Comparing a Modified Structured Mix with a Modified Random Rotation Procedure to Teach Auditory-Visual Conditional Discriminations to Children with Autism

Brittany Marie DiSanti[a,*], Svein Eikeseth[b], Sigmund Eldevik[c]

[a,*] **Corresponding Author:** Brittany Marie DiSanti, Department of Behavioral Science, Oslo Metropolitan University, Oslo, Norway.
E-mail: brittmarie13@gmail.com
ORCID: https://orcid.org/0000-0002-8333-3294

[b] Svein Eikeseth, Department of Behavioral Science, Oslo Metropolitan University, Oslo, Norway.
E-mail: svein.eikeseth@oslomet.no
ORCID: https://orcid.org/0000-0001-6404-9071

[c] Sigmund Eldevik, Department of Behavioral Science, Oslo Metropolitan University, Oslo, Norway.
E-mail: sigmund.eldevik@oslomet.no
ORCID: https://orcid.org/0000-0001-7029-1665

Abstract

We evaluated two procedures to teach auditory-visual conditional discriminations (receptive labeling) to children with autism. The procedures evaluated a modified Structured Mix (SM) procedure and a modified Counterbalanced Random Rotation (RR) procedure. The modified SM procedure was based on the logic of simplifying the task by breaking it down into smaller, successive steps and by requiring mastery of each step before introducing the next. Compared to previous studies, the modified SM procedure contained fewer steps, less prompting, and a less stringent mastery criterion. The modified RR procedure targeted all three stimuli simultaneously by presenting them across consecutive trials, both during training and error correction. Sample stimuli were presented in a counterbalanced random order and the comparison stimuli were presented in counterbalanced random positions. Participants were nine children with autism. An adapted alternating treatments design was used. Results showed that the modified SM procedure was more efficient for four of the nine participants, the modified RR procedure was more efficient for one of the nine participants, both procedures were equally efficient for two participants, and neither procedure was effective for two of the nine participants. The modified SM procedure appeared more efficient than the SM procedure employed in previous studies. Despite results, further research is warranted to examine within subject comparisons between original discrimination training procedures and modified procedures.

Introduction

Children with neurodevelopmental disorders such as autism may demonstrate a limited repertoire of receptive language such as auditory-visual conditional discriminations. An auditory-visual conditional discrimination consists of hearing an auditory sample stimulus (e.g., teacher says "touch doll") and selecting the comparison

stimulus corresponding to the correct sample stimulus (e.g., learner touches the doll). This has been called an "if ... then" relation because there is a conditional relation between the sample and comparison stimuli (Sidman & Tailby, 1982). Although an auditory-visual conditional discrimination is a rather basic listener skill, it cannot be assumed to develop naturally in children with developmental delays (Iversen et al., 1986). An auditory-visual conditional discrimination requires that the visual comparison stimuli are observed and that they are discriminated from each other. Secondly, the auditory-verbal sample stimuli, which are presented successively (e.g., the child hears "touch doll" on one trial and "touch car" on the subsequent trial), are observed and discriminated from each other. Finally, and perhaps most challenging, the discrimination of the visual comparison stimulus and its reversal must be brought under instructional control of the verbal sample stimulus.

A discrimination training procedure described in several teaching manuals (Lovaas, 1977, 1981, 2003) has been used to teach auditory-visual conditional discriminations for more than four decades. This procedure was based on the logic of simplifying the task by breaking it down into smaller, successive steps and by requiring mastery of each step before introducing the next. This method has been referred to as the simple-conditional method (Grow et al., 2011) or a structured mix (SM) procedure (DiSanti et al., 2019). It has been questioned whether this procedure is optimal because the reinforcement contingency during the initial steps of training, where each sample stimuli is presented in blocks of trials, does not require discrimination of (or responding to) the verbal sample stimulus (Green, 2001). Rather, aspects of the comparison stimuli may be established as a controlling antecedent stimulus, which may interfere with subsequent control by the sample stimulus. This is because in a particular block of trials selecting the same comparison stimulus across consecutive trials is reinforced. Under this reinforcement contingency, the only stimulus controlling the child's response may be the object itself (i.e., comparison stimulus) and not the verbal sample stimulus. Indeed, the child does not need to discriminate the verbal sample stimulus since correct responding is achieved by selecting the same stimulus that produced reinforcement on previous trials (Sidman & Stoddard, 1966). Whenever selecting this stimulus is no longer reinforced (which would be the case when the reinforcement contingency has been reversed) responding to the other comparison is reinforced. Whenever an incorrect response occurs, the child may simply start responding to the previously trained comparison stimulus until another incorrect response occurs. As a result, an alternative source of stimulus control may be established that competes with the desired type of stimulus control required to establish a conditional discrimination (Green, 2001). In this way, the child may learn "not" to listen or attend to the teacher's instruction and subsequently this

learning history may make it more difficult to get the child to discriminate the sample stimulus, which is necessary to establish a conditional discrimination.

To avoid these potential problems Green (2001) proposed that within a session or a block of trials: (a) a different sample stimulus should be presented on each trial, (b) there should be a minimum of three comparison stimuli on each trial, (c) each sample stimulus should be presented equally often, and (d) the position of the comparison stimuli should vary unsystematically across each trial.

Subsequent studies have suggested that the procedure proposed by Green (2001) is more efficient for most participants as compared to the SM procedure (or simple-conditional) (DiSanti et al., 2019; Grow et al., 2011; Grow et al., 2014; Grow & Van Der Hijde, 2017; Gutierrez et al., 2009; Holmes et al., 2015; Lin & Zhu, 2019; Vedora & Grandelski, 2015). For some learners with a limited verbal repertoire, however, the SM procedure has been found to be more efficient (DiSanti et al., 2019; Lin & Zhu, 2019). Likely, in learners with a limited verbal repertoire, some forms of deficits exist in the way verbal and nonverbal stimuli control listener and speaker behavior (Michael et al., 2011; Sundberg, 2016). Perhaps the SM procedure helps remediate some of these deficits by presenting these elements in a sequential order, introducing the next element only after the previous elements have been mastered, and establishing stimulus control. However, for those children with a more advanced verbal repertoire the SM procedure may interfere with the acquisition of conditional control. This may be due to reinforcing irrelevant sources of stimulus control which compete with the desired type of stimulus control required to eventually establish conditional discriminations.

Following this logic, the SM procedure employed in previous studies may have been less optimal because it contained unnecessary or redundant training steps, had a rigorous mastery criterion, and/or included a rigorous prompting procedure. This may have resulted in a type of overtraining and/or an increased possibility of establishing faulty stimulus control. For the purpose of the present study, we designed a modified version of the SM procedure (i.e., the modified SM procedure). First, training participants to select particular stimuli through mass trials with the inclusion of neutral distractors was eliminated and instead, stimuli that served as the discriminative stimulus (S+) on later steps were included as S-delta (S-). In previous studies, steps of mass trial teaching have involved training with neutral distractors (i.e., S- which later did not serve as S+; Holmes et al., 2015), or training stimuli in isolation (Grow et al., 2011; Grow et al., 2014; Gutierrez et al., 2009; Vedora et al., 2015). In past studies, as part of an errorless teaching procedure, prompts were gradually faded across a large number of trials. In the current study, errorless learning procedures were also included, but prompting was discontinued after one

correct response using a zero-second prompt delay. Modifications in the prompting procedure were made to reduce the potential for prompting an unnecessary number of trials and to minimize the potential for prompt dependency. Finally, the mastery criterion for each step was reduced to avoid overtraining. In previous studies, the mastery criterion for each step involved correct responding across a specified number of trials or across consecutive sessions. In the current study, the mastery criterion was less stringent; stimuli trained through mass trials required four-out-of-four consecutive correct responses. Steps that required discrimination between two or three stimuli required nine-out-of-nine consecutive correct responses.

In addition to the changes made to the modified SM procedure, we also made modifications to the RR procedure (i.e., modified RR procedure). The RR procedure employed by DiSanti et al. (2019) built on a "conditional only" method validated by Grow et al. (2011, 2014). This procedure was designed to teach responding to both the sample and comparison stimuli from the onset of training by presenting a different sample stimulus on each trial (random, but counterbalanced presentation), presenting each sample stimulus equally often and varying the position of the comparison stimuli unsystematically, but evenly across trials. However, typically following an incorrect response, error correction was implemented systematically; thus, not following the random, counterbalanced presentation of stimuli that was essential to the procedure. The same sample stimulus was presented on consecutive prompted trials and the comparison stimuli were placed in fixed positions across these trials, inadvertently diverting the procedure into a SM procedure, which the conditional-only method had been designed to avoid. Hence, the success of the RR procedure may be partly due to the fact that training during trials where error correction was implemented diverted from a RR procedure to more of a SM procedure. In the present study, the modified RR procedure presented trials during error correction, which were consistent with the RR procedure. More specifically, changes made to the RR in the current study were as follows: First, to make the mastery criterion for the modified RR condition comparable to the mastery criterion for the modified SM condition, the mastery criterion for the modified RR procedure was less stringent compared to previous studies. In the current study, mastery was defined as nine-out-of-nine consecutive correct responses, which was identical to the mastery criterion for the final step in the modified SM procedure. Second, in previous studies error correction typically entailed repeating the same trial while a prompt was being faded. That is, the exact position of the comparison stimuli remained the same across consecutive trials, while the particular sample stimulus (which was responded to incorrectly) was re-presented. In the current study, the error correction procedure for the modified RR procedure followed the logic of semi-random presentation of sample and comparison stimuli. That is, a different sample stimulus was presented on each trial, each sample stimulus was presented equally often, and the position of the comparison stimuli varied unsystematically across trials.

For some of the participants with limited and advanced verbal repertoires we trained more than one stimulus set, across both training procedures. This was done to examine whether the pattern of responding on the initial stimulus set could be replicated on additional stimulus sets within participants.

The present study was designed to examine the effectiveness of the modified SM procedure by comparing the number of trials to mastery for three auditory-visual conditional discriminations (i.e., a stimulus set) using the modified SM procedure to the number of trials to mastery using: (a) the modified RR procedure, (b) the original SM procedure (DiSanti et al., 2019), and (c) the original RR procedure (DiSanti et al., 2019). Data from the original SM and RR procedure were abstracted from the DiSanti et al. (2019) study.

Method

Participants, Setting and Training Personnel

Participants were nine, (seven males) three-to-thirteen-year-old children, diagnosed with autism spectrum disorder (ASD). The diagnosis was set by a licensed clinical psychologist. Inclusion criteria were as follows: a repertoire of at least 10 receptive labels, 10 motor imitations, and 10 visual-visual identical matching-to-sample discriminations. The purpose of these inclusion criteria was to ensure the participants were able to scan the visual array, attend to the instructor, and to acquire conditional discriminations via discrete trial teaching (Green, 2001). The number of auditory-visual conditional discriminations was reported by the participant's behavioral therapist before the start of the study.

All participants had previous exposure to both training procedures although the procedure may have appeared less strict in the applied setting compared to the current study (i.e., a less rigorous mastery criterion, flexible error correction procedures, implementation of mastered targets during acquisition training, combination of both training procedures). Table 1 exhibits participant characteristics, and scores from behavioral and developmental assessments.

Training sessions for all participants were conducted by the experimenter (all participants), lead behavior specialist (Participant 5, 6, and 7), or a registered behavior technician (Participant 4, 8, and 9). Sessions were conducted in the participants' school or clinic setting.

Training Sets and Materials

Each training set included three auditory sample stimuli in the form of spoken words. Three corresponding visual stimuli were used in the form of picture templates. A template with framed boxes for stimulus cards to be placed on the right, middle, and left was placed on an A3 laminated paper sheet (29.7 cm x 42.0 cm). A standard-sized clipboard was used for three of the participants due to lack of appropriate responding to stimuli when laminated templates were placed on the tabletop. Individual picture cards were made for the stimuli taught in each training condition. Picture cards were rotated on trials based on the sample stimulus, comparison stimuli available, and position required on the data collection sheet. Data collection sheets containing 30 trials were made for all steps and were randomized for position and stimuli.

For participants 1, 2, 3, 7, and 8 one set of stimuli for each condition was administered. For remaining participants, between two and three stimulus sets were taught to assess the extent to which results could be replicated across participant stimulus sets. Participants were not exposed to the training stimuli outside of the study. Furthermore, instructors who conducted the one-to-one training sessions did not include stimuli or related stimuli within the participant's other therapeutic treatment programs.

Dependent Measures and Data Collection

The dependent variable was the number of trials to mastery for three auditory-visual conditional discriminations (i.e., a stimulus set) in each training condition. One procedure was considered more efficient than the other procedure if the difference in number of trials to mastery exceeded ten percent (Ledford et al., 2019). If the difference in number of trials to mastery were ten percent or less, they were considered equally efficient. A correct response was defined as the participant pointing to the correct visual comparison stimulus within 5 seconds of the presentation of the auditory sample stimulus. A prompted response was defined as the participant selecting the correct comparison stimulus within 5

seconds of the experimenter providing a prompt. An incorrect response was scored if the participant selected the incorrect comparison stimulus or did not respond within 5 seconds after the presentation of the auditory sample stimulus. Data was also collected on the number of prompts required, number of errors that occurred, and maintenance at 4 and 6 weeks.

Interobserver Agreement (IOA)

To measure IOA, a second independent observer recorded participant responses for each training condition. Trial-by-trial agreement was calculated, and an agreement was scored if both the primary and secondary observer recorded (a) a correct response, (b) a prompted response, (c) an incorrect response, and (d) the position of the visual comparison stimulus. Interobserver agreement was calculated by taking the number of trials in agreement divided by the total number of trials in the session, multiplied by 100. Interobserver agreement was collected across all participants for a mean 39.8 % of the trials (range, 33% to 40%). Mean scores for interobserver agreement were 99.9% (range, 98.6% to 100%).

Preference Assessment

To identify putative reinforcers to be used during training, stimuli were chosen based on a teacher report of 20 preferred items. The twenty items were used in a Multiple-Stimulus-Without- Replacement-Preference assessment (MSWO) to identify the top five preferred items for each participant (DeLeon & Iwata, 1996). Before each training session, a brief MSWO using the preferred five items was conducted to identify preference. For two of the participants, the establishment of a token economy system had previously been in place for discrete trial training sessions. Participants who used a token economy system were exposed to a brief MSWO before training sessions in order to identify back-up reinforcers.

Pretest

Six target stimuli were identified through pretests for each stimulus set to ensure targets were unknown

Table 1.

Participant Characteristics

Participant	Age	Receptive Labels	CARS	Developmental Age (Months)	Vineland ABC	Vineland Communication
1	9:7	>200	36.5*	44*	73	76
2	8:11	>200	29.5*	56*	88	86
3	11:8	>200	30*	41*	77	73
4	3:5	>100	29	80	69	76
5	13	<50	46.5	19	20	20
6	11:7	<50	33	34	45	43
7	13:3	<50	29.5	40-42	62	57
8	4:5	<50	33.5	25	69	70
9	4:5	<50	45	21	57	48

Note. Receptive Labels = receptive labels in the participant's repertoire before training; CARS = Childhood Autism Rating Scale Standard Version or High-Functioning Version (denoted with an *; Schopler et al., 2010); Developmental Age = developmental age as derived from Bayley Scales of Infant Development (Bayley, 2005) or Stanford Binet-Intelligence Scale (denoted with an *) (Roid, 2003); ABC = Adaptive Behavior Composite Score and Vineland Communication from Vineland Adaptive Behavior Scales-3 (Sparrow et al., 2016).

to the participants. All target stimuli were nouns, except one stimulus set for Participant 4, which was verbs. Three stimuli were presented on the template in front of the child. Pretest data sheets were created to ensure each stimulus was presented semi-randomly as the sample stimulus three times each, and each stimulus was presented in the right, middle, and left position three times each. The experimenter asked the participant to select one of the comparison stimuli. Stimuli were included in one of the training conditions if the participant responded less than or equal to 33% correct during the nine-trial probe. During the pretest, reinforcement was given approximately every 10 seconds for proper sitting and attending to the experimenter. Reinforcement for proper sitting and attending was not given immediately following a trial. No consequences were provided for incorrect or correct responding. When stimuli had been identified, three stimuli were randomly assigned to each discrimination training condition. Randomization of stimuli across conditions was assessed based on the initial sounds; that is, stimuli with similar, first sounds were not included within the same stimulus set (Wolery et al., 2014). Also, those stimuli that were part of a category (i.e., planets, continents, states, etc.) formed a stimulus set (e.g., Random Rotation: Mercury, Neptune, Jupiter). For one stimulus set (participant 1 modified structured mix condition), two stimuli began with the letter A, but the overlapping sounds were not the same (e.g., Asia and Africa).

Teaching Procedure

Training sessions were conducted in the morning and afternoon, five days a week. Participants received two sessions of each condition, daily. Each condition was counterbalanced during morning and afternoon training sessions. Morning training sessions took place between school arrival and lunch, and afternoon training sessions took place between lunch and school dismissal. For three participants, only one session of each condition was conducted, daily. The presentation of training conditions for these three participants were counterbalanced across days (i.e., Day 1: Random Rotation, Structured Mix; Day 2: Structured Mix, Random Rotation). Sessions consisted of 30 trials. Sessions were discontinued if the participant engaged in challenging behaviors (e.g., self-injury or aggression towards the trainer), or did not respond to instructions or prompts.

Both teaching conditions utilized a discrete trial teaching format (Eikeseth et al., 2014). The trainer presented the antecedent stimulus similarly to the comparison-first procedure (Grow et al., 2011; Grow et al., 2014; Kodak et al., 2015). Each trial began by placing the paper template containing the comparison stimuli in front of the participant on the tabletop before presenting the auditory sample stimulus. Next, the trainer presented an auditory sample stimulus (e.g.,

"cherry tree"). If the participant responded correctly by pointing to the correct comparison stimulus this was reinforced with verbal praise and a tangible or edible item. If the participant responded incorrectly or did not respond, they were told "no" or "try again," and the template was pulled away from the center of the table. On the next trial, the trainer presented the template and initiated a zero-second prompt delay by pointing to the correct comparison stimulus after the presentation of the auditory sample stimulus. If the child had not acquired the stimulus set within 500 training trials, training for that condition was discontinued.

Modified Structured Mix (SM) Condition

The modified SM condition included five steps. For all steps that included the presentation of a new sample stimulus the trainer initiated a zero-second prompt delay on the first trial only (i.e., Step 1, 2, and 4). On steps where two comparison stimuli were available (i.e., Step 1, 2, and 3) one position (right, middle, or left) was blank across the session. If the participant responded correctly with a zero-second prompt on the initial trial, the experimenter moved to the next trial on the pre-made data collection sheet and provided no prompt. If the participant responded correctly on the next trial, this was scored as correct and counted towards the mastery criterion. Next, the trainer continued to rotate the position of the comparison stimuli for the remaining trials; position of comparison stimuli and sample stimulus were dependent on the pre-made data collection sheet. If an incorrect response occurred during the session, error correction procedures were almost identical to the errorless learning procedures (implemented during Steps 1, 2, and 4). That is, following an incorrect response, the position of the comparison-stimuli was the same and the sample stimulus that was incorrect was re-presented with a zero-second prompt delay. Following a prompted response, the position of the comparison-stimuli remained the same and the sample stimulus was presented without a prompt. If the participant responded correctly, this trial was scored as correct, and the trainer rotated the position of the comparison stimuli for the next trial dependent on the data collection sheet. It should be noted that although the participant responded correctly without a prompt, this trial was not counted towards the mastery criterion. Prompted responses were reinforced with verbal praise. See Appendix for a detailed description of each step in the modified SM procedure.

Appendix

Detailed Description of Each Step in the Modified SM Condition

Step 1: Sample Stimulus 1. In Step 1, the trainer labeled stimulus 1 (e.g., "point to 'Africa'") while stimulus 1 and 2

served as the comparison stimuli. The position of both comparison stimuli rotated between the left, middle, and right position semi-randomly across trials. The mastery criterion was four consecutive trials correct within a 30-trial session.

Step 2: Sample Stimulus 2. In Step 2, the trainer labeled stimulus 2 while stimulus 1 and 2 served as the comparison stimuli. The position of both comparison stimuli rotated between the left, middle, and right position semi-randomly across trials. The mastery criterion was four consecutive trials correct within a 30-trial session.

Step 3: Sample Stimulus 1 and Sample Stimulus 2 Structured Mix. In Step 3, stimulus 1 and stimulus 2 were alternated semi-randomly as the sample stimulus across trials. Stimulus 1 and stimulus 2 served as the comparison stimuli and were rotated between the left, middle, and right positions semi-randomly across trials. The mastery criterion was nine out consecutive trials correct within a 30-trial session.

Step 4: Sample Stimulus 3. In Step 4, stimulus 3 was presented as the sample stimulus on all trials, while stimulus 1, 2, and 3 served as the comparison stimuli. The position of the comparison stimuli rotated between the left, middle, and right position semi-randomly across trials. The mastery criterion was four consecutive trials correct within a 30-trial session.

Step 5: Counterbalanced Random Rotation (Random Rotation). Within each block of nine trials, all three stimuli served as the sample stimulus three times each and the comparison stimuli appeared in each position three times each. The mastery criterion was nine consecutive trials correct within a 30-trial session.

Modified Counterbalanced Random Rotation (RR) Condition

The modified RR condition was identical to Step 5 of the modified SM condition except for some differences in the prompting procedure. At the onset of training, a zero-second prompt delay was provided for the first nine trials, during which each sample stimulus was presented three times: once in the left, middle, and right position. If the participant responded correctly across the nine prompted trials, prompting was discontinued. Subsequently, whenever an incorrect response occurred during sessions, a zero-second prompt delay was provided on the next three consecutive trials following an incorrect response. The comparison stimuli and sample stimuli were kept semi-random after an incorrect response, for example, if an incorrect response occurred on stimulus 1 in the middle position (Trial 11), but Trial 12 on the data collection sheet listed stimulus 2 as the sample stimulus in the left position, a zero-second prompt delay was initiated for stimulus 2. The next two prompted responses followed a semi-random presentation based on the position of the comparison stimuli and the sample stimulus presented on the pre-made data collection sheet. Following prompted responses, trials were presented without a prompt. The reason for the difference in the prompt procedure was that within a block of trials for the modified RR procedure a different sample stimulus should be presented on each trial, such that each sample stimulus is presented equally as often. Additionally, the position of the comparison stimuli should vary unsystematically across each trial. Nine consecutive prompted trials would ensure that each stimulus served as the sample stimulus at least once in each position. Prompted responses were reinforced with verbal praise.

Error Analysis and Additional Error Correction Procedures

If after a 30-trial session, a participant had not acquired the discrimination targeted during that session (i.e., fulfilled the mastery criterion of 4/4 or 9/9), error analysis was conducted to identify the extent to which errors occurred due to, for example, win-stay, lose-shift, win-shift, or position bias. A win-stay strategy may occur between different training steps (molar win-stay) of a discrimination training procedure or between consecutive trials (molecular win-stay) (Grow et al., 2011; Lovaas, 2003). That is, rather than attending to the change in auditory-sample stimulus the learner may respond to the stimulus which received reinforcement on the previous trial (molecular win-stay) or the previous step (molar win-stay). A win-shift strategy may occur when intermixing two sample stimuli, where the learner is reinforced for correct responding on trial 1, but the learner may shift to the other comparison stimulus on trial 2 to receive reinforcement. A lose-shift strategy may occur when the learner responds incorrectly to one of the comparison stimuli; thus, shifting to the other comparison stimulus due to a loss of reinforcement on the previous trial. A position bias may occur depending on how the comparison stimuli are arranged; that is, correct responding to the sample stimulus in a particular position may reinforce a higher percentage of responses to that specific position. Error analysis was taken across both training conditions, but additional procedures to correct for errors were only implemented for the modified SM condition. The reasoning for implementing error analysis and additional error correction procedures was from past research suggesting that the SM training condition could lead to faulty stimulus control (Green, 2001; Grow et al., 2011). Additional error correction procedures, following error analysis, are described in Table 2.

Maintenance

Maintenance tests were conducted four and six weeks after mastery for each condition. Maintenance tests followed the same format as the pretest.

Procedural Integrity

Procedural integrity was conducted by the lead classroom teacher or the experimenter. Data were collected on sections labeled: (a) preparing the session, (b) presentation of the sample and comparison stimuli, (c) prompting, (d) reinforcement, and (e) session structure. Preparing the session was scored as correct if necessary teaching materials and reinforcers were available and a positive relationship (i.e., child was smiling and interacting with the teacher) with the child had been established. During the presentation of the sample and comparison stimuli, the observer recorded whether clear instructions were provided to the child (i.e., presentation of only one sample stimulus indicated on the data collection sheet); if the instruction was appropriate (i.e., instruction followed the data collection sheet); and if the child was attentive during the presentation of the instruction (i.e., hands still, sitting upright in chair, looking at the instructor, looking at the visual array). Correct prompting included collecting data on whether the prompt led to a correct response, whether the correct prompt delay was provided and whether the correct number of responses were prompted. Reinforcement was scored as correct if praise and tangible reinforcers were used when correct responses occurred, whether reinforcement occurred within 2 seconds of a correct response, whether every correct response was reinforced during acquisition phases, whether praise was used for prompted responses, and whether incorrect responses were followed by "no" or "try again." The session structure was assessed correct if the trainer followed the correct discrete trial format (i.e., instruction, response, consequence); the inter-trial interval was no longer than 5 seconds; criterion levels were achieved before moving to different steps; and sessions ended on a correct response or with a task

the child could perform correctly (i.e., did not have to be a receptive language task). Additional items listed within session structure assessed if the appropriate amount of time (a minimum of 60 minutes) was left between the two training procedures, and that 30 trials were conducted for both conditions in the morning and afternoon unless unethical to do so. The individual conducting the training was scored on whether they completed the tasks required for each of the five sections listed above. Procedural integrity was collected across all participants for a mean of 39.8% (range, 33% to 40%) of the sessions. Mean scores for procedural integrity were 98.2 % (range, 86% to 100%).

Design

An adapted alternating treatments design (Sindelar et al., 1985) was used to compare the SM condition to the RR condition. To counterbalance sequence effects, the order of training conditions was alternated semi-randomly so that each condition occurred an equal number of times in the morning and in the afternoon.

Results

Figure 1 shows individual data across blocks of trials for all participants (e.g., blocks of 4 trials for steps 1, 2, and 4 for modified SM; or 9 trials for modified RR and 9 trials for modified SM steps 3 & 5). Table 3 shows the number of trials to mastery, number of prompts, number of errors, and percentage correct at maintenance tests conducted four and six weeks after training.

Participant 1 was taught one stimulus set in each condition and acquired the stimulus set in each condition in less than 45 trials, although with 7 trials fewer in the modified RR condition. Six prompts and 3 errors occurred in the modified SM condition

Table 2.

Error Correction Procedures for Participants 5, 8, and 9

Participant	Error pattern	Stimulus set and training step	Error correction procedure	Results
5	Bias towards comparison stimulus reinforced on previous trials.	Stimulus Set 2 and 3, training Step 3 (i.e., Discrimination between Stimulus 1 and Stimulus 2).	1. Retraining from Step 1. 2. Prompting after two incorrect responses was started.	Neither retraining from Step 1 nor prompting after two incorrect responses resulted in acquisition of Step 3. Training on Step 3 was terminated because training had reached 500 trials.
8	Bias towards comparison stimuli not yet trained.	Stimulus Set 1, training Step 1.	Replacing the comparison distractor with a sample stimulus that would serve as stimulus 3.	Resulted in the acquisition of Step 1. Step 2 was subsequently mastered without procedural modifications. Training on Step 3 was terminated because training had reached 500 trials.
9	Bias towards comparison stimulus reinforced on previous trials. Bias towards comparison stimuli not yet trained.	Stimulus Set 1, training Step 3 (i.e., Discrimination between Stimulus 1 and Stimulus 2). Stimulus Set 2, training Step 2.	1. Retraining from Step 1. 2. Replacing the comparison stimulus (distractor) with a neutral distractor that was not intended to be trained.	Retraining from Step 1 did not result in acquisition of Step 3. Training on Set 1 was terminated because training had reached 500 trials. Resulted in the acquisition of Step 2. Steps 3 and 4 were subsequently mastered without procedural modifications. Training on Step 5 was terminated because training had reached 500 trials.

compared to 16 prompts and 2 errors in the modified RR condition. Maintenance was high for both conditions.

Participant 2 was taught one stimulus set in each condition and acquired both in fewer than 40 trials. There was a two-trial difference in the number of trials to mastery across conditions. Six prompts and 1 error occurred in the modified SM condition, compared to 15 prompts and 2 errors in the modified RR condition. Maintenance for both conditions was low.

Participant 3 was taught one stimulus set in each condition and acquired the stimulus set in the modified SM condition in considerably fewer trials (74), than in the modified RR condition (128). The modified SM condition had fewer prompts and fewer errors, but maintenance was higher in the modified RR condition.

Participant 4 was taught two stimulus sets for each condition. Both training procedures were equally effective. Across both stimulus sets, fewer prompts were required for the modified SM condition, but in one stimulus set more errors occurred for the modified SM condition. Maintenance was 100% at four- and six-week follow-up, for both conditions and across both stimulus sets.

Participant 5 was trained on three stimulus sets and mastered stimulus set 1 with considerably fewer trials in the modified SM condition (93), with fewer prompts and errors, compared to the modified RR condition (252). Participant 5 failed to acquire the next two stimulus sets in either condition. For this participant, error analysis was conducted for stimulus set 2 and 3, and the error analysis and additional error correction procedures that were implemented are shown in Table 2. However, these procedures did not lead to acquisition of the discriminations in the modified SM condition. Maintenance was higher in the modified RR condition for stimulus set 1.

Participant 6 was taught three stimulus sets and acquired the first stimulus set with fewer trials in the modified SM condition, but acquired the two next stimulus sets with fewer trials in the modified RR condition. Participant 6 acquired stimulus set 1 in the modified SM condition in 33 trials, which was the minimum number of trials required to master a stimulus set without making any errors. Moreover, in the modified RR condition, stimulus set 3 was acquired in 18 trials, which was the minimum number of trials required to master a stimulus set without making any errors in this condition. In total, the number of prompts and number of errors were higher in the modified RR condition compared to the modified SM condition. For stimulus set 1, maintenance was higher in the modified SM condition at four- and six-week follow-up (67%, 78%), compared to the modified RR condition (44%, 0%). For stimulus set 2, maintenance was higher in the modified RR condition at four- and six-week follow-up (67%, 89%), compared to the modified SM condition

(33%, 22%). For stimulus set 3, maintenance was higher in the modified RR condition (55% at four- and six-week follow-up), compared to the modified SM condition (44% at four- and six-week follow-up).

Participant 7 was taught one stimulus set and acquired the stimulus set in both conditions, but with fewer trials in the modified SM condition. The number of prompts and errors were fewer in the modified SM condition. Maintenance was not collected for this participant.

Participant 8 and 9 both failed to acquire any discriminations in both training conditions for one stimulus set (Participant 8) and two stimulus sets (Participant 9). Error analysis and additional error correction procedures were implemented but did not result in acquisition (see Table 2). The number of prompts were fewer in the modified SM condition, but also more errors occurred in this condition.

Considering participants as a group, the modified SM procedure was more efficient for four of the nine participants, the modified RR procedure was more efficient for one of the nine participants, both procedures were equally efficient for two participants, and finally, neither procedure was effective for two of the nine participants. The mean number of trials needed to acquire each stimulus set across participants in the modified SM condition was 65 trials compared to 105 trials per stimulus set in the modified RR condition. When tallying the total number of teaching trials required, a total of 654 trials were conducted in the modified SM condition compared to 1047 in the modified RR condition.

To compare the results from the present study to the results of DiSanti et al. (2019), we computed the mean number of trials to mastery for a stimulus set for each condition across participants in both studies. In the current study, seven participants were taught a total of 10 stimulus sets in the modified SM condition and 11 stimulus sets in the modified RR condition. In the DiSanti et al. (2019) study (experiment 2), four participants were taught a total of eight stimulus sets in the SM condition and another eight stimulus sets in the RR condition. Participants in the DiSanti et al. (2019) study were comparable in language skills to the participants in the present study. Table 4 exhibits mean number trials (and range) to acquire three auditory-visual conditional discriminations (i.e., a stimulus set) across the two different conditions, across both studies. The table also shows mean cognitive score and mean Vineland Adaptive Behavior Scale (VABS) scores for the participants in both studies. On average, the participants were similar in functioning across the two studies. Results show that a stimulus set was acquired in fewer trials, on average, in the modified SM condition (65 trials), followed by the RR condition (89 trials), the modified RR condition (119 trials), and the SM condition (132 trials).

Table 3.

Number of Trials to Mastery, Number of Prompts, and Number of Incorrect Responses for each Participant

Participants	Stimulus Sets Taught	Trials to Mastery		Number of Prompts		Number of Errors		Percentage Maintenance 4 Weeks		Percentage Maintenance 6 Weeks	
		SM	RR	SM	RR	SM	RR	SM	RR	SM	RR
1	1	44	37	6	16	3	2	100	100	100	100
2	1	38	36	6	15	1	2	33	22	22	33
3	1	74	128	11	54	7	15	33	35	55	89
4	2	192	189	36	92	28	24	100	100	100	100
5	3	93*	252*	331	780	351	281	33	44	33	55
6	3	134	246	19	136	8	37	47	55	48	48
7	1	79	159	10	73	7	22	**	**	**	**
8	1	*	*	161	197	166	117	***	***	***	***
9	2	*	*	326	641	302	222	***	***	***	***

Note. SM = modified structured mix condition; RR = counterbalanced random rotation condition. Bolded numbers in either the SM or RR condition under Trials to Mastery represent the most efficient condition. *= Mastery was not achieved for one or more stimulus sets. **= Not assessed. ***= Not assessed because the discriminations were not acquired. The 500-trial limit includes the sum of the number of trials to mastery, number of prompts and number of errors for one condition.

Table 4.

*Comparison of Results from The Present Study and DiSanti et al. (2019)*

| Condition	Mean number of trials to mastery per stimulus set	Range	N	Number of stimulus sets mastered	Mean Developmental age (months)	Mean VABS ABC	Mean VABS Communication
Modified SM	65	33-107	7	10	45	62	62
RR*	89	57-173	4	8	41	67	64
Modified RR	119	18-263	7	11	45	62	62
SM*	132	119-147	4	8	41	67	64

Note: * Data extracted from DiSanti et al. (2019), Experiment 2.

Figure 1.

Individual data across modified SM and RR conditions.
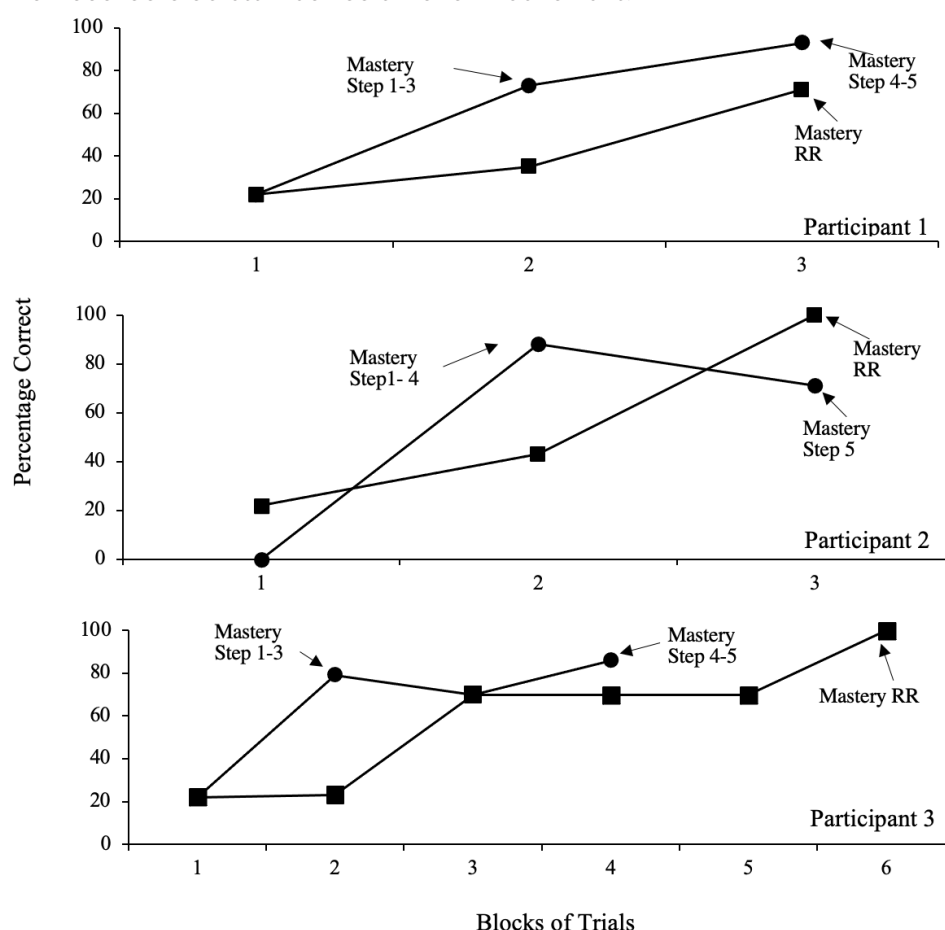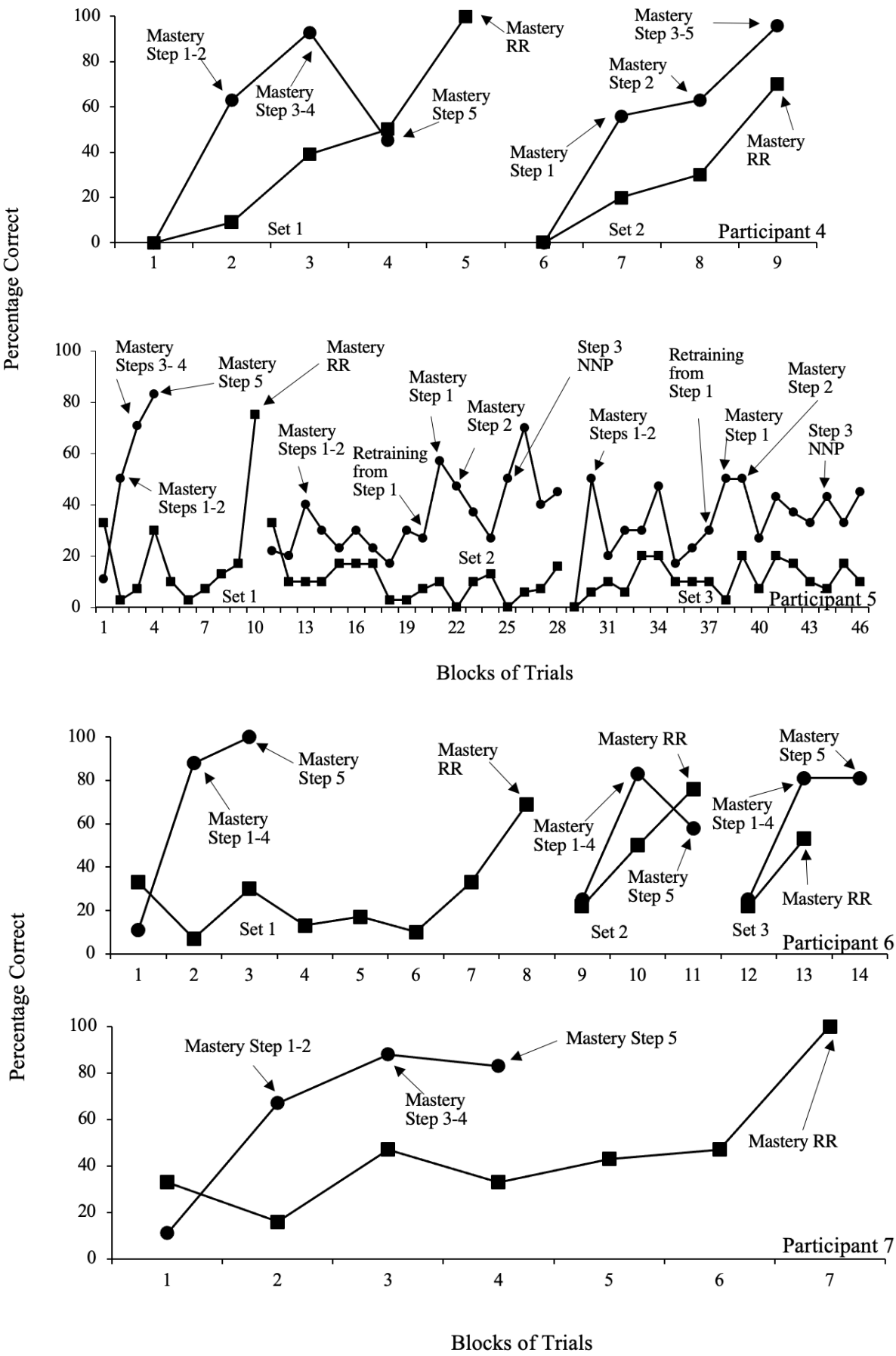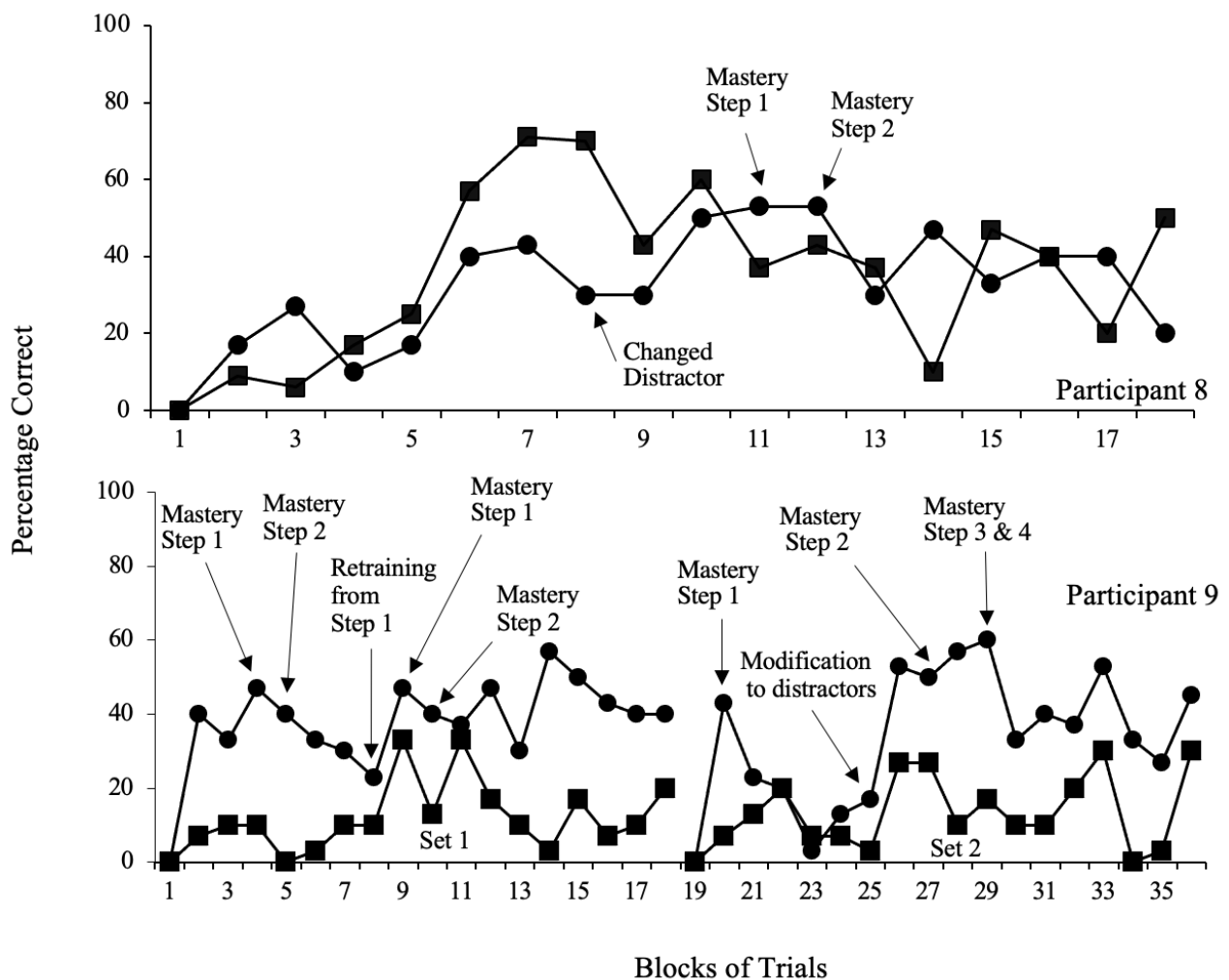
Note. Percentage of correct responses per session for each participant across both conditions (Modified Structured Mix and Modified Random Rotation). Individual data points are represented across blocks of trials (e.g., blocks of 4 trials for Steps 1, 2, and 4 for modified SM; or 9 trials for modified RR and 9 trials for modified SM Steps 3 & 5). Sessions contained a maximum of 30 trials. The numbered arrows represent where each step of the Structured Mix condition was mastered. The arrows also represent where error correction procedures were conducted in the Structured Mix condition (Participants 5, 8, and 9).

Discussion

We evaluated the efficiency of a modified SM procedure and a modified RR procedure for teaching auditory-visual conditional discriminations (receptive labels) to children with autism. The modified SM procedure contained fewer training steps, fewer prompts, and a less stringent mastery criteria compared to the SM method used in previous studies (DiSanti et al., 2019; Grow, et al., 2011; Grow et al., 2014; Grow & Van Der Hijde, 2017; Gutierrez et al., 2009; Holmes et al., 2015; Lin & Zhu, 2019; Vedora & Grandelski, 2015). Refer to the introduction section for a detailed description of the modifications made to the training steps, prompt procedures, and mastery criteria for the SM and RR procedures.

The efficiency of the two procedures were assessed by comparing the number of trials to mastery for establishing three auditory-visual conditional discriminations (i.e., a stimulus set) using the two different discrimination training procedures. Results showed that the modified SM procedure was more efficient for four of the nine participants, the modified RR procedure was more efficient for one of the nine participants, both procedures were equally efficient for two participants, and neither procedure was effective for two of the nine participants. For the stimulus sets which were mastered, the number of errors across the two conditions were comparable, except for two participants who were characterized as having limited verbal repertoires.

Maintenance was assessed four and six weeks after training for nine stimulus sets (Table 3). Across four of the stimulus sets, better maintenance scores were observed in the modified RR condition. For three stimulus sets, better maintenance was observed in the modified SM condition. Maintenance was equally effective for two stimulus sets (participant 4, stimulus set 1 & 2).

For one third of the participants, the modified RR was either equally efficient (two participants) or more efficient (one participant). These three participants had the highest scores on the communication subscale of the Vineland Adaptive Behavior Scale. This may suggest that for participants with an advanced verbal repertoire the modified RR procedure may be most efficient, which could be a topic for further research. For example, future research could compare the two procedures when teaching auditory-visual conditional discriminations to individuals with an intact verbal repertoire.

Based on available data in the present study, there is no clear-cut answer to the question of which approach is most efficient when teaching auditory-visual conditional discriminations to children with autism. Results from the current study suggest that a structured mixed procedure, in general, is more efficient when training steps are reduced, and prompt and mastery criterion are less stringent (i.e., the modified SM procedure used in the current study). Preliminary data suggest that the modified SM procedure is more efficient compared to the RR procedure, but this finding warrants replication, considering there was no intra-participant replication for most participants, except for two (P4 and P6). The modified SM procedure appeared to be more efficient for children with a limited verbal repertoire (DiSanti et al., 2019; Lin & Zhu, 2019); whereas preliminary data may suggest the RR procedure is more efficient for children with a more advanced verbal repertoire. This observation warrants further research.

For those children with an advanced or intact verbal repertoire, perhaps the SM procedure interferes with the acquisition of conditional control due to the reinforcement contingencies involved; that is, by reinforcing irrelevant sources of stimulus control that competes with the desired type of stimulus control required to eventually establish the conditional discrimination. This could be to select the comparison stimulus that was the S+ on the previous trial, select the comparison stimulus which has a stronger reinforcement history across trials or sessions, or to select the stimulus occurring in a particular position; all of which may interfere with sample-stimulus control. However, some children with a more limited verbal repertoire may not have learned to be affected by such reinforcement contingencies. Consistently selecting the S+ and not S- across successive trials and responding to reversal of S+ and S- functions may likely be a prerequisite for learning conditional discriminations (McIlvane et al., 1990). The SM procedure and the modified SM procedure may facilitate the establishment of these prerequisites for some participants.

Another reason why the SM procedure may be efficient is that a two-choice format is used when the first two stimuli are intermixed. This may have simplified the discrimination task and somehow facilitated learning of the sample-comparison relation. Correct responding during this part of training cannot occur unless the sample stimuli exert some type of control over comparison selection. The desired type of stimulus control would be to select S1 when hearing the name of S1 and select S2 when hearing the name of S2 (when S1 and S2 comprise the comparison array). However, it cannot be inferred from the data whether this type of stimulus control had emerged. Alternatively, the learner may have selected S1 and not S2 when hearing the name of S1, and excluded or rejected S1, hence selecting S2 by default when hearing the name of S2. In this case, a reject-relation has emerged (e.g., when the learner hears the name of S2, the learner rejects the previously reinforced S1 and selects the other comparison stimulus; Johnson & Sidman, 1993). Stimulus control of S2 selection when hearing the name of S2 may subsequently transfer from rejecting S1 to selecting S2, consequently resulting in the desired type of stimulus control. To ensure that a conditional discrimination has been established and that the child has learned receptive labeling, it is necessary to increase the stimulus array from two stimuli to three or more stimuli (Carrigan & Sidman, 1992; Sidman, 1987), but perhaps only after the discriminations in the two-choice format has been established. This could be a topic for future research.

Comparing the results from the modified RR procedure (current study) with results from the RR procedure (DiSanti et al., 2019), suggests that the RR procedure was more efficient than the modified RR procedure (Table 4). The mean number of trials were greater for the modified RR condition (119, compared to 89 in the RR condition), the maximum number of trials required for the modified RR condition were greater (263, compared to 173 in the RR condition), and the modified RR procedure had a higher range of trials (245, compared to 116 in the RR condition). Though preliminary, this could be because the error correction procedure in the RR procedure contained elements of a SM procedure. It is possible that the modified RR error correction procedure did not provide meaningful influence on learning and inflated the number of prompts required for the modified RR condition; thus, providing fewer opportunities for the participant to respond independently in the modified RR condition. Future research may explore this topic further.

Individual tailoring should be used when deciding which procedure to use for individual children. The current study provides some guidelines for which procedure to use for which children, but as we have seen variability in data exists. Hence, for this reason, which procedure to use for individual children must be

determined empirically. In addition, which procedure is most efficient may gradually change over time. For example, a particular child may initially benefit more from using the SM procedure when learning their first receptive labels. Later, the modified SM procedure may be more beneficial once initial labels are established. Finally, after becoming a prominent learner of receptive labels, the learner may not require the systematic approach offered by the SM or modified SM procedure, and training with the RR procedure will be more beneficial. This could be a topic for future research.

A problem with studying conditional discriminations is that it cannot be observed directly when and how the sample starts exerting comparison control. This process can only be inferred from manipulating various training procedures and observing the extent to which conditional relations result from these manipulations. For example, during mass trial teaching with the modified SM procedure, it is possible that the participants discriminated the sample stimuli to some extent and came to relate them with their corresponding comparison stimuli, even though this was not required by the reinforcement contingency (McIlvane et al., 1990). Unfortunately, it is difficult to verify this experimentally in an auditory-visual conditional discrimination task. This question is more available for experimental examination when using an arbitrary visual-visual conditional discrimination task. Consider the following experiment: In the presence of the visual sample stimulus A1, selection of the visual comparison stimulus A2 and not B2 is reinforced. This training will continue until performance is stable. Next, in the presence of B1, selection of B2 and not A2 is reinforced, and training is continued until performance is stable. So far, A and B stimuli have been trained using mass trials. To assess whether the participants came to discriminate the sample stimuli (A1 and B1) after mass trials, participants can be presented with a choice task, under extinction conditions, where A1 and B1 is presented together with two other stimuli (say X1 and Y1) which have not been part of any previous training. If A1 and B1 have not been discriminated between as part of the baseline training, participants should choose the four stimuli, on average, equally often. Conversely, if there is a bias towards selecting A1 and B1, these stimuli likely are selected because they previously have been discriminated between and associated with reinforcement. Unfortunately, it is difficult to conceive how this experiment would be carried out using auditory-visual conditional discrimination, and indeed, this may be the reason why there is more research available examining visual-visual conditional discriminations compared to auditory-visual conditional discriminations. This is unfortunate, since understanding auditory-visual conditional discriminations is important for understanding the acquisition of listener behavior.

Whenever ineffective, neither procedure appeared to have advantages over the other. Three participants failed to acquire auditory-visual conditional discriminations and every time a participant failed to acquire a stimulus set this happened concurrently in both the modified SM and modified RR conditions. Hence, when a participant did not acquire a stimulus set, one procedure did not produce advantages over the other. Training novel auditory-visual discriminations in isolation before systematically progressing to a conditional discrimination could have produced better outcomes for these learners (Lovaas, 1977). A lack of motivation could also account for these findings. A brief MSWO was conducted before each training session, but the participant could have become satiated on all five stimuli tested in the MSWO. If so, the preference assessment assessed preference between stimuli that were no longer effective as reinforcers. At this stage in training a MSWO identifying 20 new putative reinforcers could have been performed.

Error analysis was conducted across both the modified SM and RR conditions, but modifications were only made to the modified SM procedure; this could be interpreted as a limitation. In the modified SM condition, two types of faulty stimulus control were observed, and both were related to a stimulus bias. One type of stimulus bias was to select the comparison stimulus that had previously been associated with reinforcement (i.e., molar win-stay or molecular win-stay), which in individuals with developmental disabilities is the most typical error pattern observed (McIlvane & Stoddard, 1981). Another type of stimulus bias included selecting the novel comparison stimulus, that is the stimulus not yet trained (i.e., comparison stimulus bias) and selecting a different comparison stimulus after failing to contact reinforcement (i.e., lose-shift). Similar error patterns were observed in the modified RR condition.

Additional error correction procedures were attempted for three of the participants (5, 8, and 9). One error correction procedure consisted of retraining previously mastered steps. This was done because we were concerned that the mastery criterion of four consecutive correct responses was insufficient to maintain correct responding. This error correction procedure was included for two participants (5 and 9) but was not effective. This suggests that the progressive mastery criterion was not the problem. A second type of error correction procedure consisted of prompting after two incorrect responses (rather than after one incorrect response). We wondered whether corrective feedback alone (e.g., "try again") could facilitate learning and reduce the potential for prompt dependency. This procedure was employed for one participant (5) but did not improve performance. A third type of error correction procedure consisted of replacing the S- with a different S- when participants

showed a preference for comparison stimuli not yet trained. This happened for two participants (8 and 9), and for both participants this resulted in the acquisition of the target step. Additional procedural modifications were not made with the modified RR procedure when learners were failing to acquire the discriminations, and this is a limitation of the study.

Future research could compare the modified SM with the original SM procedure, the modified RR procedure with the original RR procedure, and the modified SM with the original RR procedure. Comparisons could also be made where error correction procedures were not included for either procedure, or the modified SM (or RR) procedure with two different error correction procedures. In addition, future research could consider additional error correction procedures, such as manipulation of the sample stimulus to facilitate its discrimination (Saunders & Spradlin, 1989). Examples include: requiring the participants to echo the verbal sample stimulus before presenting the comparison stimuli (providing that the learner has an echoic repertoire); requiring a manual sample stimulus response, such as requiring the participants to point to the sample stimulus before presenting the comparison stimuli; or re-presenting the auditory sample stimulus every two seconds until a comparison response is emitted (Green, 2001). Research indicates that effectiveness of different error correction procedures may vary across individuals (Carroll et al., 2015; Carroll et al., 2018; Leaf et al., 2020). Furthermore, future research may continue to compare the effectiveness of comparison-first or sample-first presentation, as this has been found to be idiosyncratic across literature (Cubicciotti et al., 2019).

For ethical reasons, training was terminated if a stimulus set was not acquired within 500 training trials (DiSanti et al., 2019). It is possible that some participants would have acquired the auditory-visual conditional discriminations if more training had been conducted. For example, one participant reached the final step in the modified SM condition, but training was discontinued after relatively few training trials because the 500-trial-limit was reached. Rather than discontinuing training for this participant, teaching could have continued if progress was made across sessions. Further evaluations should also be conducted to ensure the difficulty between receptive labels trained is comparable across conditions.

In addition to the limitations already mentioned, some additional limitations should be considered. Maintenance scores were relatively low across some participants who acquired the auditory-visual conditional discriminations in either both or one of the conditions. This could be due to the participants acquiring the auditory-visual conditional discriminations within very few sessions and that

no maintenance training was conducted before four- and six-week follow-up. Future research might consider conducting additional maintenance training and follow-up assessments after the discriminations have been mastered.

Declarations

Ethics approval

This study was performed in line with the principles of the Declaration of Helsinki. Approval was granted by the Norwegian Centre for Research Data (NSD).

Consent to participate

Informed consent was obtained from legal guardians.

Conflict of interest: none.

References

Bayley, N. (2005). *Bayley scales of infant and toddler development (3rd ed.).* HarcourtAssessment, Inc.

Carrigan P. F. Jr. & Sidman M. (1982). Conditional discrimination and equivalence relations: A theoretical analysis of control by negative stimuli. *Journal of the Experimental Analysis of Behavior, 58*, 183–204. https://doi.org/10.1901/jeab.1992.58-183.

Carroll, R. A., Joachim, B. T., St Peter, C. C. & Robinson, N. (2015). A comparison of error correction procedures on skill acquisition during discrete-trial instruction. *Journal of Applied Behavior Analysis, 48,* 257-273. https://doi.org/10.1002/jaba.205.

Carroll, R. A., Owsiany, J. & Cheatham, J. M. (2018). Using an abbreviated assessment to identify effective error correction procedures for individual learners during discrete-trial Instruction. *Journal of Applied Behavior Analysis, 51,* 482-501. https://doi.org/10.1002/jaba.460.

Cubicciotti, J. E., Vladescu, J. C., Reeve, K. F., Carrol, R. A. & Schnell, L. K. (2019). Effects of stimulus presentation order during auditory-visual conditional discrimination training for children with autism spectrum disorder. *Journal of Applied Behavior Analysis, 52,* 541-556. https://doi.org/10.1002/jaba.530.

DeLeon, I. G. & Iwata, B. A. (1996). Evaluation of a multiple-stimulus presentation format for assessing reinforce preferences. *Journal of Applied Behavior Analysis*, *29*, 519-533.https://doi.org/10.1901/jaba.1996.29-519.

DiSanti, B. M., Eikeseth, S., Eldevik, S., Conrad, J. M. & Cotter-Fisher, K. (2019). Comparing structured mix and random rotation procedures to teach receptive labeling to Children with Autism. *Behavioral Interventions, 35*, 38-56. https://doi.org/10.1002/bin.1694.

Eikeseth, S., Smith, D. & Klintwall, L. (2014). Discrete trial teaching and discrimination training. In P. Sturmey, J. Tarbox, D. Dixon & Matson (Eds.), *Handbook of early intervention and autism spectrum disorders: Research, practice, and policy* (pp. 229- 253). Springer.

Green, G. (2001). Behavior analytic instruction for learners with autism: Advances in stimulus control technology. *Focus on Autism and Other Developmental Disabilities*, 16, 72-85. https://doi.org/10.1177%2F108835760101600203.

Grow, L. L., Carr, J. E., Kodak, T. M., Jostad, C. M. & Kisamore, A. N. (2011). A comparison of methods for teaching receptive labeling to children with autism spectrum disorders. *Journal of Applied Behavior Analysis*, 44, 475-498. https://doi.org/10.1901/jaba.2011.44475.

Grow, L. L., Kodak, T. M. & Carr, J. E. (2014). A comparison of methods for teaching receptive labeling to children with autism spectrum disorders: A systematic replication. *Journal of Applied Behavior Analysis*, 47, 600- 605. https://doi.org/10.1002/jaba.141.

Grow, L.L. & Van Der Hijde, R. (2017). A comparison of procedures for teaching receptive labeling of sight words to a child with autism spectrum disorder. *Journal of Behavior Analysis Practice*, 10, 62-66. https://doi.org/10.1007/s40617-016-0133-0.

Gutierrez, A., Hale, M. N., O'Brien, H. A., Fischer, A. J., Drocher, J. S. & Alessandri, M. (2009). Evaluating the effectiveness of two commonly used discrete trial procedures for teaching receptive discrimination to young children with autism spectrum disorders. *Research in Autism Spectrum Disorders, 11,* 1-12. https://doi.org/10.1016/j.rasd.2008.12.005.

Holmes, E. J., Eikeseth, S. & Schulze, K. A. (2015). Teaching individuals with autism receptive labeling skills involving conditional discriminations: A comparison of mass trial and intermixing before random rotation, random rotation only, and combined blocking. *Research in Autism Spectrum Disorders*, 11, 1-12. https://doi.org/10.1016/j.rasd.2014.11.013.

Iversen, I. H., Sidman, M. & Carrigan, P. (1986). Stimulus definition in conditional discriminations. *Journal of the Experimental Analysis of Behavior*, 45, 297-304. https://doi.org/10.1901/jeab.1986.45-297.

Johnson, C. & Sidman, M. (1993). Conditional discrimination and equivalence relation: Control by negative stimuli. *Journal of the Experimental Analysis of Behavior, 59,* 333-347.

Kodak, T., Clements, A., Paden, A., LeBlanc, B., Mintz, J. & Toussaint, K. A. (2015). Examination of the relation between an assessment of skills and performance on auditory visual conditional discriminations for children with autism spectrum disorder. *Journal of Applied Behavior Analysis, 48,* 52- 70. https://doi.org/10.1002/jaba.160.

Leaf, J. B., Cihon, J. H., Ferguson, J. L, Milne, C. M., Leaf, R. & McEachin, J. (2020). Comparing error correction to errorless learning: A randomized clinical trial. *The Analysis of Verbal Behavior, 36,* 1-20. https://doi.org/10.1007/s40616-019-00124-y.

Ledford, J. R., Chazin, K. T., Gagnon, K. L., Lord, A. K., Turner, V. R., & Zimmerman, K. N (2019). A systematic review of instructional comparisons in single-case research. *Remedial and Special Education, 1-14.* https://doi.org/10.1177/0741932519855059.

Lin, F. Y. & Zhu, J. (2019). Comparison of two discrimination methods in teaching Chinese children with autism. *Journal of Applied Behavior Analysis, 53,* 1145-1152. https://doi.org/10.1002/jaba.652.

Lovaas, O. I., Ackerman, A., Alexander, D., Firestone, P., Perkins, M., Young, D. B., Carr, E. G & Newsome, C. (1981). *Teaching developmentally disabled children: The ME book.* PRO-ED.

Lovaas, O. I. (1977). *The autistic child.* Irvington Publishers, Inc.

Lovaas, O. I. (2003). *Teaching individuals with developmental delays: Basic intervention techniques.* Pro-Ed.

Michael, J., Palmer, D. C. & Sundberg, M. L. (2011). The multiple control of verbal behavior. *The Analysis of Verbal Behavior, 27,* 3-22. https://doi.org/10.1007/BF03393089.

McIlvane, W. J. & Stoddard, T. (1981). Acquisition of matching-to-sample performances in severe retardation: Learning by exclusion. *Journal of Mental Deficiency Research, 25,* 33- 4 8 . https://doi.org/10.1111/j.1365-2788.1981.tb00091.x.

McIlvane, W. J., Dube, W. V., Kledaras, J. B., Iennaco, F. M. & Stoddard, L. T. (1990). Teaching relational discriminations to individuals with mental retardation: Some problems and possible solutions. *American Journal of Mental Retardation, 95*, 283-296.

Roid, G. H. (2003). *Stanford Binet intelligence scale (5th ed.).* Riverside Publishing.

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. https://doi.org/10.1901/jeab.1989.52-1.

Schopler, E., Van Bourgondien, M. E., Wellman, G. J. & Love, S. R. (2010). *Childhood Autism Rating Scale (2nd ed.).* Western Psychological Services.

Sidman, M. (1987). Two choices are not enough. *Behavior Analysis, 22*, 11-18.

Sidman, M. & Stoddard, L. T. (1966). Programming perception and learning for retarded children. *International Review of Research in Mental Retardation, 2,* 151-208. https://doi.org/10.1016/S0074-7750(08)60206-2.

Sidman, M. & Tailby, W. (1982). Conditional discrimination vs. matching to sample: an expansion of the testing paradigm. *Journal of the Experimental Analysis of Behavior*, 37, 5–22. https://doi.org/10.1901/jeab.1982.37-5.

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

Sparrow, S. S. Cicchetti, D.V. & Saulnier, C. A. (2016). *Vineland adaptive behavior scales (3rd ed.).* Pearson.

Sundberg, M. (2016). Verbal stimulus control and the intraverbal relation. *Analysis of Verbal*

Behavior, 32, 107-124. https://doi.org/10.1007/s40616-016-0065-3.

Vedora, J. & Grandelski, K. (2015). A comparison of methods for teaching receptive language to toddlers with autism. *Journal of Applied Behavior Analysis, 48,* 188-193. https://doi.org/10.1002/jaba.167.

Wolery, M., Gast, D. L., & Ledford, J. R. (2014). Comparison designs. In D. L. Gast & J. R. Ledford (Eds.), *Single case research methodology: Applications in special education and behavioral sciences (*pp. 297-345). Routledge.