed portions of matrices with known components. They found that teaching one relation was sufficient at promoting RG to the other known components. As discussed previously, much of the stated efficiency of matrix training would be lost by including only known components in matrices. Instead, it may be helpful for future researchers to determine whether embedding some known components would maximize the effectiveness of matrix training, independent of the extent to which overlapping relations are selected for teaching.

A goal of this experiment was to highlight the amount of time saved when skills are taught systematically. Matrix training involves the systematic selection of target relations in order to decrease the amount of teaching time required for whole classes of relations to emerge. The development of recombinative repertoires ensures maximum amount of learning with a minimum amount of teaching, and may be one of the more important things applied behavior analysis has to offer to the educational field (Alessi, 1987). The current evidence suggests there is much efficiency to this technology. For example, in the current analysis John was taught 10 relations in the OMT condition, and RG was demonstrated to 8 untaught relations. Put in approximate proportional terms, for each relation that was taught in the OMT condition, another untaught emerged. If all relations in the OMT grid (including the grey relations; see Figure 14) were assessed, then a total of 26 relations may have emerged. In this scenario, for every two relations taught in the OMT condition, 5 more may have emerged.

Matrix training is an appealing procedure because it can presumably promote generalized skill acquisition for any set of conditional discriminations. Perhaps the most important future research is in the application of matrix training to develop concept classes of social significance. Early intensive behavioral intervention (EIBI), which has been repeatedly demonstrated to be a successful treatment for improving outcomes of children diagnosed with an ASD (Cohen, Amerine-Dickens, & Smith, 2006; Helt et al., 2008; McEachin, Smith, & Lovaas, 1993), involves the direct teaching of perhaps hundreds of concept classes. An important skill commonly targeted in EIBI programs is imitation. Imitation is critically important to human learning (Schlinger, 1995), and it would be beneficial to increase the efficiency of teaching imitation to children diagnosed with an ASD via matrix training. Teaching imitation through a matrix could be accomplished by arranging objects in the rows (e.g., *pencil, paper, marker*) and actions in the columns of a matrix (e.g., *tap, shake, pick-up*). In order to teach the target relations, experimenters would state "*do this*" and model a specified action with a specified object. Experimenters would then conduct a test for RG after the target relations are performed to a specified mastery criterion. After students independently demonstrate imitation with one matrix, then more matrices with unknown actions and objects could be implemented. Teaching the target relations and demonstrating RG to untaught relation across multiple matrices may ultimately develop a strong repertoire of imitation.

Another important extension of matrix training is a systematic approach to stimulus generalization by arranging matrices with different individuals, settings, or activities as components in the matrix. For example, a matrix could be developed that targets one skill (e.g., responding to one's name), and promotes the generalization of that skill across settings and individuals. The settings (e.g., classroom, gym, hallway) could be placed in the horizontal rows of the matrix, and the people (e.g., teacher, parent, peer) could be place in the vertical columns of the matrix. A matrix of this kind has yet to be examined, and it is therefore unknown whether this design would be effective at promoting stimulus generalization.

The findings of this study contribute the matrix training literature in a variety of ways. First, the majority of matrix training research has occurred with individuals with developmental disabilities and this study extended matrix training to typically developing individuals. The application of matrix training to typically developing individuals is pragmatic for same reasons that it is applied to individuals with developmental disabilities; one does not need to teach each and every relation. Second, we implemented matrices with stimulus components that were unknown to the participants. This allowed for our analysis to study the critical variables affecting RG outcomes independent of the blending of known and unknown components. The demonstration of RG in the current analysis was, therefore, not a result of a prior history of reinforcement with the stimulus components. Instead, RG occurred as a result of the manner in which the target relations were taught. Third, no reinforcement or corrective feedback were provided contingent on correct or incorrect responding in the RG tests. This allowed for us to analyze RG throughout the course of the experiment, rather than study just the first instance. Fourth, RG tests were temporally removed from teaching conditions allowing for a more rigorous test of RG. As a result, the current demonstrations of RG contributed to the work of Nigam, Schlosser, and Lloyd (2006) who were the only previous experimenters to demonstrate RG under more rigorous conditions. Fifth, RG was demonstrated for all the participants when tested with a different experimenter and in a different setting. These results demonstrated that relations acquired through RG persisted when tested outside of the teaching conditions. Sixth, this study compared the relative effectiveness of matrix training to other teaching procedures. By comparing two teaching procedures, we discovered that a critical component to development of RG in matrix training was teaching overlapping relations.

References

Axe, J. B., & Sainato (2010). Matrix training of preliteracy skills with preschoolers with autism. *Journal of Applied Behavior Analysis, 43*, 635-652.

Bunce, B. H., Ruder, K. F., & Ruder, C. C. (1985). Using the miniature linguistic system in teaching syntax: Two case studies. *Journal of Speech and Hearing Disorders 50*, 247-253.

Catania, C.A. (2007). *Learning; Interim 4th edition*. Hudson, NY: Sloan Publishing.

Dauphin, M., Kinney, E. M, & Stromer, R. (2004). Using video-enhanced activity schedules and matrix training to teach sociodramatic play to a child with autism. *Journal of Positive Behavioral Interventions, 6*, 238-250.

Dickson, C. A., Wang, S. S., Lombard, K. M., & Dube, W. V. (2006). Overselective stimulus control in residential school students with intellectual disabilities. *Research in Developmental Disabilities, 27*, 618-631.

Dube, W. V. (2009). Stimulus overselectivity in discrimination learning. In Phil Reed, Editor (Ed), *Behavioral Theories and Interventions for Autism* (pp.23-46). Hauppauge, NY: Nova Science Publisher, Inc.

Ducharme, D. E., & Holdborn, S. W. (1997). Programming generalization of social skills in preschool children with hearing impairments. *Journal of Applied Behavior Analysis, 30*, 639-651.

Ezell, H. K., & Goldstein, H. (1989). Effects of imitation on language comprehension and transfer to production in children with mental retardation. *Journal of Speech and Hearing Disorders, 54*, 49-56.

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

Goldstein, H. (1983). Recombinative generalization: Relationships between environmental conditions and the linguistic repertoires of language learners. *Analysis and Intervention in Developmental Disabilities, 3,* 279-293.

Goldstein, H., Angelo, D., & Mousetis, L. (1987). Acquisition and extension of syntactic repertoires by severely mentally retarded youth. *Research in Developmental Disabilities, 8*, 549-574.

Goldstein, H., & Mousetis, L. (1989). Generalized language learning by children with severe mental retardation: Effects of peers' expressive modeling. *Journal of Applied Behavior Analysis, 22*, 245-259.

Karlan, G. E., Brenn-White, B., Lentz, A., Hodur, P., Egger, D., & Frankoff, D. (1983). Establishing generalized, productive verb-noun phrase usage in a manual language system with moderately handicapped children. *Journal of Speech and Hearing Disorders, 47,* 31-42.

Kinney, E. M., Vedora, J., & Stromer, R. (2003). Computer-presented video models to teach generative spelling to a child with an autism spectrum disorder. *Journal of Positive Behavioral Interventions, 5*, 22-29.

Light, P., Watson, J., & Remington B. (1990). Beyond the single sign: the significance of sign order in a matrix-based approach to teaching productive sign combinations. *Mental Handicap Research, 3,* 161-178.

Lovass, O. I. (1987). Behavioral treatment and normal educational and intellectual functioning in young autistic children. *Journal of Consulting Clinical Psychology, 55*, 3-9.

Mahon, C., Lyddy, F., & Barnes-Holmes, D. (2010). Recombinative generalization of subword units using matching to sample. *Journal of Applied Behavior Analysis, 43,* 303-307.

*McCuller, W. R., & Salzberg, C. L. (1984). Generalized action-object verbal instruction-following by profoundly mentally retarded adults. *American Journal of Mental Deficiency, 88,* 442-445.

Mueller, M. M., Olmi, D. J., & Saunders, K. J. (2000). Recombinative generalization of within-syllable units in prereading children. *Journal of Applied Behavior Analysis, 33,* 515-531.

Mineo, B. A., & Goldstein, H. (1990). Generalized learning of receptive and expressive action-object responses by language-delayed preschoolers. *Journal of Speech and Hearing Disorders, 55*, 665-678.

Nigam, R., Schlosser, R. W., & Lloyd, L. L. (2006). Concomitant use of the matrix strategy and the mand-model procedure in teaching graphic symbol combinations. *Augmentative and Alternative Communication, 22*, 160-177.

Remington, B., Watson, J., & Light, P. (1990). Beyond the single sign: A matrix-based approach to teaching productive sign combinations. *Mental Handicap Research, 3*, 33-50.

Reynolds, G. S. (1961). Attention in the pigeon. *Journal of Experimental Analysis of Behavior, 4,* 203-208.

Saunders, K. J., O'Donnell, J., Vaidya, M., Williams, D. C. (2003). Recombinative generalization of within-syllable units in nonreading adults with mental retardation. *Journal of Applied Behavior Analysis, 36*, 95-99.

Saunders, K.J. & Spradlin, J.E. (1989). Conditional discrimination in mentally retarded adults: The effect of training the component simple discriminations. *Journal of the Experimental Analysis of Behavior, 52*, 1-12.

Schover, L. R., and Newsom, C. D. (1976). Overselectivity, developmental level, and overtraining in autistic and normal children. *Journal of Abnormal Child Psychology, 4*, 289-298.

Sidman, M. (1986). Functional analysis of emergent verbal classes. In T. Thompson & M. D. Zeiler (Eds.), *Analysis and integration of behavioral units* (pp. 213-245). Hillsdale, NJ: Erlbaum.

Sindelar, P. T., Rosenberg, M. S., & Wilson, R. J. (1985). An adapted alternating treatments design for instructional research. *Education and Treatment of Children, 8*, 67–76.

Striefel, S., Wetherby, B., & Karlan, G. (1978). Developing generalized instruction-following in the severely retarded. In C. Meyers (Ed.), *Quality of life in severely and profoundly mentally retarded people: Research foundations for improvement* (pp.267-326). Washington, DC: American Association on Mental Deficiency.

Table 1

Components of sessions across conditions in Experiment 1

Trial Type	Pre-Test		Baseline		Training		Recombinative Generalization	
	Actions	Objects	NOMT	TDT	NOMT	TDT	NOMT RG	TDT RG
Target relations	25 actions (x2)	30 objects (x2)	3(x2)	4(x2)	3(x6)	4(x6)	-	-
Recombinative generalization	-	-	6(2)	5(x2)	-	-	6(x2)	5(x2)
Gross motor imitations	20	24	6	6	-	-	4	4
Total trials in a block	70	84	24	24	18	24	16	14
Trials in a session	14	14	12	12	9	12	8	7

Note. NOMT = non-overlapping matrix training, TDT = traditional discrimination training, and RG = recombinative generalization. Some relations were assessed multiple times in one condition. For example, 3(x2) denotes that three relations were assessed twice, and 3(x6) denotes that three relations were assessed six times.

Table 2

Components of sessions across experimental conditions in Experiment 2

Trial Type	Pre-Test		Baseline		Training		Recombinative Generalization	
	Actions	Objects	NOMT	OMT	NOMT	OMT	NOMT RG	OMT RG
Target relations	25 actions (x 2)	30 objects (x 2)	3(x2)	5(x2)	3(x6)	5(x6)	-	-
Recombinative generalization	-	-	6(x2)	4(x2)	-	-	6(x2)	4(x2)
Gross motor imitations	20	24	6	6	-	-	4	4
Total trials in a block	70	84	24	24	18	30	16	12
Trials in a session	14	14	12	12	9	15	8	6

Note. NOMT = non-overlapping matrix training, OMT = overlapping matrix training, and RG = recombinative generalization. Some relations were assessed multiple times in one condition. For example, 3(x2) denotes that three relations were assessed twice, and 3(x6) denotes that three relations were assessed six times.

	Hand me	Lift	Rub	Drag	Pinch	Capture
Drainer	T	G	G			
Mixer	G	T	G			
Salamander	G	G	T			
Spatula				T	G	G
Oven mit				G	T	G
Highlighter				G	G	T

Figure 1: T denotes an example of relations taught in the NOMT condition, G represents untaught relations that were tested for RG with Ernie.

	Cover	Crush	Toss	Flick	Twist	Flip
Clip	G	T	T			
Foam	T	G	G			
Pterodactyl	T	G	G			
Mushroom				G	T	T
Cannon				T	G	G
Screwdriver				T	G	G

Figure 2: T denotes an example of relations taught in the TDT condition, G represents untaught relations that were tested for RG with Ernie.

	Drag	Push	Touch	Crush	Lift	Tilt
Foam	T	G	G			
Clamp	G	T	G			
Radish	G	G	T			
Drainer				T	G	G
Salamander				G	T	G
Ziti				G	G	T

Figure 3: T denotes an example of relations taught in the NOMT condition, G represents untaught relations that were tested for RG with Shelly.

	Roll	Twist	Capture	Flip	Cover	Flick
Snow mobile	G	T	T			
Pterodactyl	T	G	G			
Hammerhead	T	G	G			
Jalapeno				G	T	T
Mixer				T	G	G
Caribou				T	G	G

Figure 4: T denotes an example of relations taught in the TDT condition, G represents untaught relations that were tested for RG with Shelly.

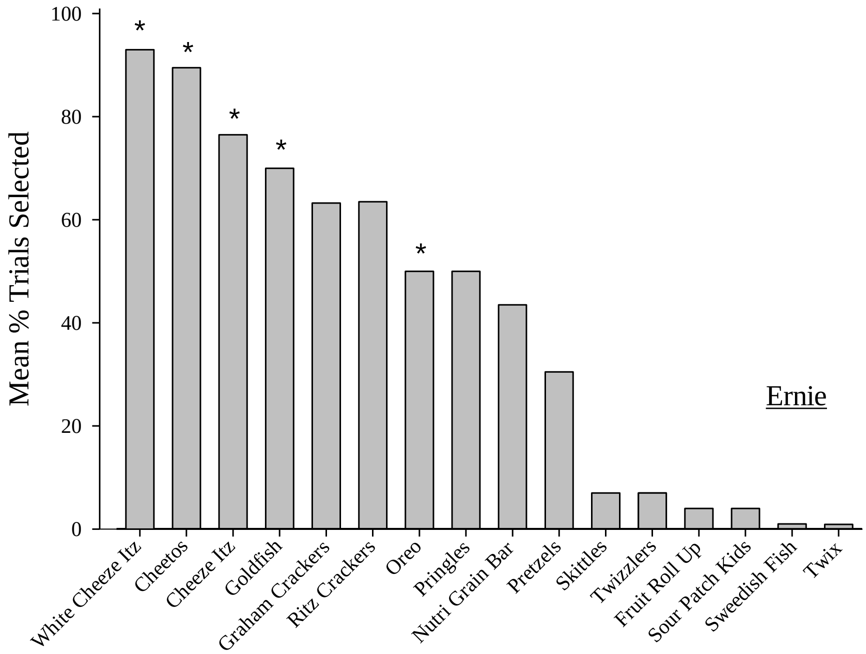

Figure 5: Mean percent of trials edibles were selected across the two Fisher et al. (1992) preference assessments conducted with Ernie. Asterisks denote edibles selected for use in the experiment.

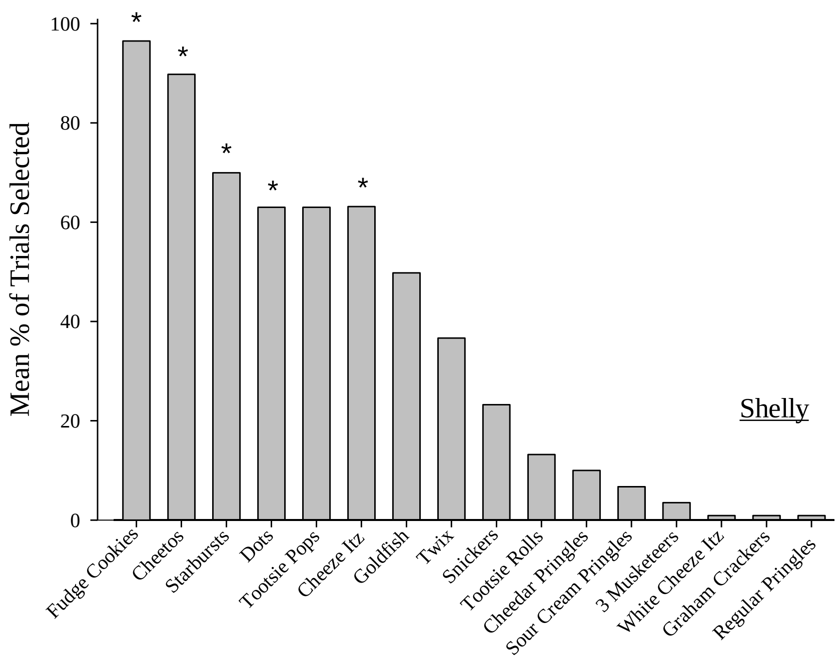

Figure 6: Mean percent of trials edibles were selected across the two Fisher et al. (1992) preference assessments conducted with Shelly. Asterisks denote edibles selected for use in the experiment.

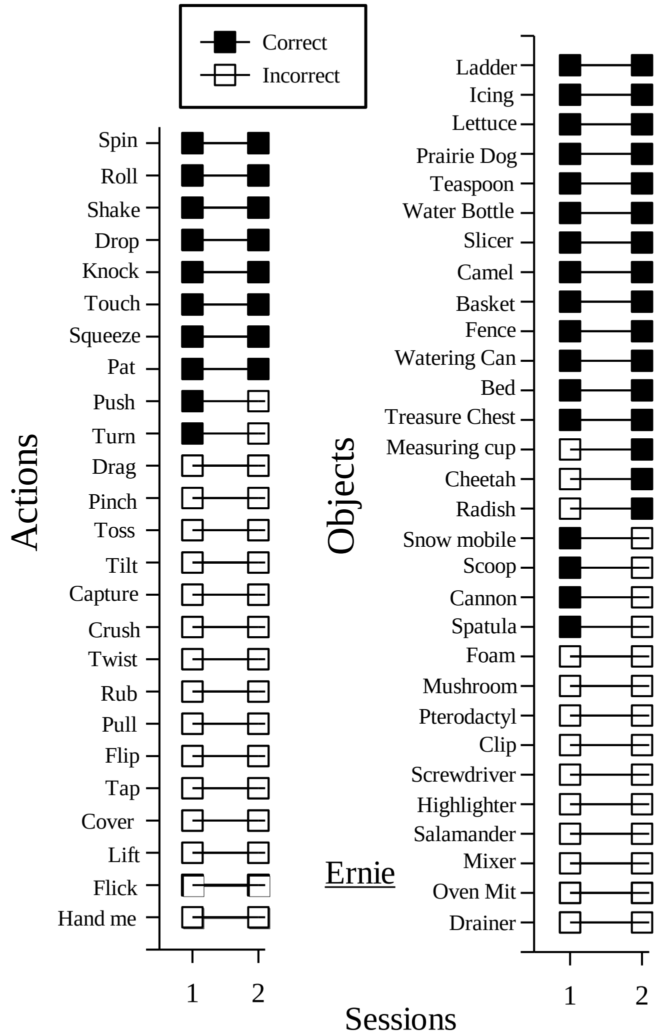

Figure 7: Pre-test data for following instructions with respect to actions and objects for Ernie.

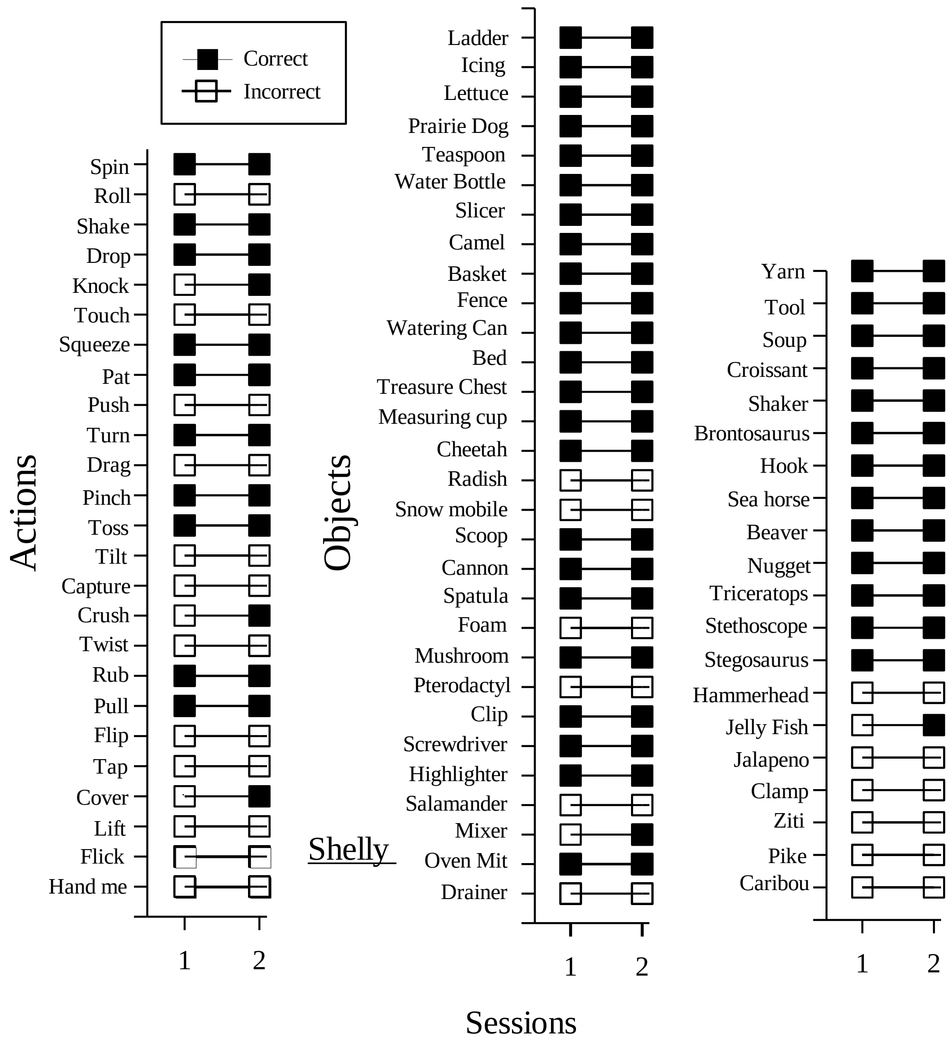

Figure 8: Pre-test data for following instructions with respect to actions and objects for Shelly.

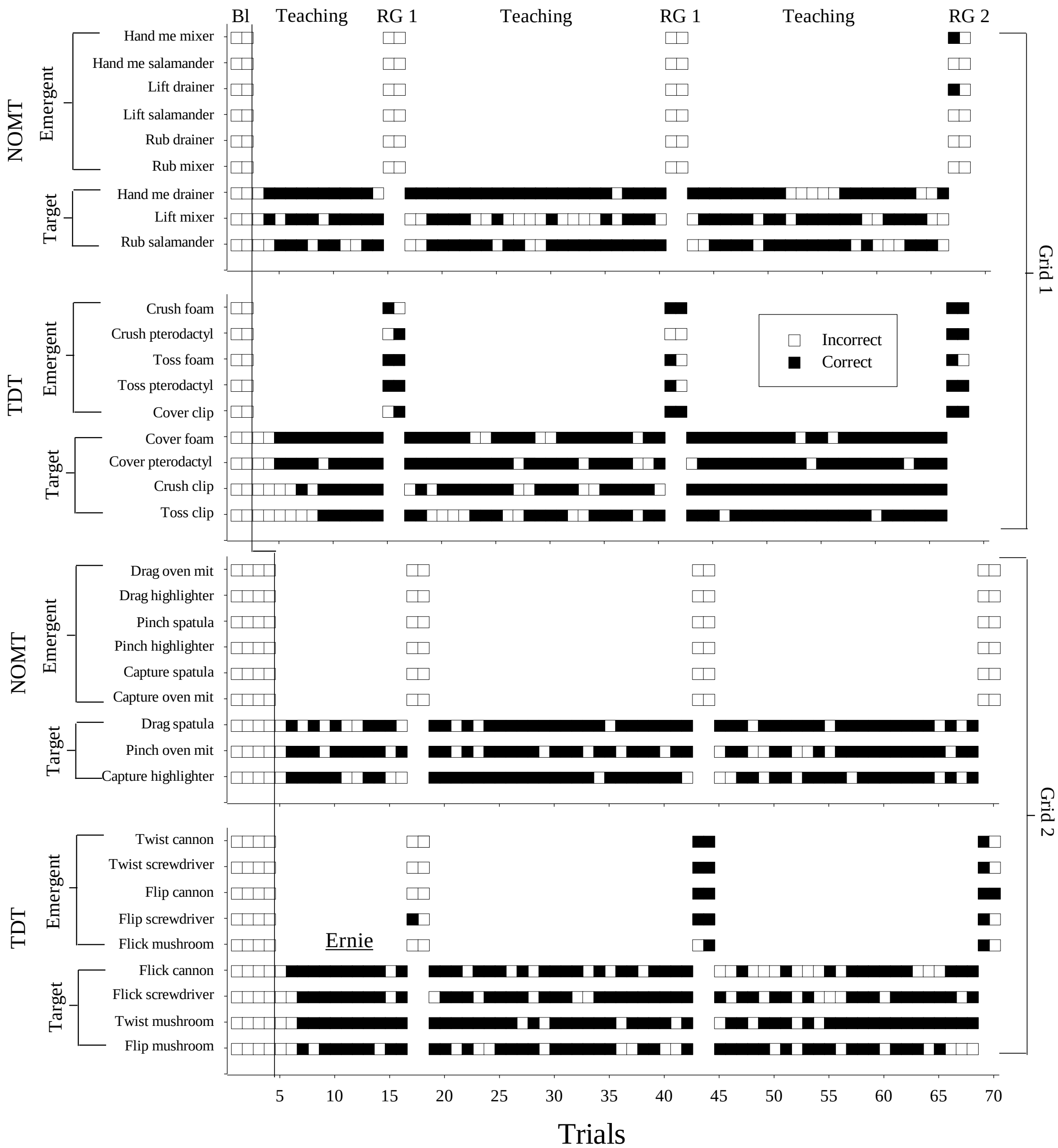

Figure 9: Acquisition and RG performance across the NOMT and TDT teaching conditions for Ernie. Open squares represent incorrect responses and closed squares represent correct responses.

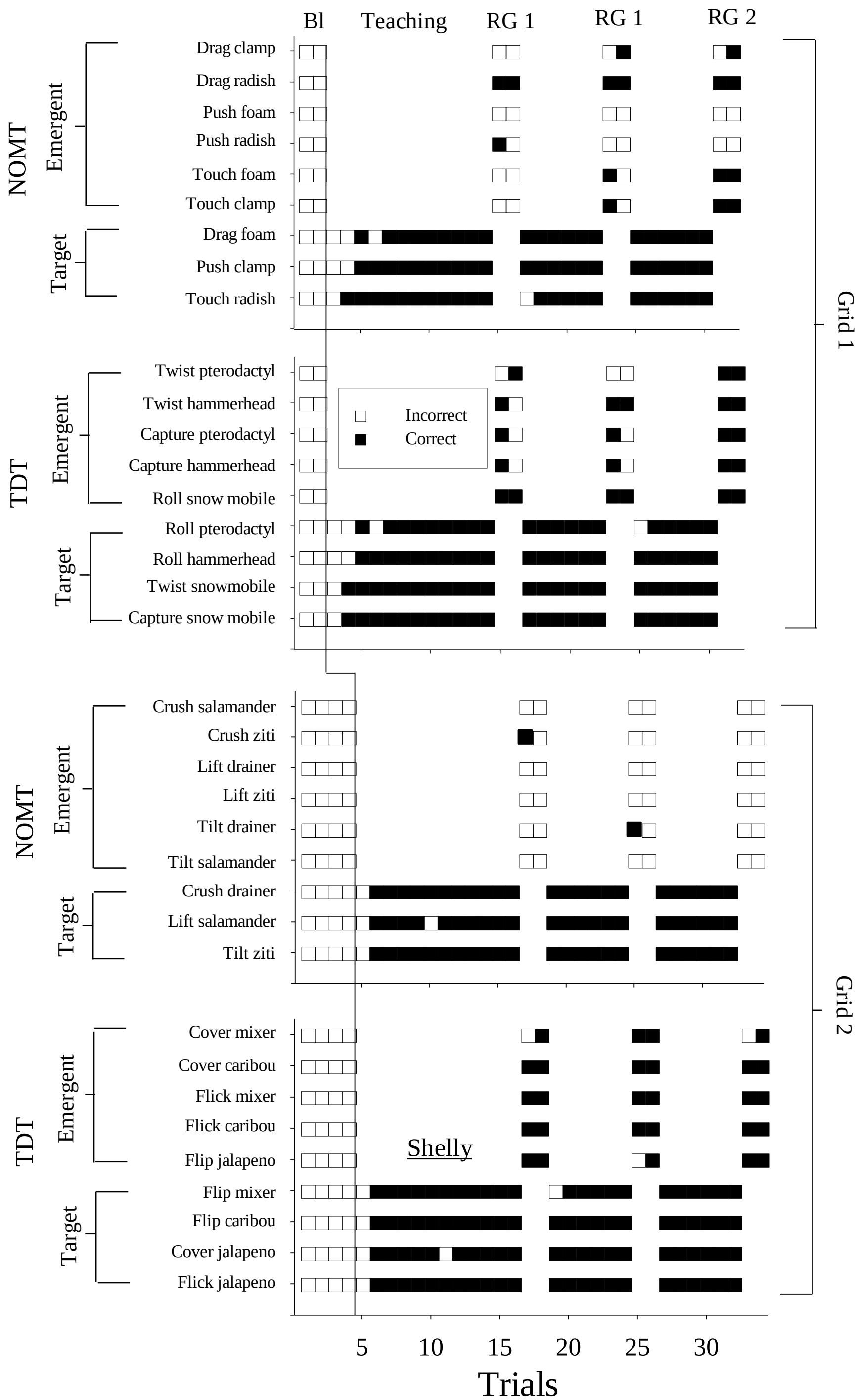

Figure 10: Acquisition and RG performance across the NOMT and TDT teaching conditions for Shelly. Open squares represent incorrect responses and closed squares represent correct responses.

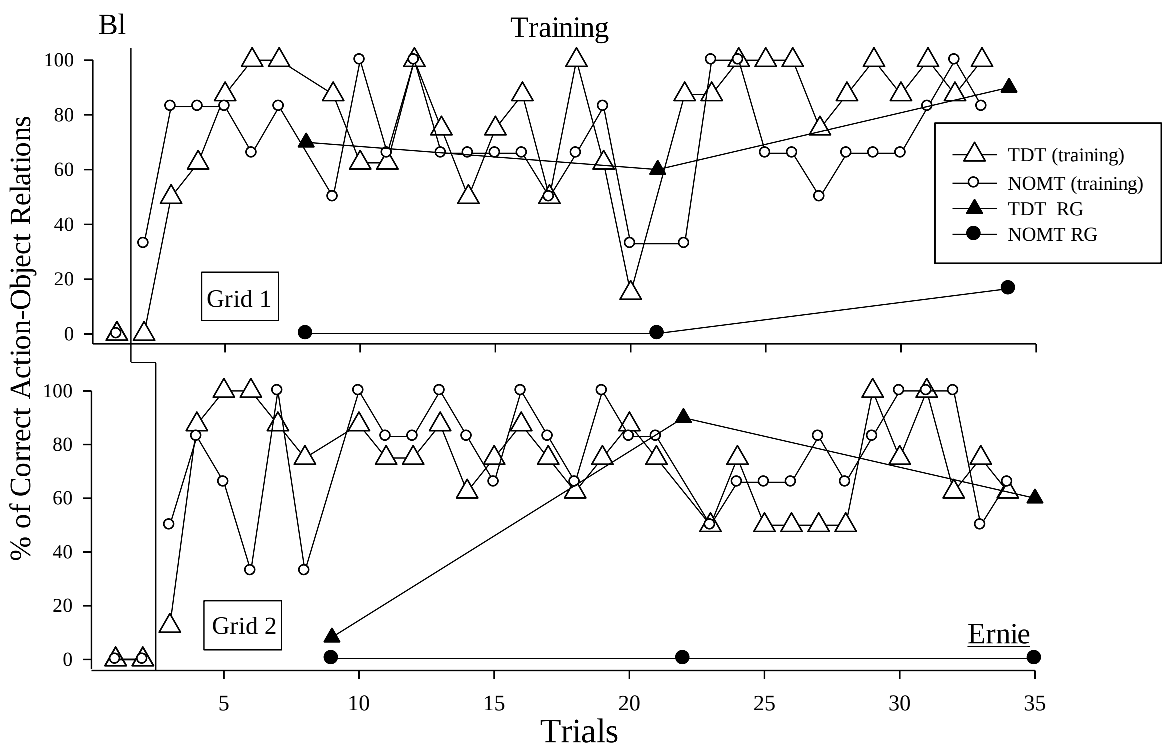

Figure 11: Acquisition and RG performance for Ernie. Each data point represents average performance of all relations across 2 teaching trials.

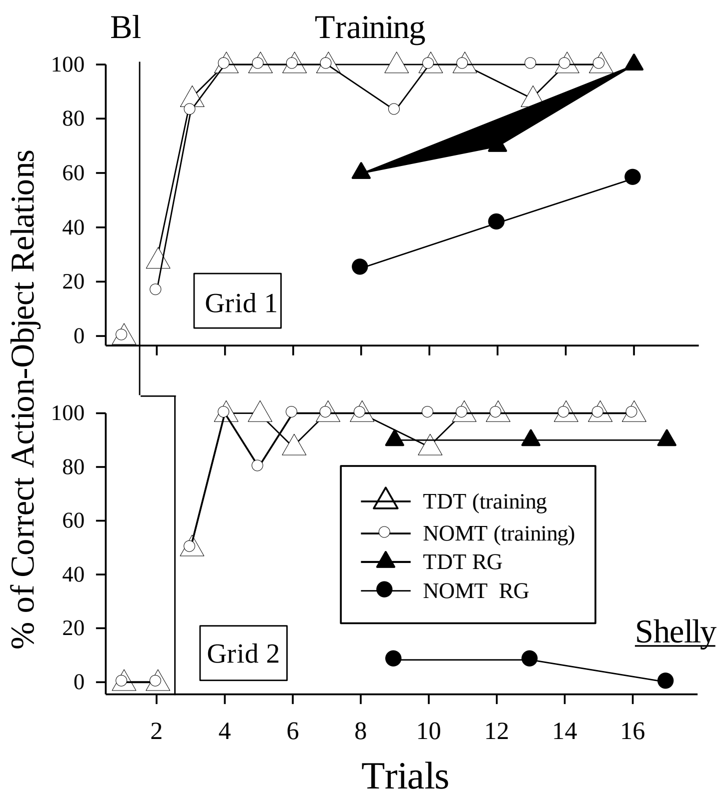

Figure 12: Acquisition and recombinative generalization performance for Shelly. Each data point represents average performance of all relations across 2 teaching trials.

	Tilt	Knock	Flick	Tap	Pinch	Cover
Yarn	T	G	G			
Ziti	G	T	G			
Jellyfish	G	G	T			
Wire				T	G	G
Cork				G	T	G
Sherbet				G	G	T

Figure 13: T denotes an example of relations taught in the NOMT condition, G represents untaught relations that were tested for RG with John.

	Turn	Rub	Lift	Capture	Twist	Pat
Tool	T	G	G			
Soup	T	T	G			
Jalapeno	G	T	T			
Pike				T	G	G
Beatle				T	T	G
Tri-pod				G	T	T

Figure 14: T denotes an example of relations taught in the OMT condition, G represents untaught relations that were tested for RG with John.

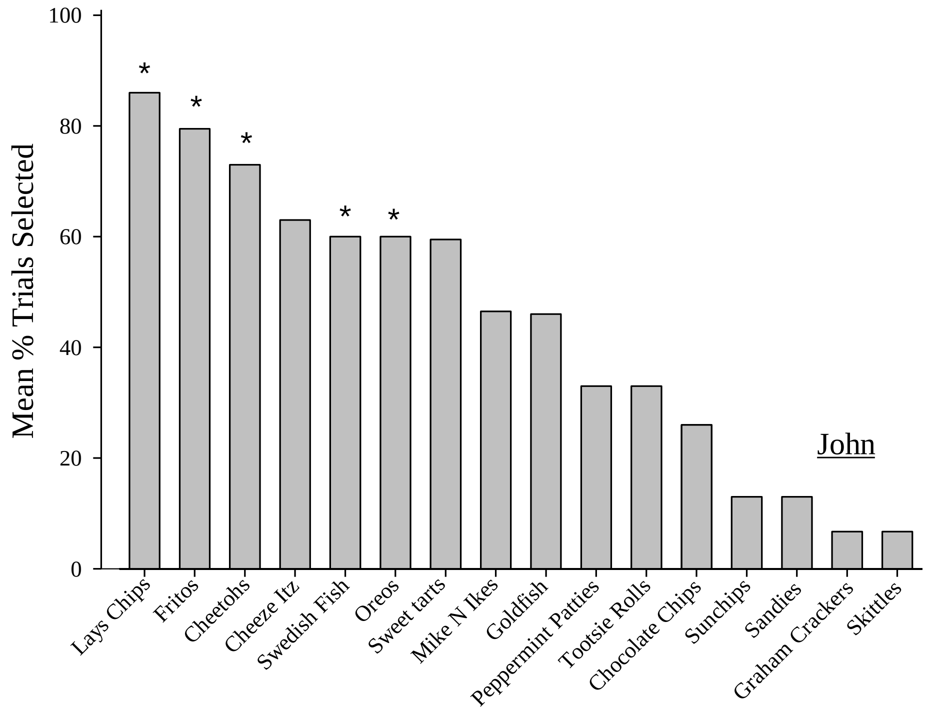

Figure 15: Mean percent of trials edibles were selected across the two Fisher et al. (1992) preference assessments conducted with John. Asterisks denote edibles selected for use in the experiment.

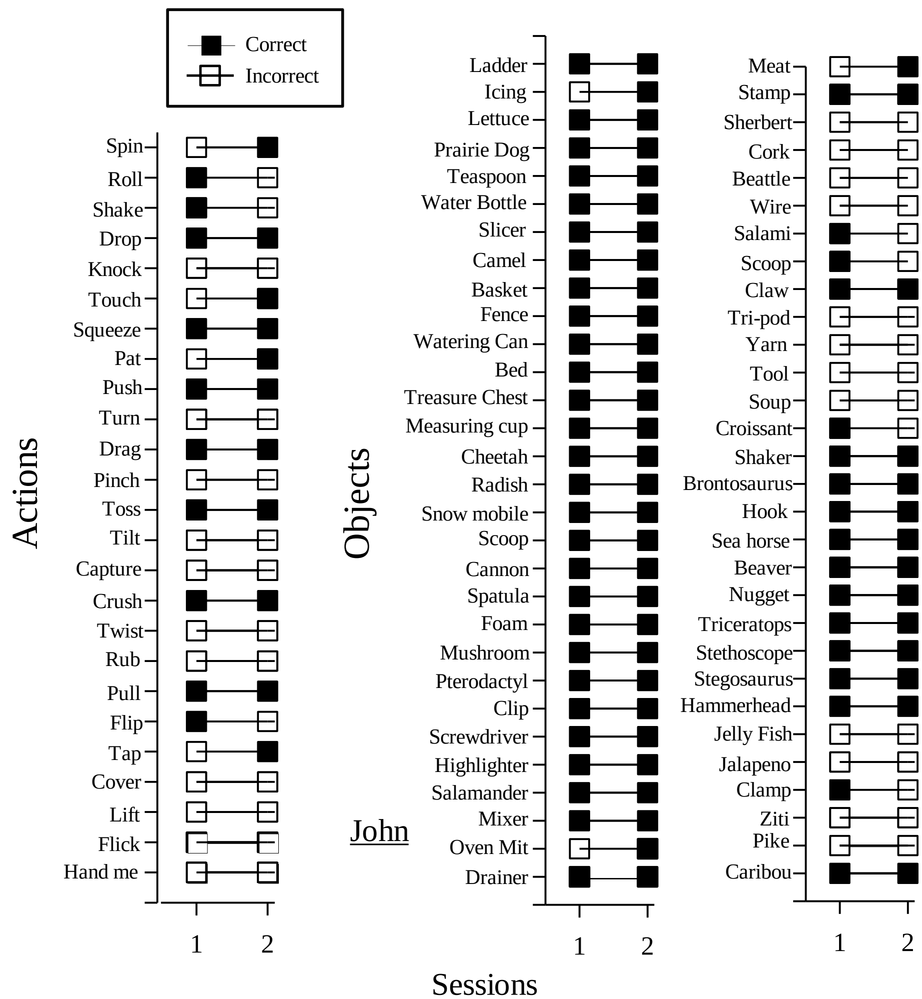

Figure 16: Pre-test data for following instructions with respect to actions and objects for John.

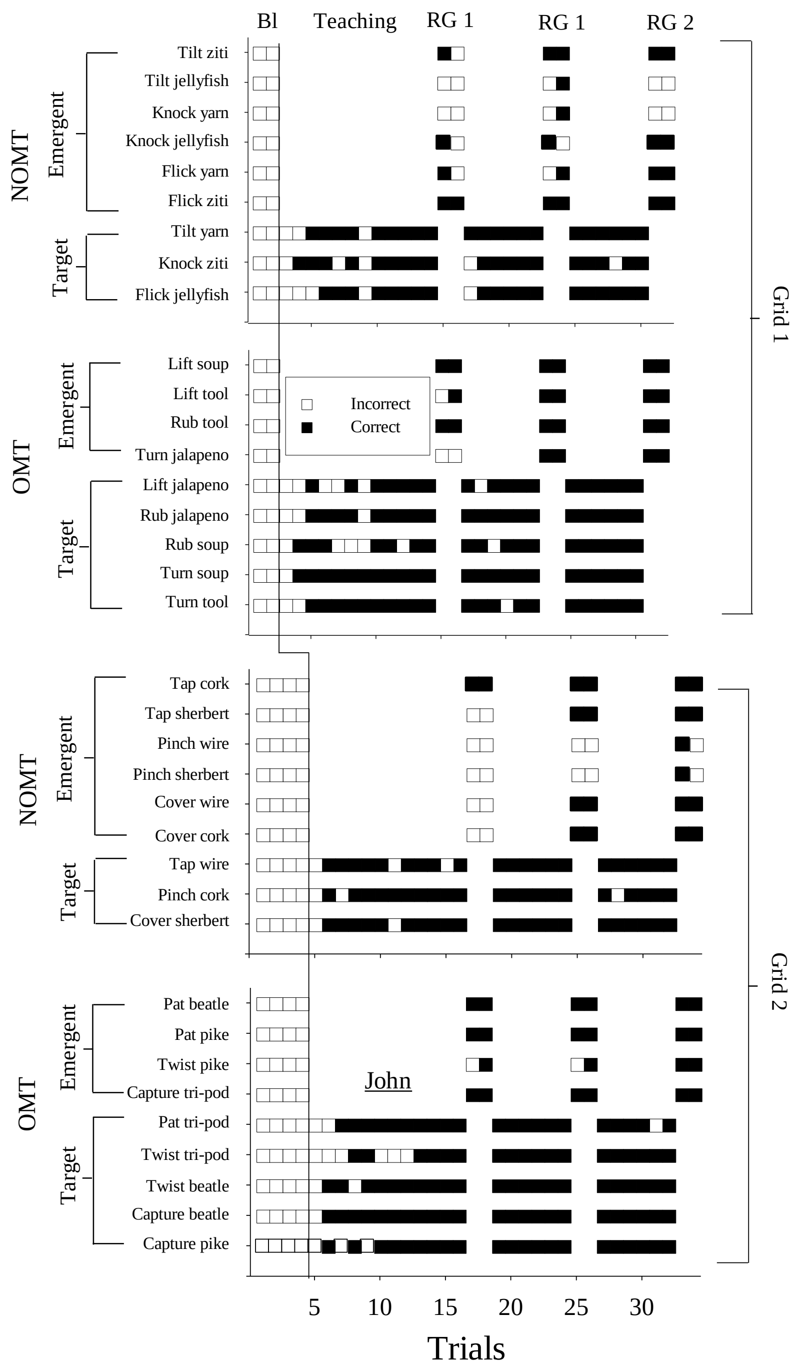

Figure 17: Acquisition and RG performance across the NOMT and OMT teaching conditions for John. Open squares represent incorrect responses and closed squares represent correct responses.

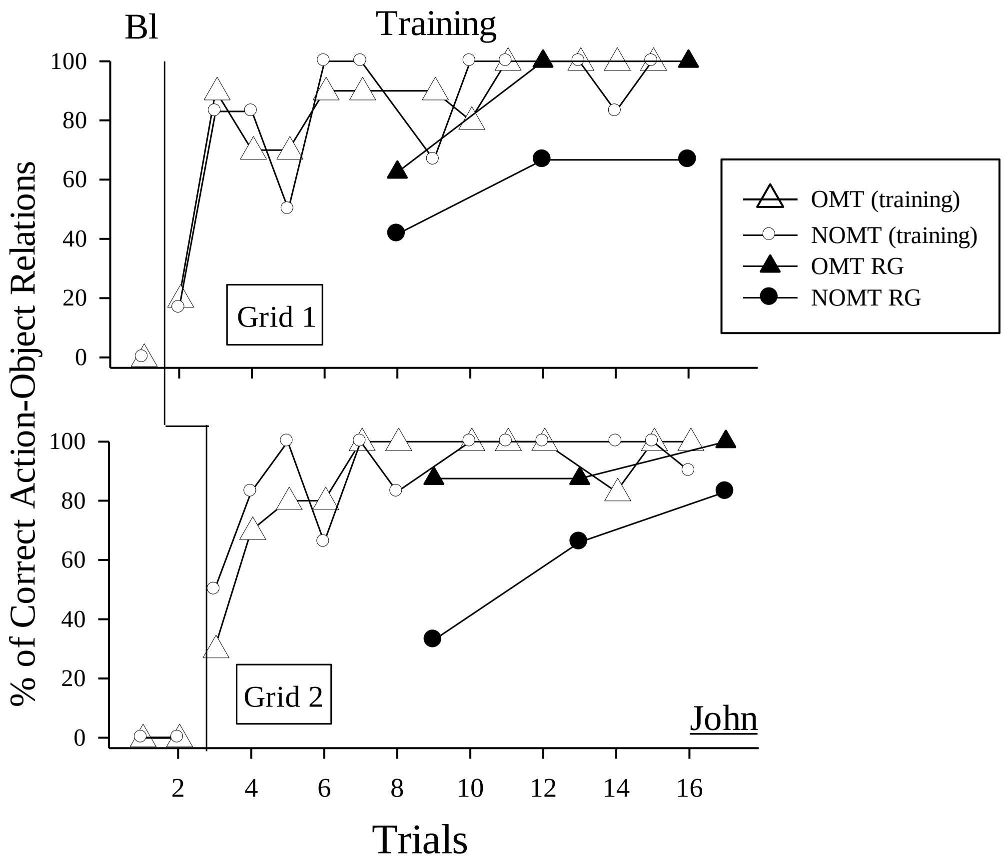

Figure 18: Acquisition and recombinative generalization performance for John. Each data point represents average performance of all relations across 2 teaching trials.

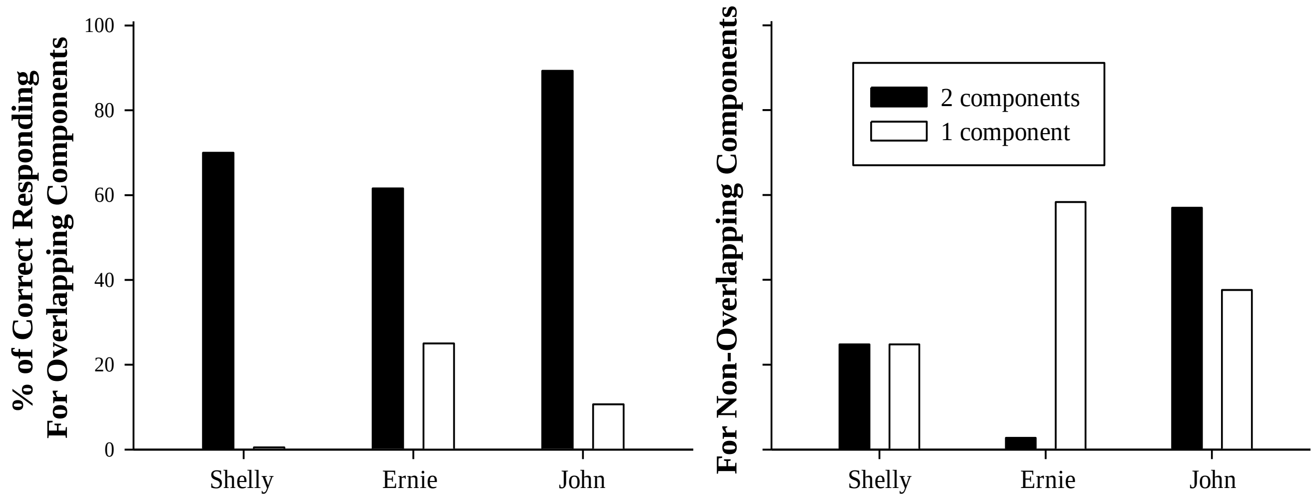

Figure 19: Number of components controlling responding in RG tests for all the participants. Control by one component suggests overselectivity in RG tests.