

Marquette University

e-Publications@Marquette

Psychology Faculty Research and Publications

Psychology, Department of

6-2022

Comparing Matrix-Training Procedures with Children with Autism Spectrum Disorder

Samantha Bergmann
University of North Texas

Gabriella Van Den Elzen
University of Nebraska

Tiffany Kodak
Marquette University, tiffany.kodak@marquette.edu

Haven Niland
University of North Texas

Desiree Dawson
University of Nebraska

Follow this and additional works at: https://epublications.marquette.edu/psych_fac



Part of the [Psychology Commons](#)

Recommended Citation

Bergmann, Samantha; Van Den Elzen, Gabriella; Kodak, Tiffany; Niland, Haven; and Dawson, Desiree, "Comparing Matrix-Training Procedures with Children with Autism Spectrum Disorder" (2022). *Psychology Faculty Research and Publications*. 560.

https://epublications.marquette.edu/psych_fac/560

Marquette University

e-Publications@Marquette

Psychology Faculty Research and Publications/College of Arts and Sciences

This paper is NOT THE PUBLISHED VERSION.

Access the published version via the link in the citation below.

The Analysis of Verbal Behavior, Vol. 38, No. 1 (June 2022): 24-53. [DOI](#). This article is © Springer and permission has been granted for this version to appear in [e-Publications@Marquette](#). Springer does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from Springer.

Comparing Matrix-Training Procedures with Children with Autism Spectrum Disorder

Tiffany Kodak

Department of Psychology, Marquette University, Milwaukee, WI, USA

Samantha Bergmann

Department of Behavior Analysis, University of North Texas, 1155 Union Circle

Gabriella Van Den Elzen

University of Nebraska Medical Center's Munroe-Meyer Institute, Omaha, NE, USA

Haven Niland

Department of Behavior Analysis, University of North Texas, 1155 Union Circle

Desiree Dawson

University of Nebraska Medical Center's Munroe-Meyer Institute, Omaha, NE, USA

Abstract

Recombinative generalization is the production of responses in the presence of novel combinations of known components. For example, after learning “red triangle” and “blue square,” recombinative generalization is observed when a child can tact “red square” and “blue triangle.” Recombinative generalization can emerge from a history of matrix training, which involves carefully selecting and arranging stimuli and responses along at least two axes and training a subset of responses. With three children with autism spectrum disorder, we compared recombinative generalization of object–action or feature–object tacts when the component stimuli were trained

before combination stimuli, trained along with combination stimuli, or untrained (i.e., combination only). For two participants, training the components along with some combinations led to the most untrained targets acquired without direct teaching. For the other participant, training the combinations only led to the greatest proportion of untrained targets acquired without direct teaching. We discuss stimulus control promoted by each teaching arrangement and suggestions for future research on recombinative generalization.

The emission of novel verbal responses is a crucial component of teaching functional language repertoires to individuals with language deficits, such as a proportion of individuals with autism spectrum disorder (ASD). Interventions that program for generalization might increase intervention efficiency by producing novel verbal behavior without training all responses directly (Stokes & Baer, 1977). Furthermore, procedures that lead to generalization may avoid training rote or inflexible verbal repertoires, a common critique of interventions based on the principles of applied behavior analysis (Greer & Ross, 2008; Stewart et al., 2013) and related to diagnostic criteria for ASD (American Psychiatric Association, 2013).

An example of generalization of verbal behavior is a speaker's emission of novel combinations of known components (Goldstein & Moussetis, 1989). To illustrate this stimulus control phenomenon, consider the child who can tact (i.e., verbal behavior evoked by a nonverbal stimulus and strengthened with generalized reinforcers; Skinner, 1957) after learning "red triangle" and "blue square" but can also tact "red square" and "blue triangle" without additional instruction. These responses occur because the component stimuli (i.e., red, blue, square, and triangle) retain precise and appropriate control when presented in new combinations (Goldstein & Moussetis, 1989). This phenomenon is recombinative generalization, which may result from matrix training.

Matrix training is a strategy that involves arranging instructional components (also referred to as constituents), separable units that can be combined in different ways, along two axes. Individual cells between the axes include the components arranged in combinations, thereby forming a matrix (e.g., Esper, 1925). To illustrate this, an instructor may use matrix training to teach feature–object tacts to expand the verbal repertoire of a child with ASD. The instructor prepares the instructional program by arranging features (e.g., colors; Fig. 1 C1–C5) along the vertical axis and objects (e.g., utensils; Fig. 1 C'1–C'5) along the horizontal axis. The instructor combines the components to fill each cell of the matrix with a feature–object combination (e.g., color–utensil combinations; Fig. 1 D1–D5) and presents each combination to assess the child's responding before training. The instructor carefully selects a subset of combinations, which may or may not include overlapping components (i.e., diagonal or overlap training; described below), to teach with an evidence-based skill-acquisition procedure. After the child learns the subset of combination targets, the instructor presents the untrained combinations to assess correct responding. If the child responds correctly to the untrained combinations, then recombinative generalization is observed. For example (Fig. 1), the instructor selects the stimuli in D1 to teach directly. The instructor teaches "sapphire spatula" and "blush ladle" to the child. After reaching the mastery criterion, the instructor probes responding in the presence of *blush spatula* and *sapphire ladle*. Recombinative generalization occurs when the child says "blush spatula" and "sapphire ladle" without direct training on these stimulus combinations. In the literature, matrix training has led to recombinative generalization across various skills (e.g., tacts, listener discriminations) with adults and children with and without ASD (see reviews by Curiel et al., 2020; Kemmerer et al., 2021).

	<u>Spatula</u> (C'1)	<u>Ladle</u> (C'2)	<u>Peeler</u> (C'3)	<u>Baster</u> (C'4)	<u>Tongs</u> (C'5)
<u>Sapphire</u> (C1)	Sapphire Spatula (D1-1)	Sapphire Ladle (D5-1)	Sapphire Peeler (D4-1)	Sapphire Baster (D3-1)	Sapphire Tongs (D2-1)
<u>Blush</u> (C2)	Blush Spatula (D2-2)	Blush Ladle (D1-2)	Blush Peeler (D5-2)	Blush Baster (D4-2)	Blush Tongs (D3-2)
<u>Gold</u> (C3)	Gold Spatula (D3-3)	Gold Ladle (D2-3)	Gold Peeler (D1-3)	Gold Baster (D5-3)	Gold Tongs (D4-3)
<u>Forest</u> (C4)	Forest Spatula (D4-4)	Forest Ladle (D3-4)	Forest Peeler (D2-4)	Forest Baster (D1-4)	Forest Tongs (D5-4)
<u>Crimson</u> (C5)	Crimson Spatula (D5-5)	Crimson Ladle (D4-5)	Crimson Peeler (D3-5)	Crimson Baster (D2-5)	Crimson Tongs (D1-5)

Fig. 1 Example matrix with feature–object tacts. Example Matrix (D = diagonal; C = component). Colors are along the y-axis and utensils are along the x-axis; diagonal targets are depicted in shaded cells and probe targets are italicized. A 5 x 5 matrix includes a total of 35 targets (i.e., 10 components and 25 combinations) arranged in five diagonals with five combinations in each diagonal. Nonoverlap training includes only a single diagonal (e.g., D1-1 to D1-5; dark shading). Overlap training may include at least two diagonals (e.g., D1-1 to D1-5 and D2-1 to D2-5; light shading)

When considering matrix training, one must identify the type of skill for which recombinative generalization would be a meaningful outcome. An instructor must select responses that have two separable components that can be arranged in different, novel combinations. Multiple components can be targeted in a multi-dimensional matrix (e.g., three components in Kohler & Malott, 2014), but most matrix-training studies taught two-component targets (Kemmerer et al., 2021). The matrix-training literature includes targets such as two-component tacts (e.g., object–action tacts, feature–object tacts, Frampton et al., 2016), spelling words (Tanji & Noro, 2011), subject–action–object sentences (Kohler & Malott, 2014), and action–picture listener discriminations (Axe & Sainato, 2010). Curiel et al. (2020) and Kemmerer et al. (2021) found that vocal tacts were taught most frequently in the matrix-training literature.

After identifying a meaningful skill, the instructional history of the components is determined. Components can be known (i.e., previously acquired) or unknown (i.e., novel, not yet acquired). Although one-quarter of studies included in the review by Kemmerer et al. (2021) did not report a training history of the components, known components have been used most often in matrix-training arrangements (Curiel et al., 2020; Kemmerer et al., 2021). Using known components in matrix training may facilitate component stimulus control when stimuli are presented in different combinations, thereby creating the conditions necessary for recombinative generalization. Nevertheless, unknown components can be used and may increase instructional efficiency if recombinative generalization is achieved without direct training on the components. Alternatively, although it has not been investigated in previous research to our knowledge, it could be possible to achieve recombinative generalization by targeting unknown components and combinations simultaneously. The present study explored the effects of these component-training arrangements on recombinative generalization.

In matrix training, some combinations are taught directly, and other combinations are reserved for generalization assessments. Because of the different preparations used in the extant literature on matrix training, there are multiple training arrangements to consider. The most common teaching arrangement is nonoverlap training (Curiel et al., 2020). Nonoverlap training, also referred to as diagonal training or training along the diagonal, involves teaching combination targets that do not share the same components (e.g., targets labeled D1 in Fig. 1). Another teaching arrangement is overlap training (Curiel et al., 2020). Overlap training

involves teaching combination targets that share the same components (e.g., targets labeled D1 and D2 in Fig. 1; variations include submatrices, stepwise introduction, and edgewise, e.g., Goldstein & Moussetis, 1989). Sometimes, overlap training is used as a remedial procedure when recombinative generalization does not occur with nonoverlap training (Frampton et al., 2019; Pauwels et al., 2015).

Whereas the majority of participants in the matrix-training literature have acquired the combinations taught directly (Kemmerer et al., 2021), the outcomes for recombinative generalization have been more variable (Curiel et al., 2020; Kemmerer et al., 2021). In a systematic review of matrix-training research conducted with individuals with ASD, Curiel et al. (2020) found that matrix training led to an average yield of 69% of targets acquired without direct teaching, but the range was large with outcomes from 0% to 94%. In a review of matrix-training literature conducted across populations, Kemmerer et al. (2021) found that 95% of studies reported mixed findings for recombinative generalization; that is, most studies reported that participants achieved some level of recombinative generalization but to varying degrees. It is possible that differences in instructional history with components, arranging combinations with nonoverlap or overlap training, or the participants' repertoires could help account for some mixed findings. Therefore, the most optimal arrangement for recombinative generalization and yield continues to be an area in need of additional research.

Given variable component-skill training histories and recombinative generalization outcomes in previous research (Curiel et al., 2020; Kemmerer et al., 2021), the purpose of the current study was to evaluate the effects of component training history—whether the components were directly trained or not—on the emission of independent correct responses to untrained combinations with three children with ASD to whom we taught two-component tacts (object–action or feature–object). Specifically, we evaluated whether and the degree to which recombinative generalization occurred when components were (a) trained to mastery before combination training (referred to as Sequential Training), (b) not trained before nor during combination training (referred to as Combination Only), and (c) trained during combination training (referred to as Simultaneous Training). To our knowledge, no previous studies on matrix training evaluated a Simultaneous Training condition. We analyzed which arrangement led to the greatest yield (i.e., the most untrained targets acquired) and duration of training per target.

Method

Participants, Settings, and Materials

Three male children with medical or educational diagnoses of ASD on a waitlist to receive behavior analytic intervention or currently receiving in-home behavior-analytic services participated in the study. All participants had goals related to expanding vocal–verbal repertoires. All participants had histories of restricted and inflexible vocal and non-vocal behavior (e.g., tacting all balls as “bumpy” despite different textures, playing with toy trains in the same order and manner each time) and were reported to have difficulty acquiring new responses in the presence of previously acquired stimuli (e.g., learning “animal” as a category when previously taught to tact “tiger”) or learning new combinations (e.g., bumpy ball, smooth ball, bumpy block). The institution’s human subjects review board approved all procedures, and parents provided consent. Participation was not compensated. No other participants were recruited for this study.

Arthur was a six-year-old, White boy with an educational diagnosis of ASD (i.e., he had not been evaluated by a psychologist). Arthur lived with his mother, father, and sibling. His family spoke English at home. He began receiving six hours per week of in-home behavior-analytic intervention to assess and treat problem behavior after the current study began. Arthur obtained an age-equivalent score of 7 years, 11 months on the Peabody Picture Vocabulary Test, 4th Edition (PPVT-4; Dunn & Dunn, 2007), and he scored 97.5 out of 100 on the Early Echoic Skills Assessment (EESA; Esch, 2008). Arthur communicated using complete phrases and sentences.

Felix was a five-year-old, Middle Eastern boy with a medical diagnosis of ASD. Felix lived with his father, mother, sibling, and grandmother at the time of this study. His family spoke both English and their native language at home. He received approximately 30 hours of in-home behavior-analytic intervention per week. Felix obtained an age-equivalent score of 4 years, 1 month on the PPVT-4 and 3 years, 1 month on the Expressive Vocabulary Test, 2nd Edition (EVT-2; Williams, 2007). He scored 87 on the EESA. Felix emitted several 3- to 5-word mands, tacts, and intraverbals. He had previously learned object–action tacts and listener responses in a matrix-training arrangement similar to Combination Only but failed to exhibit recombinative generalization.

Trenton was a seven-year-old, White boy with a medical diagnosis of ASD. Trenton lived with his mother, father, and sibling. His family spoke English at home. He received approximately 12 hours of in-home behavior-analytic intervention per week. Trenton obtained an age-equivalent score of 4 years, 2 months on both the PPVT-4 and EVT-2. He scored 94 on the EESA and communicated using complete phrases and sentences.

Sessions occurred in private areas of the participants' homes where therapy was typically conducted (i.e., Arthur's playroom, Felix's dining room and basement, and Trenton's living room). Prior to beginning the experiment, the first author described the study to Arthur using a form approved by the Institutional Review Board; Felix and Trenton were not asked to provide vocal assent due to concerns regarding comprehension. Arthur provided vocal assent to participate. The participant and experimenter sat at a table throughout all sessions.

All sessions included toys, edibles, and electronic devices (e.g., tablet) that the participants vocally nominated. Sessions also included data collection materials, timers, and a digital video camera on a tripod. Arthur's instructional materials were four-legged animal figurines (e.g., goat, deer; Table 1) that ranged in size from approximately 6.35 x 1.27 x 5 cm to 15.24 x 7.62 x 7.62 cm. His materials also included laminated tokens and a token board, which he had been exposed to in the context of prior studies. Arthur exchanged his tokens for 200-s access to preferred items after earning 10 tokens (i.e., fixed-ratio 10 schedule; FR10). Edited images of colors and utensils were Felix and Trenton's instructional materials. We created the colors using Adobe™ Photoshop™ software and sized each to occupy the entire screen. We obtained utensil images from a search engine. We used the Lasso Tool™ to remove the background from the images, if necessary, and all utensils appeared on a solid white background. Each image of a utensil was edited to create six stimuli of different colors including silver (a non-target color) and each of the colors in the matrix (e.g., crimson, aqua; see Fig. 1 for an example matrix; see Table 1 for list of stimuli). We used the Paint Bucket™ tool to color the stimuli at 75% opacity. We arranged the stimuli in a Microsoft PowerPoint slideshow with a blank white slide between each trial slide. We converted the slideshow to an Adobe PDF and presented it on an Amazon Fire 8™ tablet.

Table 1 Targets organized by participant and condition

	Condition					
	Sequential		Combination Only		Simultaneous	
Participant	Component 1	Component 2	Component 1	Component 2	Component 1	Component 2
Arthur	Ciervo [deer]	Tose [coughs]	Cerdo [pig]	Come [eats]	Carnero [ram]	Llora [cries]
	Onza [cheetah]	Ronca [snores]	Oveja [sheep]	Ríe [laughs]	Oso [bear]	Bebe [drinks]
	Caballo [horse]	Corre [runs]	Castor [beaver]	Salta [jumps]	Cabra [goat]	Cae [falls]
	Toro [bull]	Gira [spins]	Lobo [wolf]	Cruza [swings]	Gato [cat]	Camina [walks]
	Alce [moose]	Baila [dances]	Vaca [cow]	Escala [climbs]	Mapache [raccoon]	Vuele [flies]
Felix	Olive	Tablespoon	Sapphire	Spatula	Mustard	Pastry brush
	Cream	Masher	Blush	Measuring cup	Aqua	Funnel
	Maroon	Whisk	Gold	Peeler	Coral	Grater
	Violet	Pepper mill	Forest	Baster	Fuchsia	Can Opener
	Navy	Colander	Crimson	Tongs	Moss	Ladle
Trenton	Mustard	Grater	Olive	Masher	Gold	Measuring cup
	Aqua	Funnel	Maroon	Cheese knife	Crimson	Baster
	Moss	Garlic press	Cream	Colander	Forest	Melon baller
	Coral	Zester	Violet	Pepper mill	Blush	Tongs
	Fuchsia	Can opener	Navy	Tablespoon	Sapphire	Peeler

Targets are listed in the order they were arranged in the matrix. Diagonal combinations were formed by presenting Components 1 and 2 together (e.g., carnero llora, mustard pastry brush)

Measurement

Our primary dependent measure was the percentage of *independent correct untrained responses* emitted in the presence of untrained stimuli. An *independent correct untrained response* was scored if the correct tact was emitted within 10 s of the presentation of the untrained visual stimulus (i.e., discriminative stimulus, S^D) and vocal instruction. The participant needed to begin emitting his response within 5 s, but he had an additional 5 s to complete the response due to the number of syllables in some target responses. If independent correct responses were emitted on at least 80% of trials during probes of untrained combinations, then recombinative generalization was considered achieved. We also used independent correct responses to calculate *yield*. Yield was defined as the percentage of untrained targets—components and combinations depending on the condition—to which the participant emitted independent correct untrained responses. We calculated yield for each condition by dividing the number of untrained targets to which the participant responded correctly (i.e., at least 80% in probes) by the total number of learned targets (i.e., the sum of learned untrained and trained targets) in the matrix and converting the quotient to a percentage.

During training sessions, we collected data on *independent correct trained responses*. These responses were emitted in the presence of trained targets. Only independent correct trained responses contributed to the mastery criterion (i.e., two consecutive sessions with at least 90% independent correct responses) and are graphed in the acquisition figures (i.e., Figs. 4, 6 and 8). An *error* was scored if any vocal–verbal behavior other than the target tact was emitted within 5 s of the vocal instruction. A *no response* was scored if no vocal–verbal behavior was emitted during the 5-s response interval. A *prompted correct response* was scored if the correct tact was emitted following the experimenter’s vocal–model prompt.

We also assessed *duration* for each condition. Our main dependent variable for efficiency was total duration per target. To compute total duration per target, we needed to calculate the *total duration of training* and *total duration of thinning*. The *total duration of training* was computed by summing the minutes of each prompt delay session conducted. The *total duration of thinning* was computed by summing the minutes of each praise and tangible thinning session conducted. The *total duration per target* was computed by summing the total duration of training and thinning and dividing by the total learned targets. Baseline and probe sessions were excluded from each duration measure. We considered a condition complete when either (a) we observed recombinative generalization (i.e., at least 80% independent correct untrained responses to untrained combinations in probe sessions), (b) all combination targets were trained, or (c) participation was discontinued.

Interobserver Agreement and Treatment Integrity

An independent observer collected data on independent and prompted correct responses and session duration for 39% of Arthur’s, 33% of Felix’s, and 33% of Trenton’s sessions. Interobserver agreement (IOA) for independent and prompted correct responses was scored using the trial-by-trial method. An agreement was scored when the two observers recorded the occurrence of the same response during the trial (e.g., both observers recorded the occurrence of an independent correct response). We divided the trials for which an agreement was scored by the total number of trials and converted this quotient to a percentage. Calculations of independent and prompted correct responses yielded a mean agreement score of 99% for Arthur and Felix (range, 80–100) and 99% for Trenton (range, 70–100). We used a dichotomous measure to assess IOA for session duration. That is, if the individual session time recorded by each observer differed by no more than 2 s, the session was scored as an agreement. If the total session time recorded by each observer differed by more than 2 s, the session was scored as a disagreement. Data collectors agreed on duration for 99% of Arthur’s, 95% of Felix’s, and 96% of Trenton’s sessions.

An independent observer collected treatment integrity data for 39% of Arthur’s, 33% of Felix’s, and 33% of Trenton’s sessions. A trial was given a score of 1 if the experimenter implemented each step of the experimental

protocol, and a trial was given a score of 0 if the experimenter failed to implement any components (i.e., error of omission) or implemented any additional components (i.e., error of commission). The trial components included (a) presenting the materials; (b) securing attending; (c) delivering the instruction; (d) allowing the correct response interval; (e) delivering a prompt, if necessary; (f) implementing error correction, if necessary (Trenton only); and (g) delivering praise, tokens, tangible items, or edible items according to the current schedule. We calculated treatment integrity by dividing the number of trials implemented correctly by the number of trials in a session and converting this quotient to a percentage. These calculations yielded mean treatment integrity scores of 99% for Arthur (range, 90–100), 97% for Felix (range, 80–100), and 94% for Trenton (range, 60–100).

Stimulus Selection and Echoic Probes

We arranged 15 objects and 15 actions (Arthur) or 15 objects and 15 features (Felix and Trenton; see Table 1 for complete list of targets) into three 5 x 5 matrices (see Fig. 1 for an example matrix) for a total of 75 combination tacts. Each matrix had a total of 35 targets that could be learned (i.e., 10 components and 25 combinations, see Fig. 1). Therefore, a total of 105 new responses were included for each participant. We used a logical analysis to try to equate the targets in each matrix (Wolery et al., 2018) based on visual similarity, overlapping sounds, and number of syllables. For example, we assigned one of three green-hued colors (i.e., moss, olive, forest) to each matrix rather than assigning them to the same matrix. We randomly assigned each matrix to one of the three conditions.

Arthur

Arthur expressed interest in learning Spanish to his mother. Therefore, we selected object–action tacts in Spanish (Arthur’s evaluation was not part of his in-home behavioral intervention services), and we did not conduct probes prior to stimulus selection because it was highly unlikely that he would contact tacts in Spanish in his surroundings. We confirmed that he could not emit any component and combination tacts in baseline assessments. The objects were four-legged animals, and the actions could be demonstrated without additional materials. We altered Spanish reflexive verbs (i.e., removed the reflexive pronoun, e.g., *reírse* became *reír* which was conjugated to *ríe*) to equate grammatical requirements across targets. All verbs were conjugated in the present tense, and we used the endings appropriate for singular, formal, third-person use (Table 1). Six of the actions required the experimenter to emit sounds (e.g., chuckle for *reír* [laugh]), and these were distributed evenly across matrices. Each component tact was two or three syllables. Thus, each target response was two to six syllables (e.g., *oso* [bear], *mapache vuela* [raccoon flies]).

Felix and Trenton

Both participants had treatment goals to increase feature–object tacts and generalized verbal operants. Felix and Trenton’s target responses were feature–object tacts in English. We selected colors as feature targets because they could be edited to appear identical across objects and variations could be shown on a tablet screen. We selected uncommon colors because Felix and Trenton had already acquired tacts of primary and secondary colors (e.g., red, green). We presented all component and combination stimuli in probe trials (identical to baseline conditions described below), and we selected stimuli to which the participant never emitted correct responses for evaluation. Each component tact was one to four syllables. Four-syllable targets were not placed in the same matrix. Thus, each combination tact was two to six syllables (e.g., *blush tongs*, *forest measuring cup*).

Echoic Probes

We conducted echoic probes when selecting stimuli to ensure all participants could engage in a vocalization with point-to-point correspondence to the model or to identify the participant’s closest approximation for each target response (Frampton & Shillingsburg, 2020). Echoic probes were conducted with vocal stimuli in Spanish

for Arthur and English for Felix and Trenton. In the echoic probes, the experimenter modeled the target response without presenting a corresponding visual stimulus (e.g., said “spatula” without showing a picture of spatula) and allowed up to 5 s for the participant to echo. If the participant did not respond, the experimenter repeated the model prompt every 5 s until the participant vocalized. The experimenter recorded the vocalization and provided praise and a token or 20-s access to preferred tangibles or edibles on a variable ratio two (VR2) schedule. We conducted echoic probes for each target response until the participant’s response had point-to-point correspondence with the model or until the target was presented three times. If the participant’s response did not have point-to-point correspondence with the model, his closest approximation was identified and accepted as a correct response throughout the evaluation. Although approximations were accepted, the experimenter always modeled the complete response.

Experimental Design and Training Conditions

We used an adapted alternating treatments design (Sindelar et al., 1985) to compare recombinative generalization in three conditions. One condition leading to recombinative generalization in fewer training phases or fewer sessions demonstrates experimental control within the adapted alternating treatments design (Kazdin, 2011). We conducted the three conditions concurrently, and the order of conditions in the sequence was arranged according to random rotation without replacement. That is, the order of sessions was randomized with the restriction that one session of each condition was conducted before repeating a condition. The number of trials in each session varied based on the session type (e.g., training, probe). Each condition included the presentation of the S^D and corresponding vocal instruction (Fig. 2). The prompt delay and reinforcement procedures were implemented as described below. We began by using the most common approach to training combinations—nonoverlap training. If we did not observe recombinative generalization, we trained additional diagonals until recombinative generalization was observed, all targets were trained, or the study was discontinued.

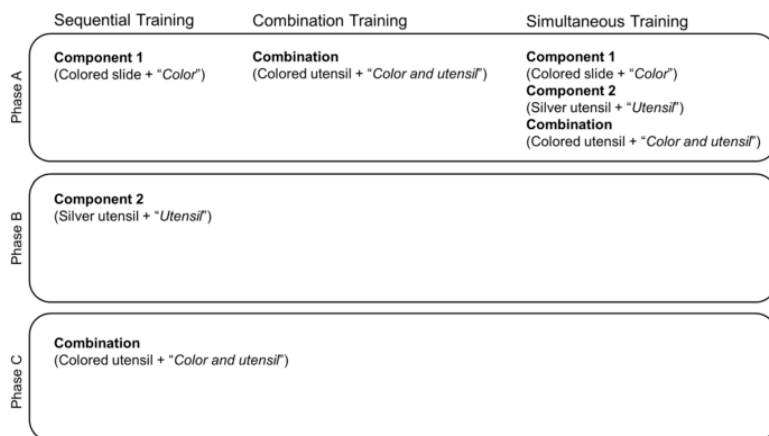


Fig. 2 Trial sequences and introduction of components and combinations in intervention across conditions. The phases represent intervention conducted until the combinations for diagonal 1 were introduced. Mastery criteria for all phases of intervention was 2 consecutive sessions at 90% accuracy. The S^D s and vocal instructions for Arthur were stationary animal + “Tell me this animal in Spanish” for Component 1, moving animal + “Tell me this action in Spanish” for Component 2, and moving animal + “Tell me this animal and action in Spanish” for Combination

Sequential Training

Five component tacts along the vertical axis (e.g., C1–C5, Fig. 1; see Fig. 2) were trained first. Following mastery of the first component, training began with the second components (e.g., C’1–C’5, Fig. 1) until mastery. Next, the experimenter probed all 25 combination targets (e.g., D1–D5 in Fig. 1). If the participants did not emit

criterion-level responding to the untrained combinations, the experimenter trained five combination tacts along the first diagonal of the matrix (e.g., D1-1–D1-5, Fig. 1) in a nonoverlapping arrangement.

Combination Only

Five combination tacts (e.g., D1-1–D1-5, Fig. 1; see Fig. 2) were taught in a nonoverlapping arrangement. The component tacts were not trained prior to nor during combination training.

Simultaneous Training

Ten component tacts (e.g., C1–C5 and C'1–C'5 in Fig. 1) and five combination tacts (e.g., D1-1–D1-5, Fig. 1) were directly trained in a nonoverlapping arrangement at the outset of this condition. The component tacts (e.g., feature, object, action) and the combination tact (e.g., object–action, feature–object) were taught within a chained-trial arrangement (Fig. 2). That is, the experimenter presented the S^D and vocal instruction for the (a) first component (e.g., C1, Fig. 1), then (b) second component (e.g., C'1, Fig. 1), and finally (c) combination (e.g., D1-1 Fig. 1).

Procedure

We conducted 1 to 16 sessions per day, 1 to 5 days per week, depending on the participants' schedules. Arthur's sessions were conducted during 2-hr research sessions, whereas Felix and Trenton's sessions were conducted during their in-home behavior-analytic services. Graduate students studying psychology and behavior analysis implemented the procedures. The experimenters were trained to implement the procedures with high integrity prior to implementing the procedures with participants.

In all phases and conditions (described below), the experimenter initiated each trial by presenting the S^D and securing attending. That is, the experimenter presented Arthur's stimuli by physically manipulating the animal figurines and presented Felix and Trenton's visual stimuli on the tablet. The experimenter waited 5 s for the participant to make eye contact with the materials before presenting a vocal prompt (i.e., "Look.") every 5 s until the participant made eye contact with the materials. Following eye contact, the experimenter delivered the vocal instruction (Fig. 2) and allowed up to 5 s for the participant to respond.

Baseline

All components and combinations were assessed prior to beginning intervention. Each stimulus was shown twice in each session. The experimenter presented the S^D and vocal instruction. They provided 5 s for the participant to initiate an independent response. The experimenter did not deliver programmed consequences for any responses to target stimuli. That is, no praise, token, or preferred item was provided following independent correct responses, and prompts were not provided following errors or no responses. The experimenter presented previously mastered visual stimuli for tact trials about every two trials. The participant had 5 s to tact the mastered targets, and the experimenter modeled a correct response if one did not occur within 5 s. Contingent on independent or prompted correct responses to mastered targets, the experimenter provided praise and a token (Arthur) or praise and 20-s access to tangibles or one small edible (Felix and Trenton). We assessed each untrained component and combination in a session before beginning training. In addition, we conducted additional baseline sessions for components and combinations until a stable or decreasing trend was observed before beginning additional training phases.

Training

Following baseline, a subset of targets—based on the condition and phase (Fig. 2)—was introduced in training. Training included (a) 0-s prompt delay, (b) 5-s prompt delay, and (c) praise and tangible thinning. The progression between training phases was based on specific criteria described below. Initially, we taught combination tacts in a nonoverlapping arrangement wherein a single diagonal was included in training. In

nonoverlap training and component tact training sessions, sessions were 10 trials with each stimulus presented twice and arranged according to random rotation without replacement (i.e., each stimulus was presented once before repeating a stimulus). If nonoverlap training did not lead to recombinative generalization, we used overlap training—wherein we taught additional diagonals. Overlap training sessions were 10 to 20 trials with each stimulus shown once or twice, depending on the number of targets in training. We continued teaching additional diagonals until either (a) the recombinative generalization criterion was achieved, (b) the combinations across all diagonals were directly trained, or (c) participation was discontinued for other reasons.

0-s Prompt Delay

After presenting the S^D and vocal instruction, the experimenter immediately provided a vocal model of the correct response and allowed 5 s for the participant to echo. If the participant echoed, the experimenter provided praise and a token (Arthur) or praise and 20-s access to tangibles or a small edible (Felix and Trenton). If the participant did not echo, the experimenter repeated her vocal model every 5 s until the participant echoed. The experimenter conducted 0-s prompt delay sessions until the participant responded correctly to the model prompt on at least 90% of trials for two consecutive sessions.

5-s Prompt Delay

After presenting the S^D and vocal instruction, the experimenter allowed the participant up to 5 s to begin to emit an independent response. If the participant erred or did not respond within 5 s, the experimenter provided a vocal model of the correct response. The experimenter continued to present her model every 5 s until the participant echoed. The therapist provided praise and a token (Arthur) or praise and 20-s access to tangibles or edibles (Felix and Trenton) following independent or prompted correct responses. Once the participant emitted independent correct responses on at least 40% of trials for two consecutive sessions, tokens and tangibles followed only independent correct responses (i.e., differential reinforcement). That is, praise followed prompted correct responses whereas praise and access to tokens and tangibles were contingent on independent correct responses. We continued 5-s prompt delay sessions until the participant met the mastery criterion, which was two consecutive sessions with at least 90% independent correct responses.

Praise and Tangible Thinning

To make the reinforcement contingencies used in probes less discriminable (Stokes & Baer, 1977), we thinned praise and tokens (Arthur) and praise and tangibles or edibles (Felix and Trenton) following correct responses. After the participant's responding met the mastery criterion, the experimenter provided praise and a token (Arthur) or tangibles or edibles (Felix and Trenton) following independent correct responses on a VR2 schedule rather than following each independent correct response. We conducted sessions on a VR2 schedule until the participant responded correctly to all components and combinations on 90% of trials for three consecutive sessions. Next, the experimenter conducted probe sessions.

The Simultaneous Training condition included an additional thinning step because of the chained-trial arrangement. The experimenter first removed praise following the component tacts (e.g., color and utensil), but she continued to provide praise and a token (Arthur) or praise and access to tangibles or edibles (Felix and Trenton) if correct independent responses occurred for both components and the combination within a trial on a FR1 schedule. Once the participant responded correctly to both component and combination tacts on 90% of trials for three consecutive sessions, the experimenter initiated the VR2 schedule as described above.

Probes

Following thinning of praise and tokens or tangible items, the experimenter assessed responding in the presence of the remaining untrained stimuli in the condition. Probe sessions were conducted for each condition separately. Probe sessions were 5 to 25 trials depending on the number of untrained targets remaining, and

each stimulus was shown once in a probe session. Probe trials were identical to baseline except that stimuli from previously mastered diagonals were interspersed rather than unrelated mastered tacts—this permitted assessment of correct responding in the presence of previously trained combinations—and no prompts were provided following errors to trained targets to reduce the discriminability between contingencies associated with trained and untrained stimuli. The experimenter provided praise and a token (Arthur) or praise and access to tangible items or edibles (Felix and Trenton) following independent correct responses to interspersed mastered targets only. The experimenter conducted at least one probe session for components (Combination Only) and combinations following each training phase with an additional probe session if the participant did not meet the 80% recombinative generalization criterion. Once 80% independent correct untrained responses occurred in a probe session, the components or combinations were considered mastered and were not probed in subsequent sessions.

Procedural Modifications

In some (i.e., 65%; Kemmerer et al., 2021) matrix-training studies, procedural modifications were necessary for a proportion of participants, and we needed to make modifications for Felix and Trenton in the current study due to limited progress. We modified the procedures based on their error patterns and implemented the following modifications across all conditions still in training.

Reinforcer Rotation (Felix)

In the presence of targets with proximal histories of correct responding (i.e., variable, intermediate levels of correct responding with D1–D4 stimuli although mastery was achieved previously), Felix began to emit more incorrect responses. Felix often selected the iPad® in the current study and in other skill-acquisition interventions; nevertheless, the experimenters hypothesized that momentary fluctuations in preference may have affected his behavior during sessions. Therefore, we conducted a paired-choice preference assessment (Fisher et al., 1992) and arranged access to top-ranked tangible items and edibles across days. The experimenter provided the items as assigned during reinforcement intervals; however, Felix’s mands for other tangible items or edibles were reinforced when the item was available.

Repeat Instruction (Trenton)

Trenton frequently echoed the vocal instruction when we began the 5-s prompt delay (e.g., “Color. Crimson” instead of “Crimson”), which violated the operational definition of a target response. Following six sessions of each condition, the experimenter re-presented the vocal instruction with a lower pitch immediately before presenting the model prompt (e.g., “Color. Crimson.”) in a conversational tone, instead of providing the model prompt without re-presenting the instruction (e.g., “Crimson.”). We continued to re-present the vocal model prompt with the vocal instruction until Trenton echoed the component only or 10 prompts had been presented. Prompted correct responses resulted in praise and 20-s access to tangible items. If Trenton did not echo correctly following the tenth model, the experimenter would have moved to the next trial; however, Trenton always echoed with fewer than 10 models. Not only was this presentation akin to other tact and intraverbal instruction in his behavior-analytic programming, but it is based on a procedure used by Kodak et al. (2012) to resolve faulty echoic control.

Error Correction (Trenton)

To increase the number of opportunities to respond correctly to antecedent stimuli and aligned with modifications used in Kodak et al. (2012), we implemented the re-present until correct error-correction procedure (Carroll et al. 2015) across conditions with Trenton. After a prompted correct response, we began the error-correction procedure. In error correction, the experimenter re-presented the trial (i.e., S^D, vocal instruction, and prompt delay procedure) with the same target, and she continued to re-present the trial until Trenton engaged in a correct response to the S^D or 10 error-correction trials had been conducted (the maximum

was never reached). Independent correct responses upon re-presentation in error correction resulted in praise and 20-s access to tangible items until Trenton’s responding met the criterion for differential reinforcement (i.e., two consecutive sessions with 40% or more independent correct responses). Independent correct responses emitted during trial re-presentations in error correction did not count toward mastery of a condition and are not shown in Trenton’s graphs.

Results

Independent correct untrained responses during baseline and probe sessions for Arthur, Felix, and Trenton are shown in Figs. 3, 5 and 7, respectively. Independent correct trained responses during intervention (excluding 0-s prompt delay sessions) for Arthur, Felix, and Trenton are displayed in Figs. 4, 6 and 8, respectively. Typically, data from all conditions in an adapted alternating treatments design are displayed in one, single-panel graph (Sindelar et al., 1985). However, we prepared probe and training data in a non-staggered, multi-panel format to show different component histories, combination targets, and thinning steps across conditions (similar to graphs in Kodak et al., 2020 and Schneider et al., 2021). We conducted all conditions concurrently, as is required in the adapted alternating treatments designs, and time is held constant across acquisition and probe graphs. Data for yield and efficiency are in Table 2.

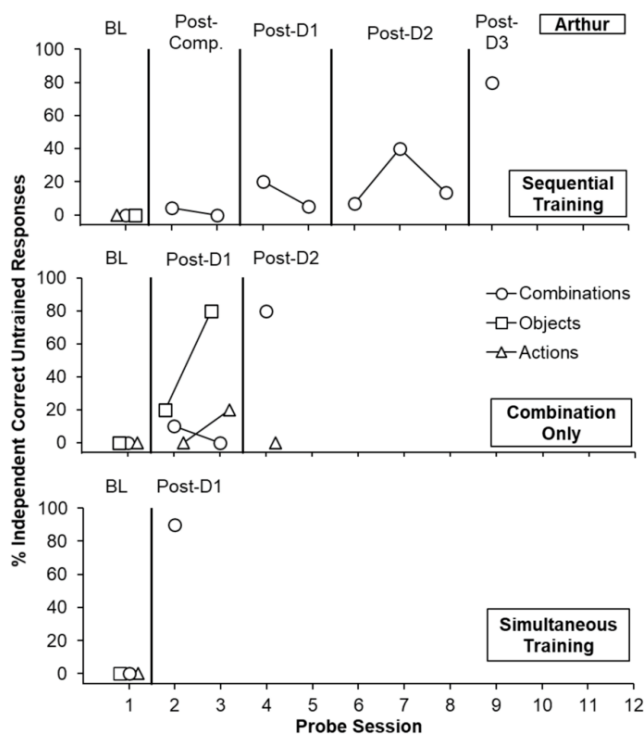


Fig. 3 Probe data for Arthur. BL = baseline, Post-comp. = post component training (C1-C5 and C’1-C’5), and Post-D = post diagonal training. All targets in the diagonal were trained (e.g., D1 = D1-1 to D1-5) during the respective training phase. Mastery criterion was one session with 80% correct and mastered stimuli were not probed in subsequent phases (i.e., objects in Combination Only post-D1)

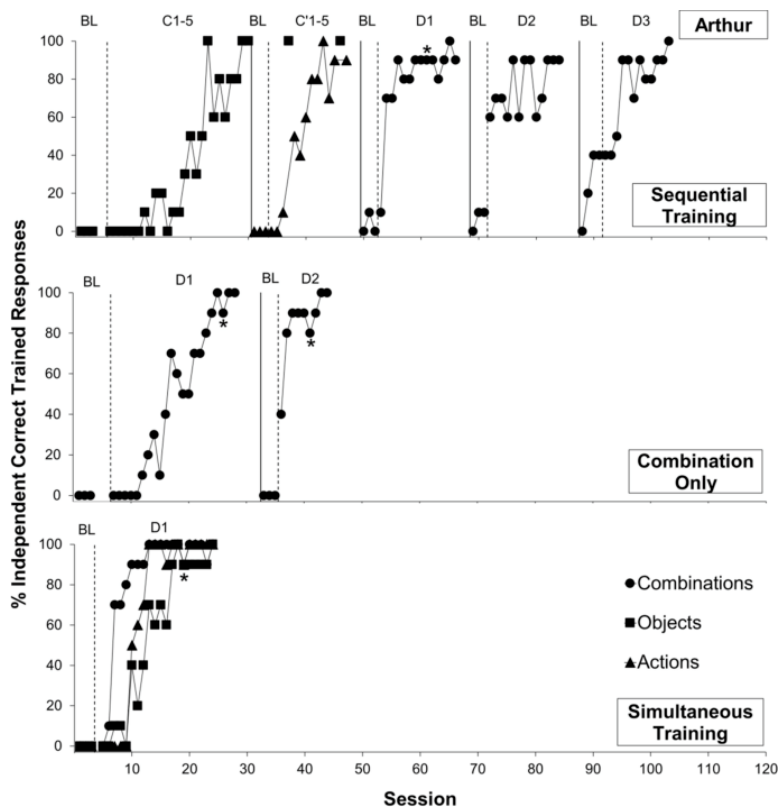


Fig. 4 Acquisition data for Arthur. BL = baseline, C1-5 = objects (animals), C'1-5 = actions, and D = diagonal. Asterisks indicate when praise and token thinning began

Table 2 Summary of diagonals trained, recombinative generalization, proportion of targets acquired, and duration

Participant	Condition	Met RG Criterion	No. of Diagonals Trained	Learned Trained Targets	Learned Untrained Targets	Total Learned Targets	% Learned Untrained Targets	Training Duration (min)	Thinning Duration (min)	Duration per Target (min)
Arthur	Sequential Training	Y	3	25	8	33	24	159.6	57.6	6.6
	Combination Only	Y	2	10	16	26	62	72	16.8	3.4 ^a
	Simultaneous Training	Y	1	15	18	33	55	72	20.4	2.8
Felix	Sequential Training	Y	3	25	10	35	29	273	79.8	10.1
	Combination Only	N ^b	4	20	4	24	17	1692	236.4	80.4
	Simultaneous Training	Y	2	20	14	34	41	233.4	62.4	8.7
Trenton	Sequential Training	N/A ^b	N/A	10	0	10	N/A	339.6	110.4	45.0
	Combination Only	N ^{b,c}	1	5 ^c	4	9	44	249.6	88.2	37.6
	Simultaneous Training	Y	1	15	18	33	54	187.2	357	16.6

Each 5 x 5 matrix included 10 components and 25 combinations for a total of 35 cells; therefore, the maximum number of learned targets was 35. The percentage of learned targets which were untrained represents the yield of each condition. The *Training Duration* column includes all prompt delay sessions prior to meeting the mastery criterion. The *Thinning Duration* column includes instructional time during the praise and tangible thinning phase. The total training time per condition can be obtained by summing the minutes in the Training Duration and Thinning Duration columns. The *Duration per Target* column depicts the total training duration per learned target including trained and untrained targets. Baseline and probe sessions are excluded from all duration totals. Durations are shown in minutes. All participants emitted at least 80% correct responses during object probes in Combination Only, and these data are included in their untrained learned targets total and yield. The format of this table was adapted from Curiel et al. (2020). *RG* Recombinative Generalization

^aComponent stimulus control for verbs did not emerge for Arthur. We continued the train and probe sequence until all diagonals were trained (an additional 184.2 min), and component stimulus control did not emerge. We implemented training with the actions, and Arthur emitted all tacts at mastery level in one session (data available from the corresponding author)

^bCondition was discontinued before recombinative generalization criterion was met

^cTrenton had begun training on the second diagonal (i.e., five additional combinations), but he had not met the mastery criterion prior to discontinuation. Therefore, these targets are not included in his learned trained targets, total learned targets, nor yield calculations

Arthur

Arthur did not engage in any correct responses to components or combinations in the initial baseline phase in any condition (Figs. 3 and 4). In Sequential Training (Fig. 4, top), Arthur required 25 sessions to meet the mastery criterion for objects (animals) and 12 sessions to meet the mastery criterion for actions. He did not emit independent correct responses when presented with the combinations (Fig. 3, Post-Comp., top). Therefore, we trained five combinations (D1-1 to D1-5) until mastery (14 sessions). We conducted probe sessions following three thinning sessions, and Arthur engaged in few independent correct responses to untrained combinations (Fig. 3, post-D1, top). Next, we implemented baseline and training sessions with five more combinations (D2-1 to D2-5) until mastery (11 sessions). We conducted probes following four thinning sessions, and Arthur emitted 40% independent correct responses to untrained combinations (Fig. 3, post-D2, top). The experimenter made an error in Arthur's sessions and trained five more combinations (D3-1 to D3-5) rather than conducting overlap training. Nevertheless, Arthur's responding met the mastery criterion in 11 sessions. We conducted probes following seven thinning sessions, and his responding met the recombinative generalization criterion (Fig. 3, post-D3, top).

In Combination Only (Fig. 4, middle), Arthur's responding met the mastery criterion with five combinations (D1-1 to D1-5) after 19 training sessions. We conducted probes following three thinning sessions, and Arthur's responding did not meet the recombinative generalization criterion (Fig. 3, post-D1, middle). He did, however, emit independent correct untrained responses on 80% of object-component probe trials. Next, we trained five more combinations (D2-1 to D2-5) until mastery (four sessions). We conducted probes following five thinning sessions, and Arthur's responding met the recombinative generalization criterion (Fig. 3, post-D2, middle).

In the Simultaneous Training condition (Fig. 4, bottom), Arthur met the mastery criterion with the components and five combinations (C1-C5, C'1-C'5, D1-1 to D1-5) following 14 sessions. We conducted probes following six thinning sessions, and Arthur's responding met the recombinative generalization criterion (Fig. 3, post-D1, bottom).

Table 2 shows summary data for Arthur's evaluation. All three conditions met the recombinative generalization criterion, but the number of diagonals trained, yield, and duration differed across conditions. Combination Only had the greatest yield (i.e., 62%). Combination Only and Simultaneous Training required the same training time; however, Simultaneous Training required less time per target because more targets were learned.

Felix

Felix did not engage in any correct responses to components or combinations in any condition in the initial baseline phase (Figs. 5 and 6). In Sequential Training (Fig. 6, top), Felix required eight sessions to reach the mastery criterion for features (colors) and 16 sessions to reach the mastery criterion for objects (utensils). Felix did not engage in any independent correct untrained responses following component training (Fig. 5, post Comp., top). We trained five combinations (D1-1 to D1-5) and continued until mastery (four sessions). We conducted probes following three thinning sessions, and Felix's responding did not meet the criterion for recombinative generalization (Fig. 5, Post-D1, top). We trained five more combinations (D2-1 to D2-5) until mastery (four sessions). We conducted probes following three thinning sessions, and Felix's responding did not meet the criterion for recombinative generalization (Fig. 5, post-D2, top). Therefore, we trained both D1-1 to D1-5 and D2-1 to D2-5 in the same session until mastery (eight sessions). We conducted probes following three thinning sessions, and Felix's responding did not meet the criterion for recombinative generalization (Fig. 5, post D1-D2, top), yet he emitted more correct responses compared to previous phases. We trained five more combinations (D3-1 to D3-5) until mastery (nine sessions). We conducted probes following three thinning sessions, and Felix's responding met the criterion for recombinative generalization (Fig. 5, post-D3, top).

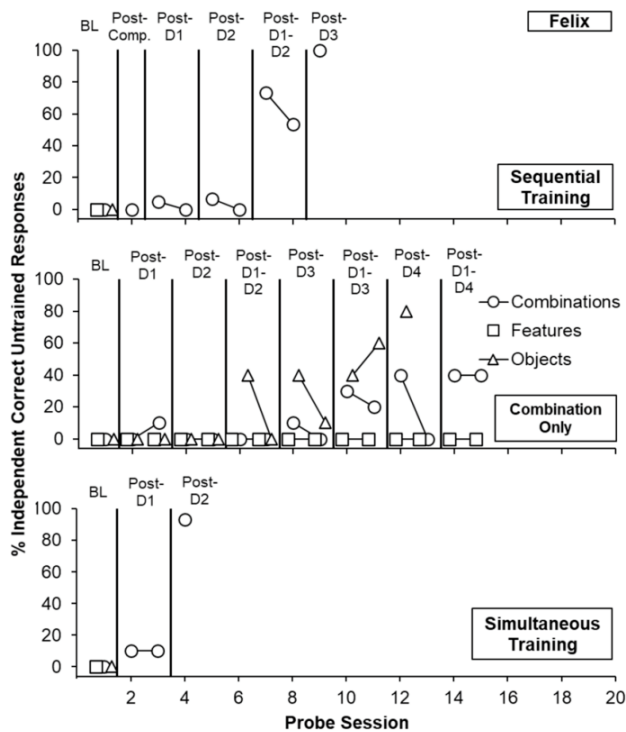


Fig. 5 Probe data for Felix. BL = baseline, Post-comp. = post component training (C1-C5 and C'1-C'5), and Post-D = post diagonal training. All targets in the diagonal were trained (e.g., D1 = D1-1 to D1-5) during the respective training phase. When multiple diagonals are listed (e.g., D1-D2) then targets from multiple diagonals were trained in the same session during that phase. Mastery criterion was one session with 80% correct and mastered stimuli were not probed in subsequent phases (i.e., objects in Combination Only post-D4)

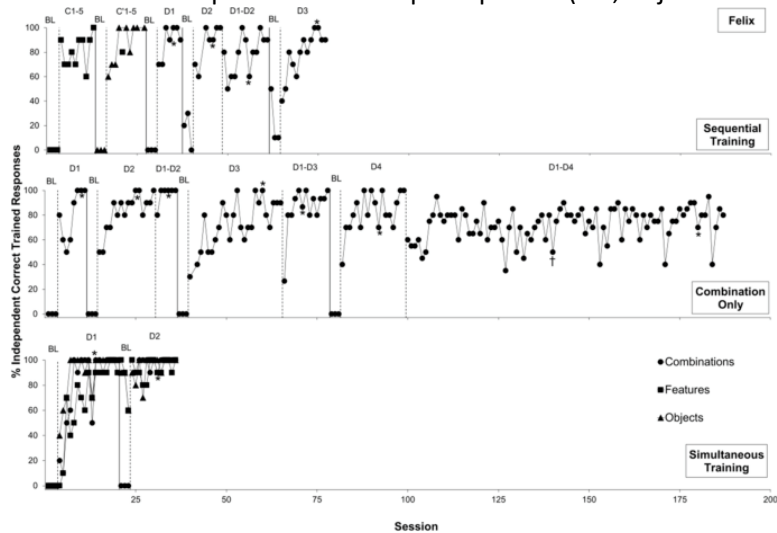


Fig. 6 Acquisition data for Felix. BL = baseline, C1-5 = features (colors), C'1-5 = objects (utensils), and D = diagonal. All targets in the diagonal were trained (e.g., D1 = D1-1 to D1-5) during the respective training phase. When multiple diagonals are listed (e.g., D1-D2) then targets from multiple diagonals were trained in the same session during that phase. Asterisks indicate when praise and tangible thinning began, and the dagger indicates when the reinforcer rotation began

In Combination Only (Fig. 6, middle), Felix required eight training sessions to reach the mastery criterion with five combinations (D1-1 to D1-5). We conducted probes following three thinning sessions, and Felix's responding did not meet the criterion for recombinative generalization (Fig. 5, post-D1, middle). Therefore, we trained five more combinations (D2-1 to D2-5) until mastery (16 sessions). We conducted probes following three thinning

sessions, and Felix's responding did not meet the criterion for recombinative generalization (Fig. 5, post-D2, middle). Next, we trained D1-1 to D1-5 and D2-1 to D2-5 in the same session until mastery (six sessions). We conducted probes following five thinning sessions, and Felix's responding did not meet the criterion for recombinative generalization (Fig. 5, Post-D1-D2, middle). Therefore, we trained five more combinations (D3-1 to D3-5) until mastery (23 sessions). We conducted probes following four thinning sessions, and Felix's responding did not meet the criterion for recombinative generalization (Fig. 5, post-D3, middle). As a result, we trained all three diagonals in the same session (e.g., D1-1 to D1-5, D2-1 to D2-5, and D3-1 to D3-5) until mastery (13 sessions). We conducted probes following eight thinning sessions, and Felix's responding did not meet the criterion for recombinative generalization (Fig. 5, post-D1-D3, middle). We trained five more combinations (D4-1 to D4-5) until mastery (18 sessions). We conducted probes following seven thinning sessions, and Felix's responding did not meet the criterion for recombinative generalization (Fig. 5, Post-D4, middle). Notably, correct responding to 80% of the objects emerged during these probe sessions. Next, we included all four diagonals in the same session (i.e., D1-1 to D1-5, D2-1 to D2-5, D3-1 to D3-5, and D4-1 to D4-5) during which Felix's responding became variable, and he emitted fewer correct responses (Fig. 6, D1–D4, middle). Therefore, we implemented the reinforcer rotation procedure (described above) after 36 training sessions. Felix completed an additional 52 sessions with this modification, but his responding was variable and did not maintain near mastery levels. Felix's participation in the study was discontinued following a discussion with his clinical team. Although his responding did not meet the mastery criterion, we conducted probes with the remaining untrained combinations as well as the features (he previously met 80% correct untrained responses to the objects), and Felix responded below criterion (Fig. 5, post D1–D4, middle). Therefore, his therapists trained the remaining targets (i.e., C1–C5 and D5-1 to D5-5) directly after the evaluation was discontinued.

In Simultaneous Training (Fig. 6, bottom), Felix's responding met the mastery criterion with the components and five combinations (C1-C5, C'1-C'5, D1-1 to D1-5) following 12 training sessions. We conducted probes following five thinning sessions, and Felix's responding did not meet the criterion for recombinative generalization (Fig. 5, post-D1, bottom). We trained five more combinations (D2-1 to D2-5) until mastery (seven sessions). We conducted probes following six thinning sessions, and Felix's responding met the criterion for recombinative generalization (Fig. 5, post-D2, bottom).

Table 2 shows summary data for Felix's evaluation. Two conditions, Sequential Training and Simultaneous Training, met the recombinative generalization criterion, but the number of diagonals trained, yield, and duration differed across conditions. Simultaneous Training produced the greatest yield (i.e., 41%). Simultaneous Training required the least amount of training per target.

Trenton

Trenton did not respond correctly to components or combinations during initial baseline sessions for any conditions (Figs. 7 and 8). In Sequential Training (Fig. 8, top), Trenton's responding to the feature component did not increase after five sessions of intervention. Trenton repeated the vocal instruction (e.g., "Colors."). We added the repeat instruction modification, but this did not change his responding. After six sessions, we added error correction. With error correction, Trenton's correct responding increased, so error correction was included across subsequent intervention phases. Trenton required 36 training sessions to reach mastery with features (colors) and 22 training sessions to reach mastery with objects (utensils). He did not respond correctly to any of the untrained combinations (Fig. 7, post-components, top). We returned to 5-s prompt delay with error-correction for features due to decreased correct responding observed during maintenance probes. Once Trenton responded at mastery levels for both components in isolation, we probed all the combinations. Trenton responded below the recombinative generalization criterion (Fig. 7, post-re-training C1–5, top); thus, training the components in isolation was not sufficient to produce correct untrained responses to combinations. Due to

a shift in treatment goals, Trenton's participation in the study was discontinued. Therefore, we did not conduct training with any combinations.

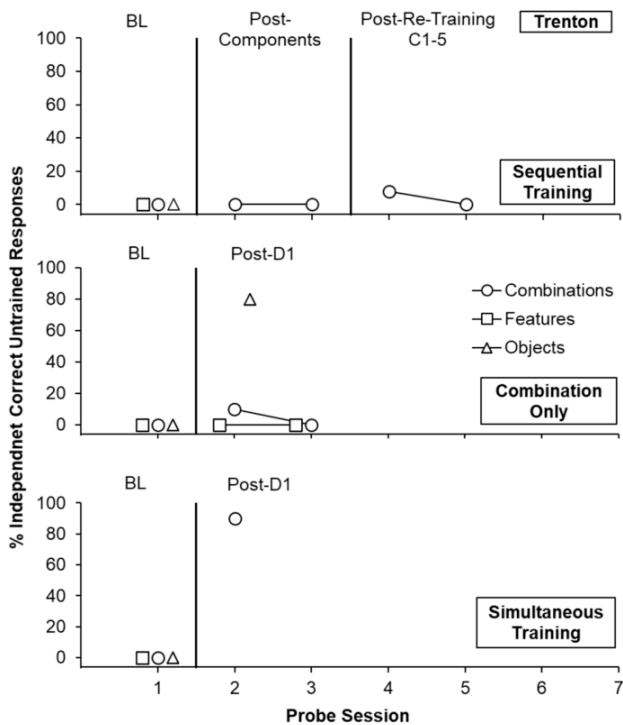


Fig. 7 Probe data for Trenton. BL = baseline, Post-components = post-component training (C1-C5 and C'1-C'5), and Post-D = post diagonal training. All targets in the diagonal were trained (e.g., D1 = D1-1 to D1-5) during the respective training phase

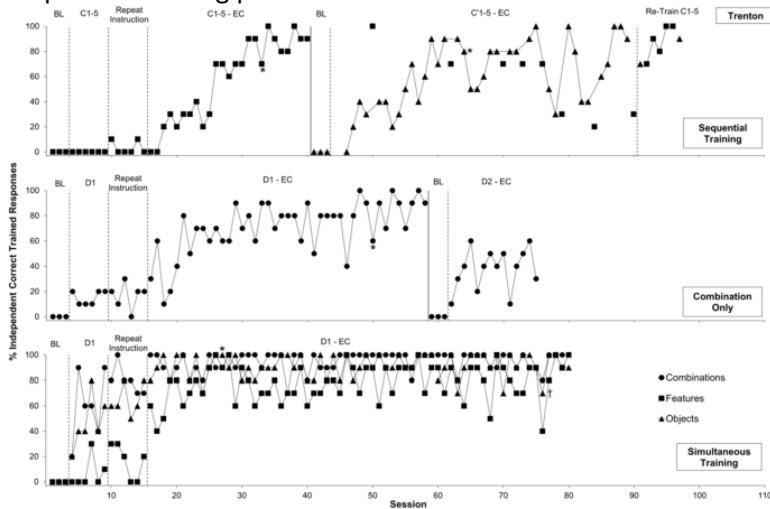


Fig. 8 Acquisition data for Trenton. BL = baseline, C1-5 = features (colors), C'1-5 = objects (utensils), D = diagonal, and EC = error correction. All targets in the diagonal were trained (e.g., D1 = D1-1 to D1-5) during the respective training phase. Asterisks indicate when praise and tangible thinning began. The dagger represents when the modified VR2 began

In Combination Only (Fig. 8, middle), Trenton's correct responding did not increase with five sessions of intervention, so we added the repeat instruction modification. His responding did not change after six sessions, so we added error correction. Trenton's independent correct responding increased with error correction. Trenton required 54 sessions to reach mastery with five combinations (D1-1 to D1-5). We conducted probes following nine thinning sessions, and Trenton's responding did not meet the recombinative generalization

criterion (Fig. 7, post-D1, middle); he did, however, respond correctly on 80% of the object probes. We trained five more combinations (D2-1 to D2-5), but Trenton's participation in this study was discontinued due to a shift in treatment goals. Therefore, we were not able to train the combinations (D2-1 to D2-5) to mastery or conduct subsequent probes.

In Simultaneous Training (Fig. 8, bottom), Trenton's responding was variable, and he repeated the vocal instruction (e.g., "Color."). Because of this, we added the instruction re-presentation modification after five training sessions. Trenton continued to repeat the vocal instruction, so we added error correction after six sessions. With error correction, Trenton required 13 training sessions to meet the mastery criterion with the components and five combinations (C1–C5, C'1–C'5, and D1-1 to D1-5). We removed praise for the individual components in Simultaneous Training, and his correct responding on the feature and object components was frequently below the mastery criterion. Trenton had not met the criterion to move to a VR2 schedule for praise and tangibles after 47 sessions. Following a discussion with his clinical team, we altered the reinforcement thinning procedure to a modified VR2 for correct independent component and combination tacts. When a trial was programmed for praise and tangibles, the experimenter provided praise following independent correct components and the combination. We conducted probes following five sessions with the modified VR2 (i.e., 52 total thinning sessions), and Trenton's responding met the criterion for recombinative generalization criterion (Fig. 7, post-D1, bottom).

Table 2 shows summary data for Trenton's evaluation. Only the Simultaneous Training condition led to recombinative generalization prior to discontinuation, and it produced a yield of 54%. At the time of discontinuation, Simultaneous Training required the least time per target.

Discussion

The current study extends the literature on matrix training and recombinative generalization by evaluating three different arrangements. We evaluated recombinative generalization with (a) Sequential Training, in which components were trained to mastery before training diagonal combinations; (b) Combination Only Training, in which components were not trained prior to or during diagonal training; and (c) Simultaneous Training, in which components were trained along with diagonal combinations. We compared whether the matrix-training procedures resulted in recombinative generalization (i.e., 80% correct untrained responses) and differences in yield (i.e., proportion of targets acquired without direct training).

We found that Simultaneous Training was the only condition to lead to recombinative generalization following training with one diagonal (Arthur and Trenton). This outcome might be anticipated because the components were trained directly, and component stimulus control creates the conditions necessary for recombinative generalization. However, Simultaneous Training and Sequential Training both included directly training component stimuli, but Sequential Training required mastering more diagonals before recombinative generalization was observed for Arthur and Felix (Trenton's Sequential Training was discontinued before recombinative generalization could be assessed).

The relative advantage of Simultaneous Training over Sequential Training may be the ongoing practice of the component responses or the use of a chained-trial presentation. Continued exposure to the components could have created the conditions necessary for the development of appropriate, complete, and accurate stimulus control by the individual components, which is crucial for recombinative generalization (Goldstein & Moussetis, 1989). Another potentially important feature of the Simultaneous Training condition was a chained-trial arrangement wherein the two components always preceded the combination stimulus. That is, the color S^D always preceded the utensil S^D , and both always preceded the colored-utensil S^D (Fig. 2). The chained-trial format may have facilitated responding to novel combinations in probes—possibly because of learning an

autoclitic frame (Skinner, 1957). We presented the probe trials in the same chained-trial arrangement, which could be a limitation of the study. That is, lower levels of correct responding to untrained combinations might have been observed if the component stimuli did not precede the combination stimuli in probes. Future comparisons of Simultaneous Training and Sequential Training should disentangle continued exposure to the components. In addition, investigations of Simultaneous Training should consider conducting probes with and without the chained-trial format.

Sequential Training may have created a lengthy instructional history of practicing the components in isolation. It is unknown whether previous studies that used known components included lengthy instructional histories with the components, or whether they trained the components sequentially (e.g., first master features then introduce objects) or simultaneously (e.g., alternating between components). For example, Frampton et al. (2019) described that component tacts were mastered in isolation prior to the start of the study but how or when was not specified. It could be ecologically valid to teach both components at the same time, perhaps in separate sessions, rather than in the sequential order we selected, and this could be examined in future studies.

Kemmerer et al. (2021) found that most participants acquired the combinations taught directly, and most studies reported mixed findings for recombinative generalization. In the current study, each participant acquired the targets taught directly in all conditions, but recombinative generalization results were mixed. The three matrix-training conditions produced different yields across participants, and the average yield in the current study (i.e., 41%; range, 17–62) was lower than the average reported in the review by Curiel et al. (2020; i.e., 69%; range, 0–94). As shown by the wide yield range reported in the review by Curiel et al., learners may respond differently to matrix-training procedures, and incoming repertoires may help account for some of these differences. For example, Arthur, a six-year-old child who attended a public school, spoke in complete sentences. He emitted correct untrained responses following Combination Only training and had the greatest yield of all participants in the current study (62%; Table 2). On the contrary, Felix was a five-year-old child who received comprehensive behavior-analytic services in his home, emitted 3- to 5-word vocalizations and achieved an age-estimated score of 3 years, 1 month on the EVT-2, suggesting a less developed tact repertoire. Felix was unable to respond correctly to untrained combinations in the Combination Only condition following overlap training with four diagonals. Felix likely acquired the feature–object tacts as inseparable units rather than separable components that retained appropriate control when presented in untrained combinations (Goldstein & Moussetis, 1989). Collecting data on whether responses emitted in the presence of untrained combinations were under the control of the feature or object component would identify sources of faulty control. Future studies should consider examining the relationship between participants' incoming repertoires and recombinative generalization. Additionally, greater attention should be paid to other variables that may impact recombinative generalization such as the history of components, selection of combinations to teach (e.g., Pauwels et al., 2015), and the size of the matrix (Solano et al., 2021) to identify the most optimal procedures.

In addition to yield, we compared the efficiency of each condition by examining the time per learned target. Conclusions regarding the efficiency of the training procedures should be approached cautiously because we did not conduct replications and did not complete the evaluation with Trenton. Due to these limitations, it is not possible to draw firm conclusions regarding the relative yield and instructional efficiency of these matrix-training procedures. That is, without multiple comparisons with each participant, we cannot rule out differences due to history with or difficulty of the targets included in each condition (Wolery et al., 2018), nor can we identify whether the procedure that is most efficient was learner-specific (e.g., Cubicciotti et al., 2019; Schnell et al., 2020). Regardless, several intersubject patterns in the data warrant discussion. We measured the duration of instruction until at least 80% correct untrained responses occurred in the presence of combinations or until participation was discontinued. The condition with the lowest training time per target across participants was Simultaneous Training (Table 2; this comparison is limited for Trenton because the study was discontinued

before Sequential Training or Combination Only met the recombinative generalization criterion). This finding is somewhat surprising given the requirement to emit three responses per trial in Simultaneous Training (i.e., 30 responses per 10-trial session) in the chained-trial arrangement (Fig. 2), whereas the other training conditions required participants to emit one response per trial (i.e., 10 responses per 10-trial session). Future research should conduct this comparison with opportunities for intrasubject replication to examine efficiency. In addition, maintenance should be assessed as this omission was a limitation of the current study.

It is possible that the Simultaneous Training could have been an even more efficient procedure had we not included two steps of thinning (i.e., thinning praise following components and then thinning praise and preferred items following combinations). Thinning praise and tokens or tangibles was designed to prepare the participants for the probe conditions (Pilgrim, 2015). Compared to the dense, fixed schedule of reinforcement in training, probe conditions were programmed with a leaner, variable schedule of reinforcement. Therefore, thinning was used to prevent disrupted performance due to abrupt changes in reinforcement schedules. We used a performance criterion (i.e., three consecutive sessions with 100% correct responding to trained targets) to determine when to move to probes, so this resulted in different thinning durations across participants (Table 2). Felix and Trenton required a considerable number of praise and tangible thinning sessions before meeting the criterion to move to probes. For example, Trenton spent nearly twice the time in the praise and tangible thinning phase compared to the training phase in Simultaneous Training (Table 2). We did not conduct probes prior to thinning, so it is unknown whether and to what degree thinning praise and tangibles was a necessary component of the experiment. Although thinning reinforcement is recommended to promote indiscriminable contingencies (LeBlanc et al., 2002; Stokes & Baer, 1977), recombinative generalization was observed in previous studies that did not report thinning reinforcement before probes (e.g., Frampton et al., 2016; Frampton et al., 2019; Kohler & Malott, 2014; Pauwels et al., 2015). It is possible that including the thinning component in our independent variable contributed to lower yield in the current study. Future research should systematically evaluate the necessity of praise and tangible thinning prior to probes for recombinative generalization.

Two other aspects of our procedures to consider are the use of vocal antecedents and our stimulus presentation. Across all trials, we presented a vocal instruction to specify the controlling component (e.g., “Color,” “Utensil,” “Color and utensil”). The use of these vocal antecedents was intended to evoke the specific response to either the component or combination. However, our stimulus presentation did not necessarily require participants to attend to the vocal instruction to emit a correct response, and the use of these vocal antecedents may be a limitation. To explain, consider the following (Fig. 2): The S^D for “color” was a screen filled with color, the S^D for “utensil” was an image of a silver utensil, and the S^D for “color and utensil” was a picture of the utensil edited to appear in the target color. Had we shown an alternative stimulus presentation—the target utensil in the target color—on each trial with alternating vocal instructions, then the vocal stimulus would have become a vocal conditional stimulus. It is possible that training conditionally from the outset (e.g., Grow et al., 2011; Grow & LeBlanc, 2013) could facilitate acquisition of component stimulus control, help establish an autoclitic frame (e.g., feature–object order), and lead to greater recombinative generalization. It is also possible that teaching without a vocal instruction or with a general one (e.g., “What do you see?”) when the S^D is designed in a way to occasion the specific target response (e.g., colored slide, silver utensil) may promote greater recombinative generalization. Future research could also examine the effects of requiring the learner to emit the autoclitic frame along with the target response (e.g., “Color, crimson” and “Utensil, grater,” degli Espinosa et al., 2020; Meleshkevich et al., 2021) on recombinative generalization.

Arthur and Trenton met the criterion for untrained combinations following nonoverlap training of a single diagonal in Simultaneous Training. Otherwise, training on subsequent diagonals was necessary before responding met the criterion for recombinative generalization, and this diverges from other studies that

observed recombinative generalization following nonoverlap training with unknown components (e.g., Frampton et al., 2019). The current study was a comparison across multiple procedures with three 5 x 5 matrices taught simultaneously with a total of 105 stimuli. It is possible that evaluating all procedures simultaneously in an adapted alternating treatments design affected the efficacy and efficiency of the instructional arrangements. That is, we do not know whether the participants' responding would have met the criterion for recombinative generalization in all conditions and with fewer training sessions had they been exposed to one procedure at a time (i.e., multiple treatment interference). In addition, we did not hold the number of sessions conducted each day constant within or across participants. It is possible that the conditions would have led to yields more comparable to the average reported by Curiel et al. (2020), and this could be addressed in future studies. Other than unique targets assigned to each condition, we did not include other strategies to reduce threats to internal validity (e.g., carryover effects) in an adapted alternating treatments design such as using condition-correlated stimuli or requiring a specific amount of time to pass before conducting the next session (Sindelar et al., 1985). Future comparative analyses should consider including these components in their experimental procedures to protect against threats to internal validity.

The present study supports the utility of matrix training to produce correct responses to novel combinations (i.e., recombinative generalization). Recombinative generalization is part of flexible, generalized verbal repertoires—a crucial goal of early intensive behavioral intervention for many individuals with ASD (Greer & Ross, 2008). To our knowledge, previous studies have not evaluated Simultaneous Training, and our results suggest that Simultaneous Training could produce greater yield relative to Sequential Training or Combination Only and may do so with fewer minutes of training per target. Our results also suggest that training—and perhaps maintenance—of the components is necessary for recombinative generalization as it is unlikely to emerge without direct training of the components, at least for some learners. More research on the Simultaneous Training arrangement with other participants and skills is warranted. In addition, future studies should assess maintenance and also consider generalization to other operants (e.g., listener discriminations of untrained combinations; Jimenez-Gomez et al., 2019).

References

- American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Author. <https://doi.org/10.1176/appi.books.9780890425596>
- Axe, J. B., & Sainato, D. M. (2010). Matrix training of preliteracy skills with preschoolers with autism. *Journal of Applied Behavior Analysis, 43*(4), 635–652. <https://doi.org/10.1901/jaba.2010.43-635>
- 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*(2) 257–273. <https://doi.org/10.1002/jaba.205>
- Cubicciotti, J. E., Vladescu, J. C., Reeve, K., Carroll, 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*(2), 541–556. <https://doi.org/10.1002/jaba.530>
- Curiel, E. S. L., Axe, J. B., Sainato, D. M., & Goldstein, H. (2020). Systematic review of matrix training for individuals with autism spectrum disorder. *Focus on Autism and Other Developmental Disabilities, 35*(1), 55–64. <https://doi.org/10.1177/1088357619881216>
- deglí Espinosa, F., Gerosa, F., & Brocchin-Swales, V. (2020). Teaching multiply controlled tacting to children with autism. *European Journal of Behavior Analysis*. Advance online publication. <https://doi.org/10.1080/15021149.2020.1737407>
- Dunn, L. M., & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test Fourth Edition (PPVT-4)*. Pearson Assessments.
- Esch, B. E. (2008). *Early echoic skills assessment*. AVB Press.

- Esper, E. A. (1925). A technique for the experimental investigation of associative interference in artificial linguistic material. *Language Monographs*, 1, 5–47.
- Fisher, W., Piazza, C. C., Bowman, L. G., Hagopian, L. P., Owens, J. C., & Slevin, I. (1992). A comparison of two approaches for identifying reinforcers for persons with severe and profound disabilities. *Journal of Applied Behavior Analysis*, 25(2), 491–498. <https://doi.org/10.1901/jaba.1992.25-491>
- Frampton, S. E., & Shillingsburg, M. A. (2020). Promoting the development of verbal responses using instructive feedback. *Journal of Applied Behavior Analysis*, 53(2), 1029–1041. <https://doi.org/10.1002/jaba.659>
- Frampton, S. E., Wymer, S. C., Hansen, B., & Shillingsburg, M. A. (2016). The use of matrix training to promote generative language with children with autism. *Journal of Applied Behavior Analysis*, 49(4), 869–883. <https://doi.org/10.1002/jaba.340>
- Frampton, S. E., Thompson, T. M., Bartlett, B. L., Hansen, B., & Shillingsburg, M. A. (2019). The use of matrix training to teach color-shape tacts to children with autism. *Behavior Analysis in Practice*, 12(2), 320–330. <https://doi.org/10.1007/s40617-018-00288-4>
- Goldstein, H., & Moussetis, L. (1989). Generalized language learning by children with severe mental retardation: Effects of peers' expressive modeling. *Journal of Applied Behavior Analysis*, 22(3), 245–259. <https://doi.org/10.1901/jaba.1989.22-245>
- Greer, R. D., & Ross, D. E. (2008). *Verbal behavior analysis: Inducing and expanding new verbal capabilities in children with language delays*. Pearson.
- Grow, L., & LeBlanc, L. (2013). Teaching receptive language skills. *Behavior Analysis in Practice*, 6(1), 56–75. <https://doi.org/10.1007/BF03391791>
- Grow, L. L., Carr, J. E., Kodak, T. M., Jostad, C. M., & Kismore, A. N. (2011). A comparison of methods for teaching receptive labeling to children with autism spectrum disorders. *Journal of Applied Behavior Analysis*, 44(3), 475–498. <https://doi.org/10.1901/jaba.2011.44-475>
- Jimenez-Gomez, C., Rajagopal, S., Nastri, R., & Chong, I. M. (2019). Matrix training for expanding the communication of toddlers and preschoolers with autism spectrum disorder. *Behavior Analysis in Practice*, 12(2), 375–386. <https://doi.org/10.1007/s40617-019-00346-5>
- Kazdin, A. E. (2011). *Single-case research designs: Methods for clinical and applied settings* (2nd ed.). Oxford University Press.
- Kemmerer, A. R., Vladescu, J. C., Carrow, J. N., Sidener, T. M., & Deshais, M. A. (2021). A systematic review of the matrix training literature. *Behavioral Interventions*, 36(2), 473–495. <https://doi.org/10.1002/bin.1780>
- Kodak, T., Fuchtmann, R., & Paden, A. (2012). A comparison of intraverbal training procedures for children with autism. *Journal of Applied Behavior Analysis*, 45(1), 155–160. <https://doi.org/10.1901/jaba.2012.45-155>
- Kodak, T., Halbur, M., Bergmann, S., Costello, D. R., Benitez, B., Olsen, M., Gorgan, E., & Cliett, T. (2020). A comparison of stimulus set size on tact training for children with autism spectrum disorder. *Journal of Applied Behavior Analysis*, 53(1), 265–283. <https://doi.org/10.1002/jaba.553>
- Kohler, K. T., & Malott, R. W. (2014). Matrix training and verbal generativity in children with autism. *The Analysis of Verbal Behavior*, 30(2), 170–177. <https://doi.org/10.1007/s40616-014-0016-9>
- LeBlanc, L. A., Hagopian, L. P., Maglieri, K. A., & Poling, A. (2002). Decreasing the intensity of reinforcement-based interventions for reducing behavior: Conceptual issues and a proposed model for clinical practice. *The Behavior Analyst Today*, 3(3), 289–300. <https://doi.org/10.1037/h0099991>
- Meleshkevich, O., Axe, J. B., & degli Espinosa, F. (2021). Effects of time delay and requiring echoics on answering questions about visual stimuli. *Journal of Applied Behavior Analysis*, 54(2), 725–743. <https://doi.org/10.1002/jaba.790>
- Pauwels, A. A., Ahearn, W. H., & Cohen, S. J. (2015). Recombinative generalization of tacts through matrix training with individuals with autism spectrum disorder. *The Analysis of Verbal Behavior*, 31(2), 200–214. <https://doi.org/10.1007/s40616-015-0038-y>
- Pilgrim, C. (2015). Stimulus control and generalization. In F. D. DiGennaro Reed & D. D. Reed (Eds.), *Autism service delivery: Bridging the gap between science and practice* (pp. 25–74). Springer.

- Schneider, A., Fienup, D. M., Gussin, R., & Moss, P. (2021). A preliminary comparison of mastery criterion frequency values: Effects on acquisition and maintenance. *Behavioral Interventions*. Advance online publication. <https://doi.org/10.1002/bin.1834>
- Schnell, L. K., Vladescu, J. C., Kisamore, A. N., DeBar, R. M., Kahng, S., & Marano, K. (2020). Assessment to identify learner-specific prompt and prompt-fading procedures for children with autism spectrum disorder. *Journal of Applied Behavior Analysis, 53*(2), 1111–1129. <https://doi.org/10.1002/jaba.623>
- Sindelar, P. T., Rosenberg, M. S., & Wilson, R. J. (1985). An adapted alternating treatments design for instructional research. *Education and Treatment of Children, 8*(1), 67–76.
- Skinner, B. F. (1957). *Verbal behavior*. Prentice-Hall, Inc.
- Solano, A. S., Reeve, S. A., Reeve, K. F., DeBar, R. M., Dickson, C. A., & Milata, E. M. (2021). Comparing matrix sizes when teaching direction following to preschoolers with autism spectrum disorder. *Behavioral Interventions, 36*(4), 778–795. <https://doi.org/10.1002/bin.1824>
- Stewart, I., McElwee, J., & Ming, S. (2013). Language generativity, response generalization, and derived relational responding. *The Analysis of Verbal Behavior, 29*(1), 137–155. <https://doi.org/10.1007/BF03393131>
- Stokes, T. F., & Baer, D. M. (1977). An implicit technology of generalization. *Journal of Applied Behavior Analysis, 10*(2), 349–367. <https://doi.org/10.1901/jaba.1977.10-349>
- Tanji, T., & Noro, F. (2011). Matrix training for generative spelling in children with autism spectrum disorder. *Behavioral Interventions, 26*(4), 326–339. <https://doi.org/10.1002/bin.340>
- Williams, K. T. (2007). *Expressive Vocabulary Test Second Edition (EVT-2)*. Pearson Assessments.
- Wolery, M., Gast, D. L., & Ledford, J. R. (2018). Comparative designs. In J. R. Ledford & D. L. Gast (Eds.), *Single case research methodology* (pp. 283–334). Routledge.

Acknowledgments

Portions of the study were conducted when the first three authors were at the University of Wisconsin-Milwaukee. We thank Dayna Costello, Mary Halbur, Katie Hoffert, Terra Jones, Teresa Mayland, Marisa McKee, Courtney Meyerhofer, and Jessi Reidy for their help.