20-7-2024

Effects of Single and Multiple Exemplar Training on Acquisition and Generalization of Functional Skills

Mariana Guimaraes
mg354052@wne.edu

Effects of Single and Multiple Exemplar Training on Acquisition and Generalization of Functional Skills

By

Mariana Chernicharo Guimarães,

M.S. Experimental Psychology: behavior analysis, Pontificia Catholic University of São Paulo, 2010

Dissertation

Submitted to the Department of Psychology

and the College of Arts and Sciences at

Western New England University in partial

fulfillment of the requirements for the

Degree of Doctor of Philosophy

____

Dr. Chata Dickson, Committee Chair

Rachel Thompson

____

Dr. Rachel H. Thompson, Committee Member

Jonathan Pinkston

____

Dr. Jonathan W. Pinkston, Committee Member

William H. Ahearn, Ph.D., BCBA-D, LABA

Digitally signed by William H. Ahearn, Ph.D., BCBA-D, LABA
Date: 2024.09.25 12:54:31 -04'00'

____

Dr. William H. Ahearn, Committee Member

Date: ____________________________

Abstract

Previous research has shown that Multiple Exemplar Training (MET) produces greater generalization than Single Exemplar Training (SET). The current study directly compares the effects of SET and MET using a parallel-treatments design. In the MET condition, we used a simplified approach to general case instruction called an instructional matrix. The instructional matrix borrowed elements from matrix training, such as organizing the relevant stimuli into a table to determine combinations for teaching and subsequent probing for generalization; and, it borrowed from general case instruction, the programming of common stimuli. The training and generalization settings were set up in the same way as multiple exemplar components are arranged in matrix training. In each axis of the matrix, there were variants of a single dimension of the context (e.g., a location or instruction). For each of 6 participants different functional skills were assigned to each of the two methods (SET and MET). Both methods, SET and MET, were effective in teaching the skills, but generalization results were superior using MET in 8 out of the 9 comparisons. Stakeholders were satisfied with outcomes, and they approved the procedures used.

Key words: autism, generalization, instructional matrix, multiple exemplar training, single exemplar training

Effects of Single and Multiple Exemplar Training on Acquisition and Generalization of Functional Skills

The objectives of education for children with developmental disabilities are to teach skills and provide opportunities for the students to increase their independence and facilitate participation in the community (Horner & MacDonald, 1982). A critical piece of these educational goals is to ensure that the learned skills occur under new and relevant conditions. Generalization is defined by behavior's occurrence under new contexts (settings, discriminative stimuli, instructions, and people) (Stokes & Baer, 1977). Generalization is a fundamental concern of applied behavior analysis, and it should be actively programmed to occur in the contexts where it would be functional for the individual and desired by caregivers (Stokes & Baer,1977).

Stokes and Baer (1977) described nine techniques found in the behavior analytic literature that were used to assess or to program for generalization. One of these techniques is training sufficient exemplars. In this case, training is programmed with sufficient exemplars of stimulus conditions or responses to enhance generalization to untrained stimulus and responses. One example of using sufficient exemplars to program for generalization is multiple exemplar training (MET).

Multiple exemplar training is a teaching strategy that focuses on programming for response and stimulus generalization across settings, people, and materials (Erhand et al., 2022). Research has demonstrated that MET is beneficial to produce generalization of community skills (Milata et al., 2019; Risley & Cuvo, 1980; Sprague & Horner, 1984), leisure skills (Bord et al., 2017), self-care skills (Day & Horner, 1986), social skills (Erhand et al., 2022; Gomes et al., 2019; Stokes et al., 1974; Striefel et al., 1976), vocational skills (Horner & McDonald, 1982), and verbal behavior (Curiel et al., 2020; Dass et al., 2018; Garcia-Albea et al., 2014; Goldstein & Mousens, 1989; Groskreutz et al., 2014).

Two approaches to MET that have been effective in producing generalization are general case instruction and matrix training. In general case instruction, the focus is on careful analysis, selection, and presentation of stimulus conditions to teach discriminations and responses across relevant stimuli and situations that the learners may encounter in their lives (Horner et al., 1985; Horner & MacDonald, 1982; O'Neill et al., 2000). For example, Milata et al. (2020) taught three adolescents with ASD a generalized repertoire of chip-debit card. The researchers employed a six-step process for selecting the responses required in each of the four stores that formed an instructional universe. First, they defined the 'instructional universe' as stores close to participants' school and home that differed in stimuli presented on the APM screens. Second, they identified the required responses of using a chip-debit card in each of the four stores. Third, they organized the responses into nine categories. Fourth, the authors defined critical and noncritical stimulus dimensions. Fifth, they identified a critical stimulus dimension specific to each response category. Sixth, they organized noncritical stimulus dimensions into groups. Participants had to emit nine responses to purchasing an item using a chip-debit card: (a) skipping store membership, (b) inserting card, (c) selecting payment type, (d) selecting cashback, (e) selecting donation, (f) entering PIN, (g) leaving card in APM, (h) removing card from APM, and (i) waiting for the cashier. The number of steps and responses required exemplifies that the general case approach requires spending a lot of time planning at the outset to identify all the possible generalization settings and how this diversity of stimuli may require different topographies of responding.

Another instructional strategy that entails the programming of multiple stimulus exemplars is matrix training (Goldstein, 1983a). The procedure consists of training targets that combine at least two sets of stimulus conditions or skill components. The target stimuli exemplars are arranged across the axes in a table. Training starts with a subset of possible combinations of stimuli. In contrast with general case instruction, matrix training is a more structured approach that programs a combination of stimulus and response features during teaching and then probing untrained combinations. Goldstein pointed out that matrix training yields recombinative generalization by promoting differential responding to novel combinations of stimulus components that have been presented in other contexts. A main advantage of matrix training is its overall efficiency (Langton et al., 2020). For example, Curiel et al. (2020) used matrix training to teach two adults with developmental disabilities to tact time increments. The matrix consisted of 12 hours displayed in 12 rows and five-minute increment displayed in 12 columns (e.g., 00, 05, 10). Participants were taught 12 time increments, and probes were conducted to assess 132 untrained time increments. During baseline the two participants responded correctly to one and two time increments that were directly taught and 15 and 24 of the untrained time increments. During the post-assessment, participants responded correctly to 10 and 11 of taught time increments, and approximately half of the untrained time increments. The researchers asserted that direct training of the increments that were seen in the post-assessment would have taken more time than teaching the twelve carefully selected targets.

Both general case instruction (Holth, 2017; Horner et al., 1985; Horner & MacDonald, 1982) and matrix training (Goldstein, 1983b) have been shown to produce generalization. These strategies require a lot of planning and careful analysis of stimuli that might influence generalization (e.g., variations of stimuli and target responses). Multiple exemplar training can be challenging for clinicians to set up or operate without specific training, and this can be a threat to social validity, specifically the appropriateness and acceptability of the procedures (Wolf, 1978). Generalization of learned skills has been studied for over four decades. Over this time strategies such as providing sufficient exemplars and programming common stimuli have been used to program for generalization. However, it is time-consuming and resource intensive for teachers and other clinicians to plan and implement the strategies to promote the generalization of the skills (O'Neill et al., 2000). This might be related to why practitioners usually evaluated generalization only after the intervention ends rather than programmed for it from the outset of the intervention (Stokes & Osnes, 1989). This is sometimes referred to as single exemplar training.

In the behavior analytic literature, we can find examples of the effects of single exemplar training (SET) and multiple exemplar training (MET) on generalization. Across three studies, Horner and his colleagues compared SET and MET and found MET to be superior for producing generalization (Day & Horner, 1986; Horner & McDonald, 1982; Sprague & Horner, 1984). Horner and McDonald (1982) compared effects of SET and MET on generalized vocational skills with four adolescents with a developmental disability. The participants learned to complete tasks of crimping/cutting electrical components. Following the single exemplar phase, the participants learned the electrical component that was trained, and they responded correctly with between two and five of the 20 untrained electrical components. The participants had significant improvement after training with MET (general case); between 15 to 20 untrained probes were performed successfully.

In extending Horner and McDonald (1982), two other studies compared single exemplar instruction (SET) and multiple exemplar training (MET). These studies varied with respect to skills, the number of exemplars, and participant characteristics. Sprague and Horner (1984) taught six adolescent boys with developmental disabilities to purchase items from vending machines, and Day and Horner (1986) taught six adults with developmental disabilities to put on a pullover shirt. In both studies, after training with MET, participants generalize the skill across ten untrained vending machines (Sprague & Horner, 1984) and put on seven or eight untrained shirts (Day & Horner, 1986). Taken together, the studies conducted by Horner and his colleagues have shown that MET produces greater generalization than SET. However, these studies had a sequence confound: in each case, SET phases preceded MET phases, making it impossible to disentangle the potential effects of SET from the effects of MET. SET does not facilitate performance across stimulus variations in the natural environment and the learning process is restricted (Horner & McDonald,1982). The authors promoted the use of MET, but only with careful consideration of which exemplars to use in training.

There is evidence that MET is a good strategy to produce generalization, however, the comparisons of MET and SET are few, and these studies have experimental designs limitations. The fundamental characteristic of any effective applied intervention includes teaching socially important behaviors and programming for generality; that is, that it will be durable over time, appear in relevant environments, and occur with a variety of related behaviors (Baer et al.,1968). Therefore, with the current study, we sought to extend the literature on methods to enhance generalization in three ways: (a) we evaluated a novel approach that combines elements from effective procedures, (b) we compared the effects of single exemplar training and multiple exemplar training on the acquisition and generalization of functional skills, and (c) we used experimental designs that allowed for a comparison between effects of SET and MET without the potential confound of sequence effects.

In the present study we developed and evaluated a strategy for streamlining programming using multiple relevant exemplars that could be used by teachers. The approach borrowed elements from two multiple exemplar teaching strategies. From matrix training, it borrowed organizing the relevant stimuli into a matrix to determine combinations for teaching and subsequent probing for generalization; from general case instruction, it borrowed programming common stimuli. Common stimuli in this case refers to those stimulus components that occur in both settings (training and generalization). The training settings and generalization settings were set up in the same way as multiple exemplar components are arranged in matrix training. In each axis of the matrix, there were variants of a single dimension of the context (e.g., a location or instruction). We will refer to this arrangement as the *instructional matrix*. This strategy is distinct from matrix training because it teaches the same response under different stimulus conditions, and it samples relevant stimulus conditions from the generalization setting. In contrast with matrix training, an instructional matrix is designed to promote stimulus generalization, rather than recombinative generalization. In addition, the instructional matrix is distinct from general case instruction as the stimuli are organized in a matrix and the response required is a simple one-step task.

Across two studies, we compared the effects of SET and MET (instructional matrix) using an experimental design that allowed for the comparison of SET and MET without the potential confound of sequence effects that was present in some of the previous research (e.g., Day & Horner, 1986; Horner & McDonald, 1982; Sprague & Horner, 1984). Given the importance of teaching functional skills to children with developmental disabilities while promoting generalization (Dickson et al., 2014), participants were taught multiple functional skills, with one method, either SET or MET, assigned for each skill. Comparisons between SET and MET were replicated across participants in both studies, and within participants in the second study.

Study 1

Method

Participants

Participants were three children diagnosed with ASD. All the participants attended a behavior analytic school for individuals with developmental disabilities. Rafael was a 12-year-old boy diagnosed with autism, intellectual disability, and attention deficit hyperactivity disorder. He communicated through single-word utterances and an AAC device. His leisure activities were accessed via caregiver report, and these consisted of coloring, listening to music, and playing games on an iPad. Challenging and interfering behavior targeted in his behavior intervention plan included self-injurious behavior, aggression, environmental destruction, bolting, dropping to the floor, and stereotypy.

Lucas was a 12-year-old boy. He communicated through single-word utterances, and his speech was not clear to most listeners. His leisure activities consisted of physical activities, playing games on an iPad, and playing games with a Wii console. Challenging and interfering behavior included perseveration and disruptive touching of people and items.

Felipe was a 9-year-old boy. He communicated through some gestures and manual signs, and his leisure activities consisted of running, watching videos, and playing with musical toys. Challenging behavior included dropping to the floor, mouthing, bolting, and out-of-seat behavior.

Settings and Materials

Sessions were conducted in multiple settings at the school (all three participants) and two off-campus retail stores (Lucas only). Items used for reinforcement were small snacks or access to toys identified by each participants' caregiver. Participants chose the reinforcers prior to the beginning of each session. All sessions were recorded using video recording equipment, and data were collected using a pen and paper. Other materials varied according to the task and are described in the Targeted Skills section below.

Measurement and Response Definitions

Each trial was scored as either a prompted correct response, an independent correct response, or an incorrect response. Independent correct responses were defined individually for each task and each participant, and these definitions can be found in the Targeted Skills section below. For each session, a percent correct score was calculated by dividing the number of independent correct responses by the total number of responses.

Experimental Design

A parallel-treatments design (Gast & Wolery, 1988) was used to compare the effects of SET and MET on the generalization of the trained skills, as measured by comparing performance in generalization probes to those from the baseline sessions. For each participant, two skills were taught in alternating sessions, one using SET and the other using MET. The parallel-treatments design required assigning each training method to its own set of skills, which were judged to be of similar difficulty and not part of the same response class (Gast & Wolery, 1988). Skills were randomly assigned to one method for each participant. The parallel-treatments design allowed for evaluating the effects of two independent variables that were simultaneously introduced for two different skills.

Procedure

Table 1 displays a generic instructional matrix. The columns and rows represent variations of stimuli that were present during training, and the cells show the training or generalization contexts. In the single exemplar training (SET) condition, one skill was taught using only one combination of stimulus conditions. For example, Rafael was taught to respond to the greeting “hi” only in the classroom (A1). In the multiple exemplar training (MET) condition, one skill was taught using three unique combinations of stimulus conditions. For example, Felipe was taught to respond to the instruction “stop” when he was in the cafeteria (A1), to respond to “Felipe” in the classroom (B2), and to respond to “hold on” in the hallway (C3). Experimental sessions for both the SET and MET conditions were conducted by graduate students working toward a MSEd degree in special education, under the author’s supervision. The skills targeted for training were selected based on the participants’ IEP objectives, parents’ and teachers’ requests, and skill repertoire. Targeted skills are listed in Table 2, along with the stimulus variant types that defined the 3 X 3 matrix. For Rafael, the skills selected were responding to a greeting while he was in his classroom (SET) and producing self-identification (MET). For Lucas, the skills selected were producing a bill to pay for an item at a store (SET) and producing self-identification (MET). For Felipe, the skills selected were responding to his name while watching videos (SET) and an instruction to "stop" while walking (MET). These tasks are described in detail in the Targeted Skills section below.

Baseline. Participants chose an item to be used as a reinforcer before the first session of the day from an array of available snacks or toys presented visually and verbally. Prior to each session the participant was brought to the prescribed location at school or store. Once the participant attended to the experimenter or task materials, instructions were delivered, and the trial began. There were no differential consequences for correct or incorrect responses. After each trial, the participant was asked to complete one known gross motor imitation response per trial and the experimenter delivered the selected item.

Each baseline session consisted of two trials. For each condition, SET and MET, nine of these two-trial sessions were conducted. For SET, one for the training context (Table 1; A1) and eight generalization contexts (A2, A3, B1, B2, B3, C1, C2, and C3). For MET, one for each of the three training contexts (Table 1; A1, B2, and C3) and six generalization contexts (A2, A3, B1, B3, C1, and C2). The details of the stimulus-variant combinations for the instructional matrices can be found in Table 2. Participants started training if they did not perform the task in all of the contexts.

Training. For the SET condition, training was conducted in context A1, and for MET training was conducted in contexts A1, B2, and C3. Training sessions were identical to baseline, except that delayed prompting (Oppenheimer et al., 1993) and differential reinforcement were used. The number of trials matched the participants' programming at school. For two participants (Rafael and Felipe), each session consisted of nine trials, and for Lucas, each session consisted of five trials. When a participant responded correctly, the chosen reinforcer and verbal praise were delivered. Incorrect responses were followed by another brief activity and then the next trial. In the MET condition, training in the three training contexts was intermixed in an ABCABC sequence (A1, B2, and C3; Table 1). The criterion to decrease the prompt hierarchy was two sessions with at least eight out of nine trials correct (prompted or unprompted) for Rafael and Felipe. For Lucas, the criterion to decrease the prompt level was one session with at least four out of five trials correct (prompted or unprompted). The mastery criterion in both conditions was two consecutive sessions in a single context with at least eight out of nine or four out of five trials correct and unprompted.

The generalization phase started for both tasks when one context in either SET or MET reached the mastery criterion. This criterion was selected to keep the requirements for an amount of exposure to the two conditions as similar as possible.

Generalization. Following training, the generalization phase began. Sessions in this phase were conducted in the same way as the baseline except that reinforcer was delivered at the end of a session rather than after gross motor imitation trials. Each session consisted of one trial. In the SET condition, eight of the nine contexts were untrained (see Table 1; B1, C1, A2, B2, C2, A3, B3, and C3), and so there were eight types of generalization probes. In the MET condition, six of the contexts were untrained (Table 1; B1, C1, A2, C2, A3, and B3).

Targeted Skills

Each of the six skills taught in Study 1 is described below, including any materials that were specific to that task. Table 2 provides details regarding the stimulus variant types that defined each of the instructional matrices.

Respond to Greetings (Rafael, SET). The stimulus variants for this task were greetings from an adult and location. At the start of each trial, Rafael was with the experimenter in the prescribed location at school. For example, he was sitting down at the chair by the table in his classroom. The experimenter waited for attending and presented one of the greetings (e.g., "hi"). The correct response included Rafael orienting his body towards the experimenter, making eye contact, and waving to the experimenter within 3-5s.

Produce Self-Identification (Rafael and Lucas, MET). The stimulus variants for this task were a question from an adult and location. At the start of each trial, Rafael or Lucas were with the experimenter in the prescribed location at school. For example, in one of the contexts, Rafael was walking in the classroom, and the experimenter walked ahead of him and stopped in front of him and asked, "what is your full name?" Rafael responded by reaching into his pocket to retrieve a self-identification card and handing it to the experimenter.

Produce a Bill (Lucas, SET). The stimulus variants for this task were questions from the cashier and location (stores). At the start of each trial, Lucas was with the experimenter in the prescribed location. For example, he was at the school store choosing a snack, and the cashier would ask, "How would you like to pay?". Lucas responded by reaching into his pocket to retrieve a bill and handing it to the cashier.

Respond to Name (Felipe, SET). The stimulus variants for this task were ongoing activity and location. At the start of each trial, Felipe was with the experimenter in the prescribed location. For example, he was at the cafeteria sitting down in the chair at the table. The experimenter turned the iPad on and gave Felipe a couple of minutes to engage with the iPad. The experimenter would call his name and the participant oriented his body towards the experimenter and made eye contact within 3-5s.

Respond to Instruction to Stop (Felipe, MET). The stimulus variants for this task were instructions and location. At the start of each trial, Felipe and the experimenter were walking together in a specified location in the school. After a short time, the experimenter presented one of three instructions to stop. Felipe put both feet on the floor or ground within 4 s of the experimenter's instruction and did not move beyond a radius of approximately 0.75 m for a few seconds until the experimenter indicated that walking was again permitted. For example, they were walking within the cafeteria and the experimenter presented one of three instructions to stop.

Interobserver Agreement (IOA)

IOA was assessed by having a second observer independently score at least 38% of sessions for each participant in each phase of the study for SET and MET. An agreement between observers was recorded when both observers scored the response for a trial as correct or incorrect. Agreement percentages were calculated by dividing the number of trials on which the observers agreed by the total number of trials in the session and multiplying by 100. Mean agreement across participants was 98% (range, 88%-100%).

Results and Discussion

During baseline, the participants did not respond correctly in any of the contexts during SET and MET. In the training phase, Rafael reached mastery criteria first in the SET condition (responding to greetings; 10 sessions; Figure 1). Rafael did not engage in any independent correct responding in the MET condition (produce self-identification). Lucas learned to produce a bill in one context (SET) in eight sessions (Figure 2). He mastered one of the three contexts of produce self-identification ("Do you need help?" in the annex hallway) within three sessions (MET). And there were some independent responses in the other two training contexts. Felipe did not reach the mastery criterion with respond to name (SET), and his performance was variable (Figure 3). He learned to respond to stop in one of the three contexts ("Felipe" in the classroom; MET) in 16 sessions, and the data were variable for the other two contexts.

During the generalization phase, Rafael responded correctly to the one context that was trained for responding to greetings ("hi" in the classroom), and he responded correctly to one untrained context ("hi" in the hallway). In contrast, in the MET condition Rafael independently gave correct responses to two training contexts (“What is your full name?” in the classroom and “Where do you live?” in the cafeteria), and he responded correctly to two untrained contexts (“What is your full name? in the cafeteria and “Who can I call?” in the cafeteria). Lucas responded correctly to the context trained during SET, but he did not respond correctly to untrained contexts. However, he responded correctly in two of the trained contexts using MET to produce self-identification and responded correctly to all six of the untrained contexts (“Where do you live?” in the student center hallway and annex hallway, “Are you lost?” in the cafeteria and annex hallway, and “Do you need help?” in the student center hallway, cafeteria, and annex hallway). Felipe performed correctly in the training context to respond to name while watching the iPad in the cafeteria (SET). He responded correctly to another untrained context (respond to name while watching iPad at the library; SET). For respond to stop, Felipe maintained the skill in one of the training contexts (“Hold on” in the hallway; MET) and responded correctly to three untrained contexts (“Felipe” in the cafeteria, “stop” in the hallway, and “Felipe” in the hallway; MET).

Both methods, SET and MET, were effective in teaching relevant functional skills to the participants. For the three participants, generalization to untrained contexts occurred in more contexts when we taught using MET instead of SET (see Figure 4). Rafael responded correctly for one untrained context when taught with SET and two untrained contexts with MET. Lucas did not respond correctly in any untrained contexts in the SET condition, but he did respond correctly in six untrained contexts in the MET condition. Felipe responded correctly in one untrained context when taught using SET, and he responded correctly to three untrained contexts in the MET condition. For all the participants, MET resulted in generalization.

Independent responding was established after different amounts of exposure to the training phase for each participant. Rafael did not have any independent responses in the training phase with MET as all responses in this condition during training were either prompted or incorrect. He experienced the generalization phase after one context condition reached mastery criteria. Rafael started emitting independent responses after five sessions in the SET condition. Lucas began emitting independent responses after three sessions of MET and five sessions of SET. Felipe's performance was variable. After eight sessions during SET, he started emitting independent responses for a couple of sessions; followed by 21 sessions without any independent responses. In the MET condition, Felipe began to emit independent responses after 15 sessions, and he had variable performance until mastering one context.

The relatively liberal criterion of training until one skill in one context reaches a mastery criterion could have influenced the degree to which generalization was seen. Rafael did not have any independent responses in the MET training condition, but during training, he was responding correctly with the prompt. It was somewhat unexpected, then, that during the probes that followed the training phase Rafael responded correctly in two of the three training contexts and in two untrained contexts. It is possible that if the prompting level had been reduced more rapidly, Rafael would have had more independent correct responses during training. For Lucas, one context reached mastery the criterion in both conditions. Even though only one of three contexts reached mastery criterion in the MET condition, Lucas responded correctly to all untrained contexts. Lucas made errors in one of the trained contexts during the post-training probes. For Felipe, the mastery criterion was not met with SET. However, he correctly responded to the trained context during probes, and he responded correctly to one untrained context. Although Felipe mastered one of the three contexts in MET, he did not correctly respond to it during generalization. He achieved a high score in another context trained even though he did not master this context during training phase, and he responded correctly to two untrained contexts. Extraneous variables could have impacted generalization for Felipe. Felipe was diagnosed with a seizure disorder during his participation in this research. During the generalization phase Felipe´s performance may have been negatively impacted by symptoms related to the diagnosis.

Study 1 had some limitations such as (a) we applied a relatively liberal mastery criterion, (b) we did not balance the assignment between tasks and conditions, and (c) we did only one trial for each generalization session. In Study 2, these limitations were addressed by (a) yoking the number of trials to equate the exposure to training across the two conditions; (b) extending training phases until all the training context combinations reached mastery criteria prior to starting the generalization phase; (c) attempting to balance the assignment between task and condition across participants; and (d) adding another trial to generalization sessions, so that each session consists of two trials. Moreover, the procedures were replicated across and within participants, and maintenance sessions were added throughout the generalization phase in case occasional contact with reinforcement would allow for generalization to be seen across untrained contexts. In addition, stakeholders' judgment of the social validity of the goals, procedures, and outcomes was assessed.

Study 2

Method

Participants

Participants were three children diagnosed with ASD. Participants were recruited based on caregivers' indications that they could benefit from learning functional skills.

Beatriz was a 4-year-old girl who attended a behaviorally based special education day program. She communicated through signs and an AAC device. Her leisure activities consisted of watching videos on her tablet, playing on the computer, bouncing on a ball, playing on the playground, and watching a DVD at the gym. Challenging and interfering behavior targeted in her behavior intervention plan included self-injurious behavior (finger biting), bolting away from teacher/parents, dropping to the floor, tantrums, motor stereotypy (hand flapping and jumping), and vocal stereotypy (making noises).

Lara was a 3-year-old girl who received early intensive behavior analytic services at home before this study. She communicated through vocal verbal responses (three-to four-word utterances). Lara attended a regular pre-school classroom. Her leisure activities consisted of playing with toys with someone, reading books, drawing on paper, playing catch in the backyard, and playing with her iPad tablet computer. Challenging and interfering behavior included tantrums, perseveration (repeatedly talking about a specific event or object), and motor stereotypy (hand flapping and jumping).

Júlia was an 8-year-old girl who attended a public school and received ABA treatment in a specialized clinic. She communicated through some single-word utterances and an AAC device. Her leisure activities consisted of watching videos on her tablet, playing with squish toys, running, and playing with letters/numbers. Challenging and interfering behavior included crying, bolting away from teacher/parents, dropping to the floor, tantrums, motor stereotypy (hand flapping and jumping), and vocal stereotypy (yelling, making noises).

Settings and Materials

Sessions were conducted in multiple familiar settings and varied according to the skill taught. For Beatriz, sessions were conducted at various locations at the school: her classroom, her desk, the hallway, the gym, the computer room, and the playground. For Lara and Júlia, sessions were conducted at various locations at their house: living room, dining room, kitchen, and porch.

Items used for reinforcement were identified by each participants' caregiver as preferred items. Participants were given a choice of toys or small snacks presented visually prior to the beginning of each session. All sessions were recorded using video recording equipment and data were collected using a pen and paper. Other materials varied according to the task and are described in the specific task descriptions below.

Measurement and Response Definitions

Measurement was conducted in the same way as in Study 1. For waiting, the duration for which the participant waited for an item was recorded in seconds. Independent correct responses were defined individually for each task and each participant, and these definitions can be found in the Targeted Skills section below.

Experimental Design

The experimental design was the same as in Study 1, except that there were two comparisons between SET and MET for each participant, allowing for a within-subjects evaluation of replication. One challenge with parallel-treatments design is to ensure that the responses assigned to each treatment are of similar difficulty. Gast and Wolery (1988) suggested counterbalancing tasks in order to address this concern. We varied the assignments between task and condition across participants. For example, if the first participant learned object imitation using SET, then, if one of the other participants was learning object imitation, we used MET to teach the skill.

Procedure

For each participant, four target tasks were selected for training. The skills were selected based on participants' IEP objectives, parents' requests, and skill repertoire. Two skills were assigned to each teaching method. In Comparison 1, the two procedures were implemented in parallel and alternating sessions according to an ABAB sequence. The procedures were then repeated in Comparison 2 with two additional skills. The number of sessions for SET and MET was yoked. Targeted Skills are listed in Table 2, along with the stimulus variant types that defined the instructional matrix.

Baseline. The procedures for baseline were the same as in Study 1. For every skill other than waiting, each baseline session consisted of two trials for each context. For waiting there was one trial per context.

Training. Training sessions were identical to baseline, except that most-to-least physical or verbal prompting and differential reinforcement were used. Each session consisted of five trials. In the case of the waiting task, each session consisted of only one trial. For waiting, a shaping procedure was used to increase the participant's target time by 10% after meeting the criterion in one session. If the participant did not meet the criterion in two consecutive sessions, the target time was decreased to the previous duration. Each participant had a different target time as we considered each participant's age and baseline performance. Beatriz was taught to wait for 60 s for a preferred item in one context (e.g., sitting in the chair at the table at her school's desk; SET). Lara was taught to wait for a preferred item for 40 s in three contexts (sitting in the chair at the table in the dining room, sitting down on the floor in the living room, and standing in the kitchen; MET). Júlia was taught to wait for 90 s for a preferred item in one context (e.g., sitting down in the chair at the table in the dining room; SET). The mastery criterion for waiting was the final target time selected for each participant for two consecutive sessions in each of the training contexts.

When the participant responded correctly, the chosen item and verbal praise were delivered. The mastery criterion was six consecutive sessions with at least four out of five trials with correct independent responses. Like in Study 1, in the MET condition, training in the three training contexts was intermixed in an ABCABC sequence. Training continued until all the three training targets met mastery criteria (six sessions in total and two sessions for each training context). For all tasks except waiting, when the participants responded incorrectly, materials and attention were removed for 3 to 5s, and the trial was represented with the most restrictive prompt, and this response was not reinforced. If the participants made two consecutive errors or three errors in one session, the session was discontinued, and the next session started with a more restrictive prompt. The criterion to decrease the prompt hierarchy was two sessions with at least four of five trials correct (prompted or unprompted).

Generalization. Like in Study 1, probes were conducted in eight untrained contexts in the SET condition, and in six untrained contexts in the MET condition. These sessions were conducted using the same procedure as in baseline, however, reinforcement was delivered at the end of the session as a reward for participation rather than after gross motor imitation. Maintenance probe sessions in which reinforcement was delivered following correct responding in each of the trained contexts were interspersed with generalization probes to maintain the likelihood of responding on unreinforced generalization probe trials. For Beatriz (Comparisons 1 and 2), and Lara and Júlia (Comparison 1), probes were conducted according to a ratio of three probe sessions to one maintenance session. In these cases, only two of the three training contexts were programmed for maintenance probes in the MET condition. For Lara and Júlia (Comparison 2), the ratio of the probe to maintenance sessions was changed from 3:1 to 2:1 so that each of the trained contexts was evaluated in maintenance probes (see Table 3). Generalization probes were judged on the same basis as in the training session for each of the tasks except for the waiting task. For waiting, the generalization criterion was a duration of at least 90% of the target time.

Targeted Skills

Each of the six skills taught in Study 2 is described below, including any materials that were specific to that task. Table 2 provides details regarding the specific stimulus variants that defined each of the nine contexts defined in the instructional matrix.

Produce Self-Identification (Beatriz, SET; Júlia, MET). The stimulus variants for this task were a question from an adult and location. This is the same as it was for Rafael and Lucas (Study 1), with the following exceptions. The target response for producing self-identification was to hold one’s wrist with the personal identification bracelet out between themselves and the experimenter (point or not to the bracelet) following the instruction. Beatriz was taught to show the bracelet to a second experimenter while walking in the hallway when the experimenter asked the question ("What is your name?"; SET). Júlia was taught to show the bracelet and point to it when the experimenter asked one of the three questions (e.g., “What is your name?” in the living room, “Are you lost?” in the dining room, and “Do you need help?” in the kitchen; MET). One context was taught in each session, and the contexts were alternated in every session. A second experimenter was not available at Júlia's home, so the experimenter alternated the questions with some known one-step directions to make the questions less predictable.

Object Imitation (Lara, SET; Beatriz and Julia, MET). The stimulus variants for this task were action and object. At the start of each trial, the experimenter displayed the objects specific for each session, waited for attending, and paired the instruction “do this” with a specific action with the object. They responded by copying the action made by the experimenter with a specific object. For Beatriz, the three training contexts alternated from session to session (e.g., shake a playdough ball, build with cups, and spin a block; MET). For Lara, a two-step imitation was taught in one context (e.g., put a bubble container upside down and tap the bottom; SET). For Júlia, three contexts were taught (e.g., shake a bubble container, build with bowls, and push a piece of a cake; MET).

Wait for the Preferred Item (Beatriz and Júlia, SET; Lara, MET). The stimulus variants for this task were position and location. A vinyl disc, approximately 0.3 m in diameter was used in the contexts that included the standing position. At the start of each trial, the participant was positioned as specified by the instructional matrix to sit or stand depending on the location. The experimenter waited for attending, presented a preferred item, and asked them if they wanted the item. If they said yes, the experimenter presented the instruction “wait.” During all trials, the experimenter modeled the correct position and was within approximately 1 m of the participant. Correct responses in the three targeted positions were as follows: (a) While sitting down at the chair by the table, or (b) sitting down at the chair, the participant’s buttocks remained in the chair, which remained stationary. Their hands could be in their lap, on the table, or by their sides; the participant was permitted to move their hands between these positions, but they could not engage in motor stereotypy. (c) while standing, the vinyl disc was on the floor, and the participant remained standing on the target with their hands at their sides; moving one foot was permitted. For Lara, correct responses while sitting on the floor were as follows. The participant’s buttocks remained on the floor, which remained stationary. Her hands could be in her lap, or by her sides; the participant was permitted to move her hands between these positions, but she could not engage in motor stereotypy. An incorrect response was recorded if the participant moved from the relevant position defined above.

Respond to Instruction to Stop (Beatriz, MET). The stimulus variants for this task were instruction and location. This is the same as it was for Felipe (Study 1). For Beatriz, the three training contexts were "stop” in the hallway, “hang on” in the gym, and “hold on” in the classroom.

Tact (Lara, SET, and MET). The stimulus variants for this task were action and person (characteristics) or facial expression (emotion) and person (occupation). At the start of each trial, Lara was with the experimenter sitting down at the chair by the table at home. The experimenter waited for attending and presented a picture specifically for each session. A pretraining session prior to baseline sessions was included in which Lara was taught to tact the individual words that she would later combine to describe the pictures. For example, a picture of a girl without doing any action or a community helper without any facial expression was presented, and the following question was asked, "who is it?". Pictures of older people doing the action for activities were presented, and the following question was asked, "what is he/she doing?". Pictures of older people or more than one person for emotions were presented, and the following question was asked, "how do they feel?".

For SET training, a picture of a person doing an action (a girl reading a book) was presented, and she learned to answer, “the girl is reading a book.” For MET training, pictures of a people with different occupations and facial expressions were presented, and she answered, “the teacher is happy,” the doctor is sad,” and “the pilot is angry.”

Respond to Greetings (Júlia, SET). The stimulus variants for this task were greeting by an adult and ongoing activity. This is the same as it was for Rafael in Study 1, with the following exceptions. At the start of each trial, Júlia was with the experimenter sitting down at the chair by the dining room table at home. The experimenter turned the iPad on, and the participant started watching one of her favorite shows. While Júlia was watching her iPad, the experimenter would present one of the greetings ("hi"). The experimenter made it unpredictable when the instruction was coming. Júlia responded by looking at the experimenter and saying "hi."

Social Validity

A social validity questionnaire assessed the acceptability of the goals, procedures, and outcomes of the MET sessions. A total of 21 stakeholders including Beatriz's mom, Lara's and Júlia's parents, teachers, and supervisors from Beatriz's school answered the questionnaire. The questionnaire consisted of 17 questions that stakeholders scored on a 7-point Likert-type scale, ranging from strongly disagree to strongly agree, or opened-ended questions. The questions were divided into (a) assessing the program goals, (b) stakeholders satisfaction with participant's performance before (baseline) and after the intervention (generalization phase), (c) whether the intervention could be carried out by participant's teachers, (d) whether they would recommend using the instructional matrix to others, (e) if there were any changes that could make the intervention more valuable or any concerns, and (f) which of the methods used was better for generalization. The social validity questionnaire can be found in Appendix A.

We gave the questionnaire to stakeholders when they were together in small groups and the experimenter followed a script during the meetings. "First, I would like to thank you for your participation. This first paper page tells you which skills were taught to your student/your child. Please go ahead and fill out the first two questions of this package (program goals) questions one and two." When the stakeholders finished answering both questions, the experimenter told them that they would watch a video projected on a screen or shown on a computer, and that after viewing the video, they should answer questions three and four. The experimenter randomly assigned the order of presentation of the pretest and posttest videos for each group of stakeholders. Each pretest (baseline) and posttest (generalization) video consisted of the first trial of each of the nine contexts taught using MET. We selected videos from the second comparison for each participant. Most of the videos were between 1 and 1.5 min in duration. The video that showed a participant engaging in the waiting task was longer: approximately 3 min for the pretest video and 8 min for the posttest video. After that, the stakeholders watched the other video and answered questions five and six. Following answering the questions, a debriefing moment took place in which the experimenter commented about the two videos watched by the stakeholders: “You watched two videos showing your student/your child performing a skill in two moments, the pretest and the posttest. The procedures used to teach the skill are called multiple exemplar training using the instructional matrix. Please refer to the sheet in front of you. Each square of the table represents one context (a combination of one instruction and one setting). The procedures consist of teaching a skill using three exemplars (three combinations of instructions and settings). After they learn the skill, tests were conducted in the other combinations that were not directly taught.” The full script used during the social validity meetings can be found in Appendix B. The stakeholders were asked to assess whether the instruction provided could be carried out by the participant´s teachers, whether the stakeholders would recommend MET using the instructional matrix to others, and if they had any other feedback to make the intervention more valuable or any concerns with the procedures (questions seven, eight, and nine). After that, the experimenter gave each stakeholder a sheet with Figure 8 with the data from SET and MET for only one of the participants. The experimenter described the axis and the legend and requested that the stakeholders answer questions 10-17.

Interobserver Agreement and Procedural Integrity

Interobserver agreement was calculated using the same parameters as Study 1 for at least 40% of sessions across participants in each phase of the study for SET and MET. Mean agreement across participants was 98% (range, 96%-100%).

Procedural integrity data were collected to ensure that the experimenter was correctly implementing the procedures throughout Study 2. A second experimenter evaluated videos from 40% of the trials across participants in each phase of the study. During baseline, training, and generalization sessions, data were collected on whether the experimenter presented (a) the correct stimulus variants for the task, (b) the correct prompt, and (c) the correct consequence. These data were summarized by calculating the total number of trials with correct responses performed by the experimenter during each session, dividing it by the total number of trials, then multiplying by 100 to obtain a percentage. The mean procedural integrity across all of the participants was 99% (range, 98%-100%).

Results and Discussion

Beatriz

During baseline, Beatriz did not emit any correct response for the four skills (Figure 5). In the training phase, Beatriz acquired the response within each of the training contexts. In the first comparison, Beatriz learned to produce her identification when a second adult asked, "What is your name?" in the hallway in 27 sessions (SET; Table 4). She learned to imitate the experimenter's model action (MET) in 40 sessions. Beatriz learned each action with one object in 13 sessions each. In the generalization phase, she did not have any correct independent response to produce self-identification in the untrained contexts, but she maintained the skill in two of three trained contexts (the third context was not evaluated). In contrast, Beatriz responded correctly to untrained contexts during the object imitation task (shake a block twice and shake a cup once), and she maintained the responses in the trained contexts evaluated.

In the second comparison, Beatriz learned to wait for a preferred item in her workspace while sitting down in the chair at the table in 30 sessions (SET; Table 4). She learned to respond to stop in the three training contexts (MET) in 28 sessions. Beatriz learned all of the three contexts in 9 sessions. During generalization, she responded correctly to one of the waiting untrained contexts (sitting in a chair in her cubby), and she maintained the skill in two of three trained contexts (the third context was not evaluated). Beatriz responded correctly and maintained the skill of responding to a request to stop in all the evaluated contexts (trained and untrained).

Lara

Before baseline, we conducted a pretraining session to evaluate whether Lara could tact all the components taught in combinations. She tacted all the stimuli (girl, boy, adult, reading, cooking, and running for SET; teacher, doctor, pilot, happy, sad, and angry for MET). During baseline, Lara did not emit any correct responses (Figure 6). In the first comparison, she learned to tact pictures with a girl reading a book in 12 sessions (SET). She learned to tact the three training pictures in 33 sessions (MET). The first two contexts using MET were learned in 8 sessions, and the last context (a pilot with an angry face) took 11 sessions to reach mastery. During the generalization phase, Lara tacted two untrained pictures (a boy reading a book and a boy running), and she maintained the trained context with SET. However, she did not tact any untrained context when we used MET. She maintained two of the trained contexts, and the third trained context was not evaluated.

In the second comparison, Lara learned to imitate the experimenter’s action with one object (turn the bubble container upside down and tap the bottom; SET) in 15 sessions. She learned to wait for a preferred item in three contexts (sitting in a chair at the table in the dining room, sitting on the floor in the living room, and standing in the kitchen) in 57 sessions (MET). Two contexts were learned in 18 sessions (sitting in a chair at the table in the dining room and sitting on the floor in the living room), and one context was learned in 19 sessions (standing in the kitchen). During the generalization phase, Lara imitated five new actions (turn a bowl upside down and tap on the bottom, turn a piece of cake upside down and tap the bottom, to put one bowl on top of another and shake them together, to put one bubble container on top of another and shake them together, and put a piece of a cake on top of another piece and shake them together), and she maintained the learned context with SET. Lara responded correctly and maintained the skill to wait for a preferred item in all the contexts (trained and untrained).

Júlia

During baseline, Júlia did not emit any correct response for the four skills (Figure 7). In the first comparison, she learned to wait for a preferred item for 90 seconds while sitting in a chair at the dining room table in 36 sessions (SET). Júlia learned to imitate the experimenter's actions with three contexts (shake a bubble container, build with bowls, and push forward a piece of cake) in 28 sessions. Each context was learned in 9 sessions. During generalization probes, she maintained the trained context and improved on waiting in all the contexts compared to baseline, but she did not reach the target time. For object imitation, Júlia responded correctly to all the untrained contexts, and she maintained two of the trained contexts, and the third trained context was not evaluated.

In the second comparison, Júlia learned to respond to a greeting while watching her iPad in 16 sessions (SET). She learned to produce her identification when an adult asked, "What is your name?" in the living room, "Are you lost?" in the dining room, and "Do you need help?" in the kitchen in 31 sessions (MET). One context was learned in 11 sessions ("What is your name?" in the living room), and the other two contexts were learned in 8 sessions each. During generalization, Júlia responded to the greeting "hi" while eating (one untrained context), and she maintained the context trained. She responded correctly and maintained the skill to produce self-identification in all the contexts (trained and untrained).

Social Validity

Social validity questionnaire responses showed that the stakeholders found the skills highly acceptable for all the participants (mean 6.9), and they highly agreed that generalization is essential for students with autism (mean 6.9; Table 5). Stakeholders were highly satisfied with the participants' performance after training (mean 6.5), with the amount of improvement observed on learning (mean 6.6), and with the amount of improvement observed in skill generalization (mean 6.7). Stakeholders also highly agreed that the instruction provided could be carried out by their teachers (mean 7.0), and they would highly recommend the use of the instruction (mean 7.0). Stakeholders strongly disagree (mean 1.8) that SET was better than MET when comparing the two methods, and strongly agreed (mean 6.5) that MET was better than SET based on graphs (see Table 5). Social validity open-ended answers are shown in Appendix C. Overall, the answers reflected that the stakeholders value generalization and that the instructional matrix approach could be a good way to teach skills in different locations (e.g.: "because it teaches the skill in multiple contexts without having to be so specific" (parent); "I think that if Julia can generalize in multiple locations, she will retain skills and knowledge in all settings" (parent).

Single Exemplar Training (SET) and Multiple Exemplar Training (MET) were effective methods to teach the targeted skills in the contexts in which they were trained for all the participants in Study 2. The functional skills learned by the participants are useful only to the extent that they occur in desired contexts, not only during training contexts. Results for two of three participants showed that generalization occurred in more contexts when we taught using MET instead of SET (see Figure 8). In the first comparison, Beatriz responded correctly to two contexts when taught with MET and zero context with SET. In the second comparison, she responded correctly to one context with SET while she responded correctly to six untrained contexts when we taught using MET. For Júlia, in the first comparison, she did not respond correctly to untrained contexts when taught with SET, however, she responded correctly to six untrained contexts using MET. In the second comparison, she responded correctly to one context with SET, and she responded correctly to six untrained contexts with MET.

Lara responded correctly to two untrained contexts when taught with SET, but she did not respond correctly in the generalization phase of the MET condition in the first comparison. In the second comparison, she responded correctly to five untrained contexts with SET, and she responded correctly to six untrained contexts with MET.

Taken together, these results support the argument that generalization needs to be actively programmed (Stokes and Baer, 1977). MET was effective in promoting generalization with all participants. There was only one comparison (Lara's first comparison) for which MET did not result in generalization. The target skill in both the MET and SET conditions for this comparison was tacting features of a person in a photo. For MET the task was to tact the emotion (facial expression) and the occupation, and for SET the task was to tact the action and characteristics of the person (i.e., boy/girl/adult). It may be important to note that this was the only comparison across both studies that required recombinative generalization. Tacting based on multiple features of a stimulus requires a different response in each context, and that is different from the other skills targeted in the study. The discrimination of both features of the stimuli could be challenging, and this could have had a negative impact on generalization. Interestingly, this participant did perform better in generalization tests with the SET condition than the MET condition in this comparison. Though we attempted to equate the difficulty of skills across comparisons, it is not possible to control completely for differences in difficulty, especially when using functional skills. There could have been a difference that led to better outcomes with SET. The number of contexts targeted for training in this case (three) may not have been sufficient for this task, and we may have seen better performance in the untrained contexts with additional exposure to MET. Previous studies have found that the number of sufficient exemplars may be greater when teaching more challenging skills. For example, in one matrix training study, Striefel et al. (1976) taught eight combinations of verbs and nouns. Goldstein (1983) identified the need to determine the minimal stimulus conditions necessary for generalization, and this remains an area for continued research.

To our knowledge, no previous studies have evaluated and compared the social validity of SET and MET. In this study, we inquired about the three components of social validity identified by Wolf (1978). We found that the skills taught were relevant to participants and stakeholders, and stakeholders viewed generalization as essential for students with autism. Stakeholders indicated that procedures used were appropriate, could be carried out by their teachers, and they would recommend the instructional matrix to others. Additionally, stakeholders indicated that they were satisfied with the amount of improvement in skill acquisition and skill generalization. Also, stakeholders agreed that MET was better than SET in producing generalization of skills.

General Discussion

This study extends the literature on generalization by evaluating a new strategy to program for generalization. A simplified approach to general case instruction was designed and implemented, an instructional matrix. In the current investigation, MET with the instructional matrix produced greater generalization than SET in eight out of the nine comparisons. The instructional matrix borrowed elements from both general case instruction and matrix training. From general case instruction, it borrowed the programming of common stimuli. From matrix training, it borrowed the organization of relevant stimuli into a table to determine combinations for teaching and generalization. In each axis of the matrix, there were variants of a single dimension of the context (e.g., a location or instruction).

Although several studies have demonstrated generalization following MET, published research reporting direct comparisons is limited. To our knowledge, only three studies compared the effects of SET and MET (Day & Horner, 1986; Horner & McDonald, 1982; Sprague & Horner, 1984). Taken together, the studies conducted by Horner and his colleagues have shown that MET produces greater generalization then SET, and this is especially the case when those exemplars are carefully selected to sample from the relevant features of the generalization contexts (Day & Horner, 1986; Horner & McDonald, 1982; Sprague & Horner, 1984). However, these studies share an experimental design element that poses a challenge for interpretation. In each case, an SET phase preceded an MET phase. The current study extended the literature by directly comparing the effects of SET and MET on the acquisition and generalization of functional skills using a parallel-treatments design. This experimental design allowed us to evaluate and identify the strengths of SET and MET separately without the threat of carryover effects of the independent variables. It would be interesting to compare in future studies the effects of instructional matrix training and general case instruction on generalization. In the current study, the skills taught had only one step, while studies on general case instruction usually teach multiple steps (e.g., Day & Horner, 1986; Horner & McDonald, 1982; Sprague & Horner, 1984). Hence, it would be also interesting to evaluate whether teaching multiple steps or one step makes a difference in relation to the number of exemplars trained and tested. Both approaches have been effective in producing generalization, but it would be useful to identify the relative strengths of the two methods. The instructional matrix approach may be faster and easier to implement when you teach only one-step skills, and in conditions where this is true, it may be a better choice for teachers looking to promote generalization of skills.

Horner and McDonald (1982) mentioned that usually one representative exemplar (SET) is selected and taught to mastery when teaching community living or vocational skills, and argued that this is a relatively ineffective and inefficient procedure for teaching generalized repertoires. One possible advantage of SET, however, is that it typically would require less planning and take less time to teach using a single exemplar than to teach using multiple exemplars. We were able to evaluate this possibility directly in Study 2, where training continued in both conditions until mastery criteria were met in both. This design feature allowed us to compare the number of sessions required to meet mastery criteria with one vs. three exemplars. For most of the comparisons (four out of six), participants reached mastery criterion faster in SET. In three of these four comparisons, generalization was greater during MET.

Our decision in Study 2 to continue teaching in both conditions until mastery criteria were met in both introduced another potentially impactful difference from Study 1. In some cases, this design feature resulted in overtraining. It may be important that the extent to which overtraining occurred was different across the two conditions. Specifically, as criteria were met in SET prior to MET in most comparisons, more overtraining occurred in SET than MET. It is possible that overtraining could have affected generalization. Although we did not have a control in place to allow us to evaluate effects of overtraining on generalization, we note greater generalization was seen in MET both in Study 1, where terminated training when mastery criteria were met, and in Study 2, where overtraining did occur. As a result, it does not appear that in this study, the difference in the about of overtraining across Studies 1 and 2 had a substantial impact on our results. In future studies, potential effects of overtraining could be evaluated by systematically programming overtraining sessions in both conditions, or by programming two SET conditions: in one, the teaching phase could end when mastery criteria are met, and in the other, the teaching phase continues until the mastery criteria are met in the MET condition.

The current investigation extends previous research evaluating effects of multiple exemplars on generalization by (a) providing a direct comparison of SET and MET, using two different experimental designs, and (b) evaluating social validity of the procedures and outcomes of MET. The social validity results should be interpreted with caution, given that only pretest and posttest videos of MET were presented before stakeholders were provided with the questionnaire, but some of the items on the questionnaire were unrelated to the video. It's possible that showing this video before asking these questions could have influenced the stakeholders' responses. Those items asked the respondent to visually inspect a graph. It is possible that, in contrast with professionals, some parents may have limited experience with interpreting graphs of their child's progress. A better method for evaluating the social validity of the two teaching procedures with stakeholders could be to show videos from both conditions.

In the past few decades there has been a call for increasing and improving the measurement of social validity (Schwartz & Baer, 1991). In future studies researchers may choose to develop further the methods for evaluating social validity. It would be useful to know if teachers prefer to use the instructional matrix in comparison to SET or other approaches to MET. Researchers could evaluate this aspect of social validity with teachers after they implement the procedures. Another way to develop further social validity methods and possibly control for potential confounds is to separate the evaluation of procedures and outcomes. In future study, the social validity questionnaire could be administered to two groups, one in which people who are not familiar with the outcomes judge the procedures, and one in which people who are not familiar with the procedures inspect the data. Having two groups separately evaluate specific aspects of the procedures and the data could provide better answers about the acceptability, satisfaction of the procedures and outcomes while minimizing the impact of possible bias. Finally, in their comments, stakeholders pointed out that they would like to see the skills in other settings and with unknown or new people. In this vein, future research may consider including various community settings such as parks or stores, at least during the pretest and generalization phases. This would allow researchers to ensure that participants learn to engage in skills in the relevant environments.

In conclusion, programming for generalization has been of concern to applied behavior analysts for many years, and there is a need for continued exploration of methods to program for generalization. In the current investigation, we evaluated the instructional matrix strategy and in eight out of the nine comparisons, the MET produced greater generalization across participants and with multiple skills. These outcomes support further investigation of the efficacy and social validity of the instructional matrix strategy of programming for generalization. A better understanding of the relative effects of SET and MET on generalization may guide clinicians and teachers in designing instruction that will produce greater generalization. This study goes beyond past research by directly comparing the effects of SET and MET using a parallel treatments design. This design was selected because it controls better the potential confound of sequence effects. This is, to the best of our knowledge, the first study to include any comparison of measures of social validity between SET and MET procedures. Stakeholders reported that they were satisfied with the outcomes and approved of the instructional matrix procedures. Given the positive outcomes we obtained with the instructional matrix, future research is warranted. This research may continue to investigate the effects of this approach on generalization, and the social acceptability of the procedures in comparison to other approaches.

References

Alessi, G. (1987). Generative strategies and teaching for generalization. *The Analysis of Verbal Behavior*, *5*, 15–27. http://doi.org/10.1007/BF03392816

Baer, D. M., Wolf, M. M., & Risley, T. (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis*, *1*, 91–97. http://doi.org/10.1901/jaba.1987.20-313

Bord, S., Sidener, T.M., Reeve, K.F., & Sidener, D.W. (2017). Teaching On-Task Rollerblading and Ice-Skating to a Child with Autism. *Behavior Analysis in Practice, 10*, 178-182. http://doi.org/10.1007/s40617-016-0150-z

Curiel, E.S., Curiel, H., & Li, A. (2020). Generative time telling in adults with disabilities: A matrix training approach. *Behavioral Interventions*, 295-305. http://doi.org/ DOI: 10.1002/bin.1714

Dass, T.K., Kisamore, A.N., Vladescu, J.C., Reeve, K.F., Reeve, S.A., & Taylor-Santa, C. (2018). Teaching children with autism spectrum disorder to tact olfactory stimuli. *Journal of Applied Behavior Analysis*, *51*, 538-552. http://doi.org/10.1002/jaba.470

Day, H.M., & Horner, R.H. (1986). Response variation and the generalization of a dressing skill: comparison of single instance and general case instruction. *Applied Research in Mental Retardation*, *7*, 189-202. http://doi.org/10.1016/0270-3092(86)90005-6

Dickson, C. A., MacDonald, R. P. F., Mansfield, R., Guilhardi, P., Johnson, C., & Ahearn, W.H. (2014). Social validation of the New England Center for Children-Core Skills Assessment. *Journal of Autism and Developmental Disorders*, *44*(1). http://doi.org/10.1007/s10803-013-1852-5

Erhard, P., Falcomata, T.S., Oshinski, M., & Sekula, A. (2022). The effects of muiltiple-exemplar training on generalization of social skills with adolescentes and young adults with autism: a systematic review. *Review Journal of Autism and*

Developmental Disorders. https://doi.org/10.1007/s40489-022-00328-4

Garcia-Albea, E., Reeve, S.A., Reeve, K.F., & Brothers, K.J. (2014). Using audio script fading and multiple-exemplar training to increase vocal interactions in children with autism. *Journal of Applied Behavior Analys*is, *47*, 325-43. http://doi.org/10.1002/jaba.125

Gast, D. L., & Wolery, M. (1988). Parallel-treatments design: a nested single-subject design for comparing instructional procedures. *Education and Treatment of Children*, *11*, 270–285. http://www.jstor.org/stable/42899073

Goldstein, H. (1983a). Recombinative Generalization: relationships between environmental conditions and the linguistic repertoires of language learners. *Analysis and Intervention in Developmental Disablities*, *3*, 279–293.http://doi.org/10.1016/0270-4684(83)90002-2

Goldstein, H. (1983b). Training generative repertoires within agent-action-object miniature linguistic systems with children. *Journal of Speech and Hearing Research*, *26*, 76–89. http://doi.org/10.1044/jshr.2601.76

Goldstein, H., & Mousens, L. (1989). Generalized language learning by children with severe mental retardation: effects of peer's expressive modeling. *Journal of Applied Behavior Analysis*, *22*, 245–259. http://doi.org/10.1901/jaba.1989.22-245

Gomes, S.R., Reeve, S.A., Brothers, K.J., Reeve, K.F., & Sidener, T.M. (2019). Establishing a Generalized Repertoire of Initiating Bids for Joint Attention in Children with Autism. *Behavior Modification, 44*, 394-428.http://doi.org/10.1177/0145445518822499

Groskreutz, N.C., Groskreutz, M.P., Bloom, S.E., & Slocum, T.A. (2014). Generalization of negatively reinforced mands in children with autism. *Journal of Applied Behavior Analysis*, *47*, 560-579. http://doi.org/10.1002/jaba.151

Holth, P. (2017). Multiple exemplar training: some strengths and limitations. *The Behavior Analyst*, *40*, 225–241. http://doi.org/10.1007/s40614-017-0083-z

Horner, R.H., Jones, D. N., & Williams, J. A. (1985). A functional approach to teaching generalized street crossing. *JASH*, *10*, 71–78. http://doi.org/10.1177/154079698501000201

Horner, R. H., & McDonald, R. S. (1982). Comparison of single instance and general case instruction in teaching a generalized vocational skill. *TASH*, *8*, 7–20. http://doi.org/10.1177/154079698200700302

Langton, E.K., Miguel, C.F., Diaz, J.E., Cordeiro, M.C. & Heinicke, M.R. (2020). Na evaluation of matrix training to teach college students piano notes and rhythms. *Journal of Applied Behavior Analysis*, *53(3)*, 1466-1484. http://doi.org/ 10.1002/jaba.690

LeBlanc, L.A., Sellers, T.P., & Ala´i, S. (2020). Structured Problem Solving Skills. In L.A. LeBlanc, T.P. Sellers, & S. Ala´i (Eds.)*, Building and sustaining meaningful and effective relationships as a supervisor and mentor* (pp.116-133). Sloan Publishing.

Milata, E.M., Reeve, S.A., Reeve, K.F., & Dickson, C.A. (2020). A blueprint for general-case procedures illustrated by teaching adolescents with autism spectrum disorder to use a chip-debit card. *Behavioral Interventions*,1-26. http://doi.org/ DOI: 10.1002/bin.1719

O'Neill, R. E., Faulkner, C., & Horner, R. H. (2000). The effects of general case training of manding responses on children with severe disabilities. *Journal of Developmental and Physical Disabilities*, *12*(1), 43–60.https://doi.org/10.1023/A:1009456210739

Oppenheimer, M., Saunders, R. R., & Spradlin, J. E. (1993). Investigating the generality of the delayed-prompt effect. *Research in developmental disabilities*, *14*(6), 425–444. https://doi.org/10.1016/0891-4222(93)90036-j

Risley, R., & Cuvo, A. J. (1980). Training Mentally Retarded Adults to Make Emergency

Telephone Calls. *Behavior Modification*, 4(4), 513–525. https://doi.org/10.1177/014544558044006

Rosales-Ruiz, J., & Baer, D. M. (1997). Behavioral cusps: a developmental and pragmatic concept for behavior analysis. *Journal of Applied Behavior Analysis*, *30*, 533–44. http://doi.org/10.1901/jaba.1997.30-533

Schwartz, I.S., & Baer, D.M. (1991). Social validity assessments: is current practice state of the art?. *Journal of Applied Behavior Analysis*, *24,* 189-204. http://doi.org/10.1901/jaba.1991.24-189

Sprague, J.R., & Horner, R. H. (1984). The effects of single instance, multiple instances, and general case training on generalized vending machine use by moderately and severely handicapped students. *Journal of Applied Behavior Analysis*,*17*, 271-278. http://doi.org/10.1901/jaba.1984.17-273

Stokes, T. F., & Baer, D. M. (1977). An implicit technology of generalization. *Journal of Applied Behavior Analysis*, *10*, 349–367. http://doi.org/10.1901/jaba.1977.10-349

Stokes, T. F., Baer, D. M., & Jackson, R. L. (1974). Programming the generalization of a greeting response in four retarded children. *Journal of Applied Behavior Analysis*, *7*, 599–610. http://doi.org/10.1901/jaba.1974.7-599

Stokes, T. F. & Osnes, P. G. (1989). An operant pursuit of generalization. *Behavior Therapy*, *20*, 337–355. http://doi.org/10.1016/S0005-7894(89)80054-1

Striefel, S., Wetherby, B., & Karlan, G. (1976). Establishing generalized verb-noun instruction-following skills in retarded children. *Journal of Experimental Child Psychology*, *22*, 247–260. http://doi.org/10.1016/0022-0965(76)90005-9

Wolf, M.M. (1978). Social validity: the case for subjective measurement or how applied behavior analysis is finding its heart. *Journal of Applied Behavior Analysis*, *11*, 203-214. http://dx.doi.org/10.19012jaba.1978.11-203

Table 1.

A Generic Instructional Matrix.

Variation Type 2	Variation Type 1		
	A	B	C
1	A1	B1	C1
2	A2	B2	C2
3	A3	B3	C3

Note. Variations Type 1 and 2 defined the contexts for evaluated generalization. These were different across tasks. For example, for the task, Respond to Stop, Variation Type 1 was the instruction, and Variation Type 2 was the location. In the SET condition, context A1 was trained and the other contexts were evaluated for generalization. In the MET condition, contexts A1, B2, and C3 were trained, and the other contexts were evaluated for generalization.

Table 2.

Overview of Tasks, Method, Participants, and Variation Type for Studies 1and 2

Task	SET: Participant (Study)	MET: Participant (Study)	Variation Type 1	Variation Type 2
Wait	Beatriz (2), Júlia (2)	Lara (2)	Position	Location
Self-Identification	Beatriz (2)	Rafael (1), Lucas (1), Júlia (2)	Question from Adult	Location
Respond to "Stop"		Felipe (1), Beatriz (2)	Instructions	Location
Produce a Bill	Lucas (1)		Question from Cashier	Location (Store)
Respond to Name	Felipe (1)		Ongoing Activity	Location
Respond to Greetings	Rafael (1)		Greeting by Adult	Location
	Júlia (2)		Greeting by Adult	Ongoing Activity
Tact	Lara (2)		Action	Person (characteristics)
		Lara (2)	Facial Expression (Emotion)	Person (occupation)
Object Imitation	Lara (2)	Beatriz (2), Júlia (2)	Action	Object

Note. SET = Single Exemplar Training; MET= Multiple Exemplar Training.

Table 3.

Example of a Sequence of Generalization Probes and Maintenance Sessions for SET and MET for Study 2.

	SET	MET
Session	Context	Context
1	“Are you lost?”/ Hallway	“Are you lost?”/ Hallway
2	“Do you need help?”/Hallway	“Do you need help?”/Hallway
3	Maintenance 1(“What is your name?”/Hallway)	Maintenance 1(“What is your name?”/Hallway)
4	“What is your name? ”/ Playground	“What is your name? ”/ Playground
5	“Are you lost?”/Playground	“Do you need help?”/Playground
6	Maintenance 1 (“What is your name?”/Hallway)	Maintenance 2 (“Are you lost?”/Playground)
7	“Do you need help?”/Playground	“What is your name?”/ Gym
8	“What is your name?”/Gym	“Are you lost?”/ Gym
9	Maintenance 1 (“What is your name? /Hallway)	Maintenance 3 (“Do you need help? /Gym)
10	“Are you lost?”/ Gym	“Are you lost?”/ Hallway
11	“Do you need help?”/Gym	“Do you need help?”/Hallway

Table 4.

Number of Sessions to Mastery and Total Training Sessions for Each Participant in Study 2.

Participant	Condition	Skill	Sessions to Mastery	Total Training Sessions
Beatriz	Comparison 1			
	SET	Produce ID	27	39
	MET	Object Imitation	40	41
	Comparison 2			
	SET	Wait	30	30
	MET	Respond to Stop	28	32
Lara	Comparison 1			
	SET	Tact	12	33
	MET	Tact	33	36
	Comparison 2			
	SET	Object Imitation	15	56
	MET	Wait	57	59
Júlia	Comparison 1			
	SET	Wait	36	36
	MET	Object Imitation	28	35
	Comparison 2			
	SET	Respond to Greetings	16	31
	MET	Produce ID	31	35

Note. SET = Single Exemplar Training; MET = Multiple Exemplar Training.

Table 5.

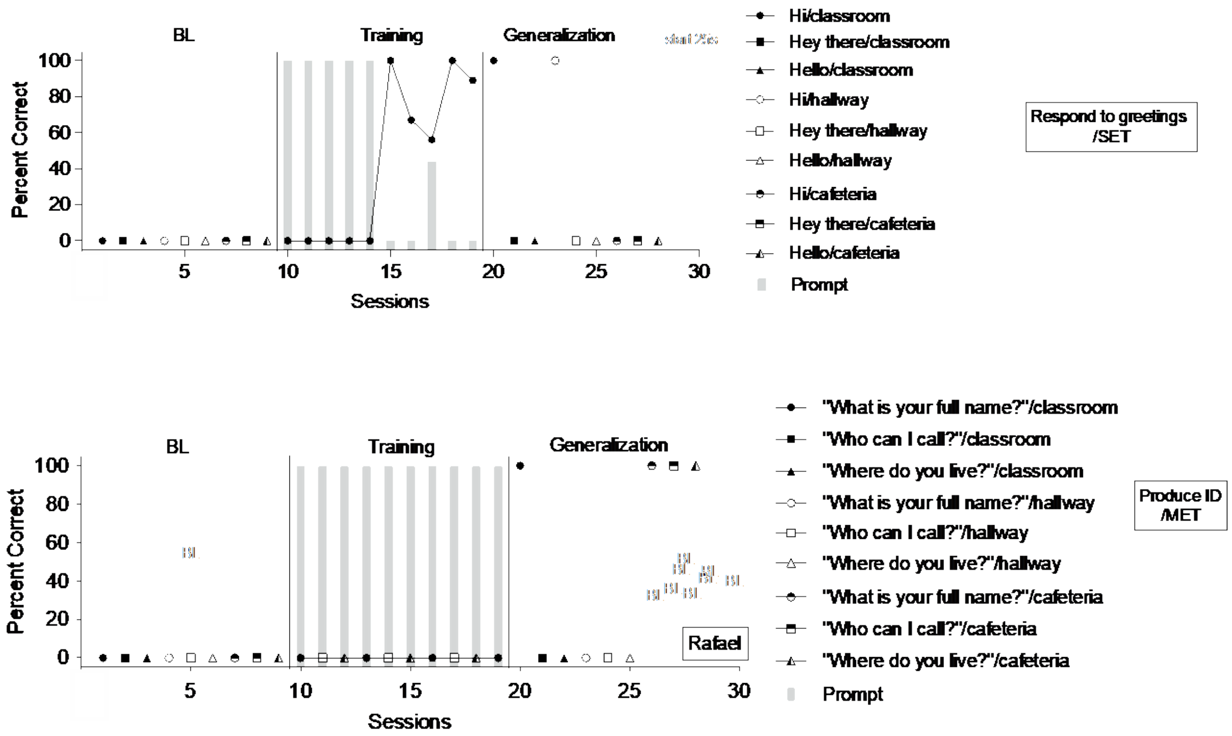
Social Validity Questionnaire Results for Study 2

	Ratings (Mean)			
Questions	Beatriz	Lara	Júlia	Total
1.Acceptability and relevance of the skills	7.0	6.7	6.9	6.9
2.Generalization essential for students with autism	7.0	6.7	6.9	6.9
3.Satisfaction with performance (pretest)	1.0	1.6	2.4	1.7
4.Satisfaction with performance (posttest)	6.7	6.6	6.1	6.5
5. Instruction could be carried out by their teachers	7.0	7.0	7.0	7.0
6. Recommend Multiple Exemplar Training using the instructional matrix to others	7.0	7.0	7.0	7.0
7. Satisfied with the amount of improvement in skill acquisition	6.9	6.9	6.1	6.6
8. SET was better than MET	1.0	2.9	1.6	1.8
9. MET was better than SET	7.0	5.9	6.7	6.5
10. Satisfied with the amount of improvement observed in skill generalization	6.7	6.9	6.4	6.7

*Note.*1 = not acceptable, not satisfied, strongly disagree; 7 = highly acceptably, highly satisfied, highly agree

Figure 1

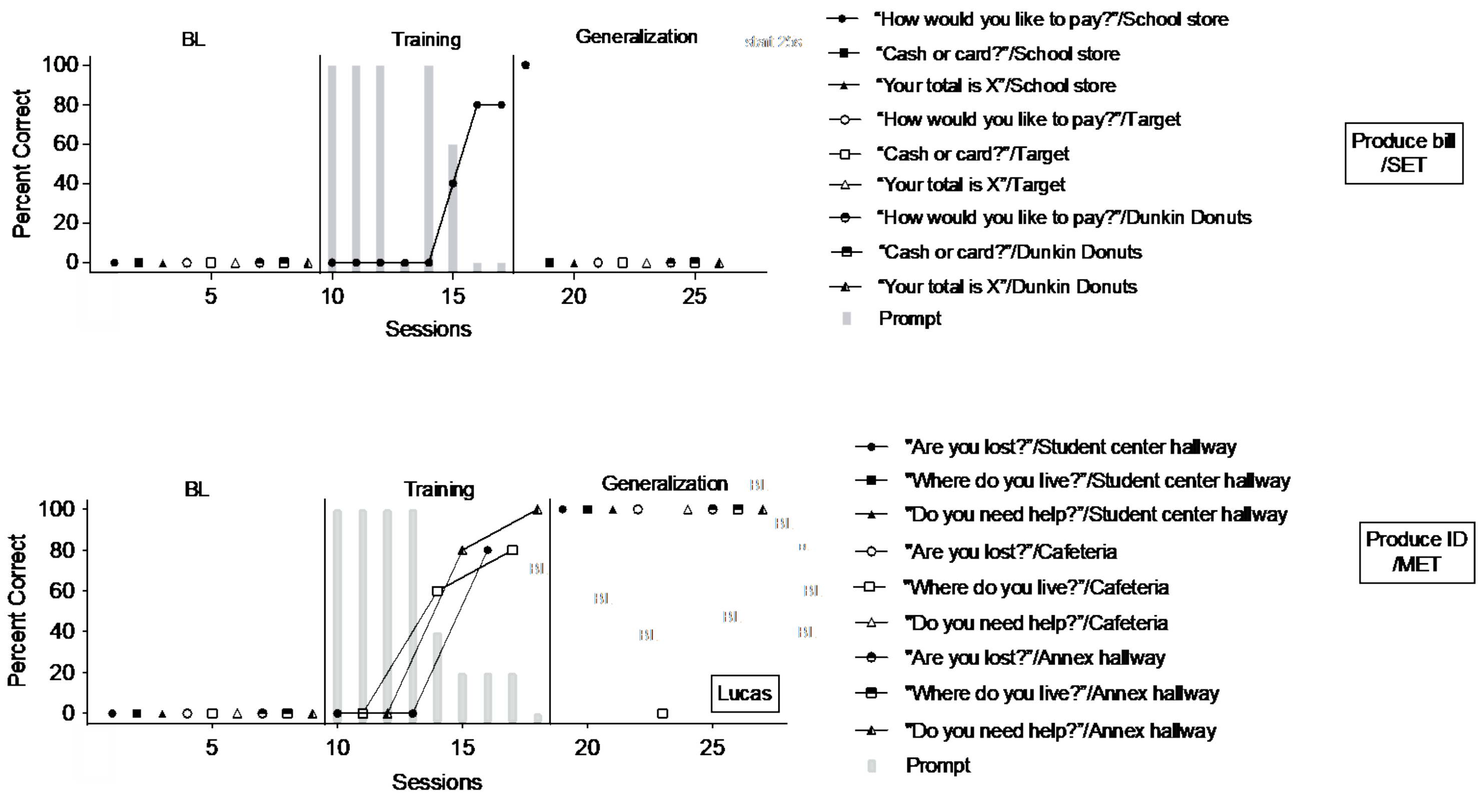
Percentage of Correct Responses Across Phases and Each Condition for Rafael in Study 1



Note. SET = Single Exemplar Training; MET = Multiple Exemplar Training. The top panel displays respond to greetings using SET and the bottom panel displays produce self-identification using MET. The grey bars show the percentage of correct prompted responses.

Figure 2

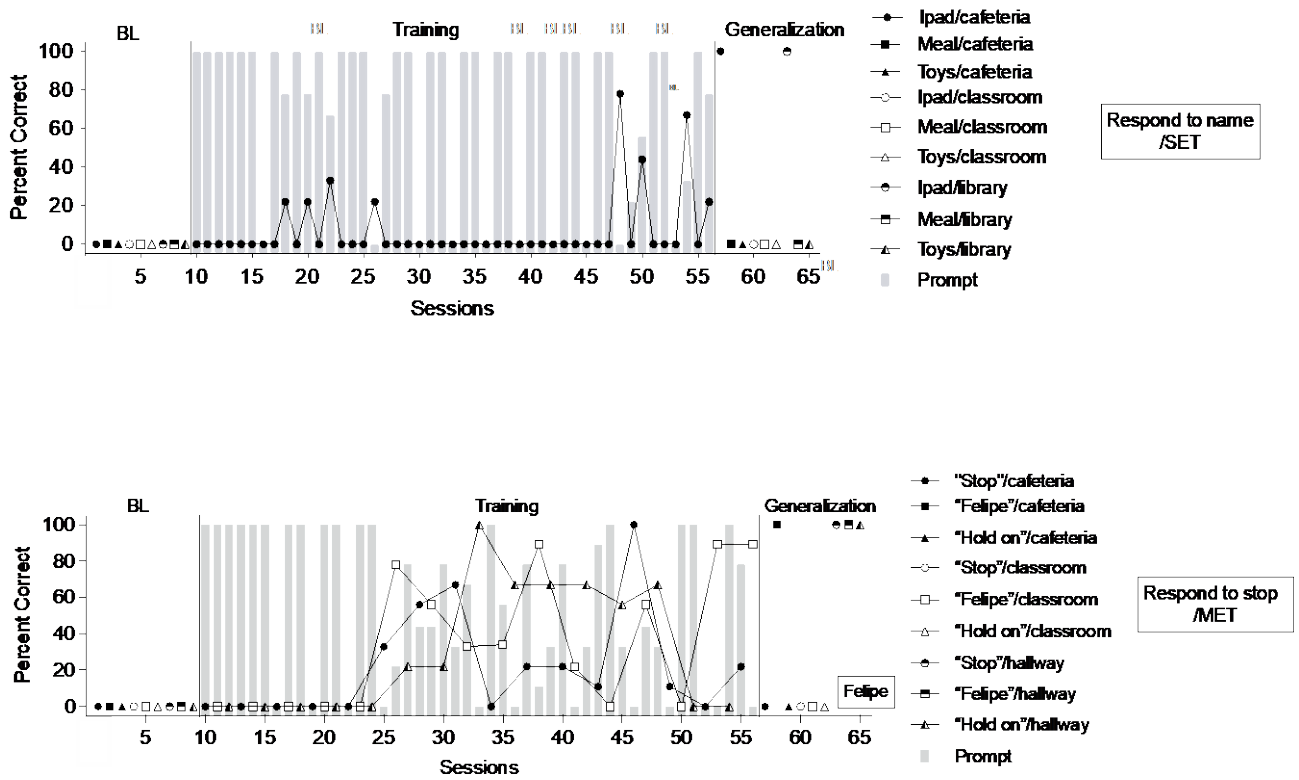
Percentage of Correct Responses Across Phases and Each Condition for Lucas l in Study 1



Note. SET = Single Exemplar Training; MET = Multiple Exemplar Training. The top panel displays produce a bill using SET and the bottom panel displays produce self-identification using MET. The grey bars show the percentage of correct prompted responses.

Figure 3

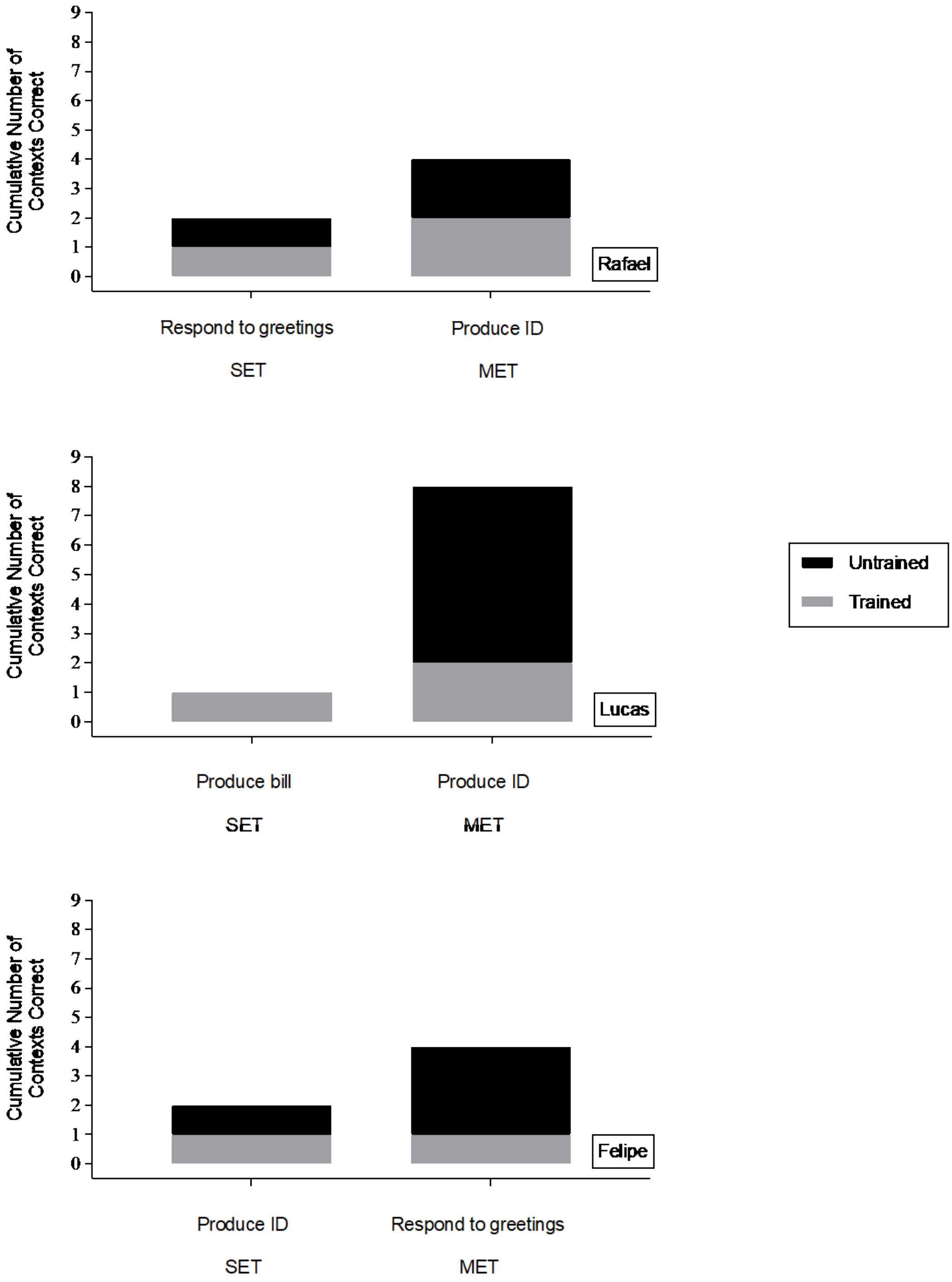
Percentage of Correct Responses Across Phases and Each Condition for Felipe in Study 1



Note. SET = Single Exemplar Training; MET = Multiple Exemplar Training. The top panel displays respond to name using SET and the bottom panel displays respond to stop using MET. The grey bars show the percentage of correct prompted responses.

Figure 4

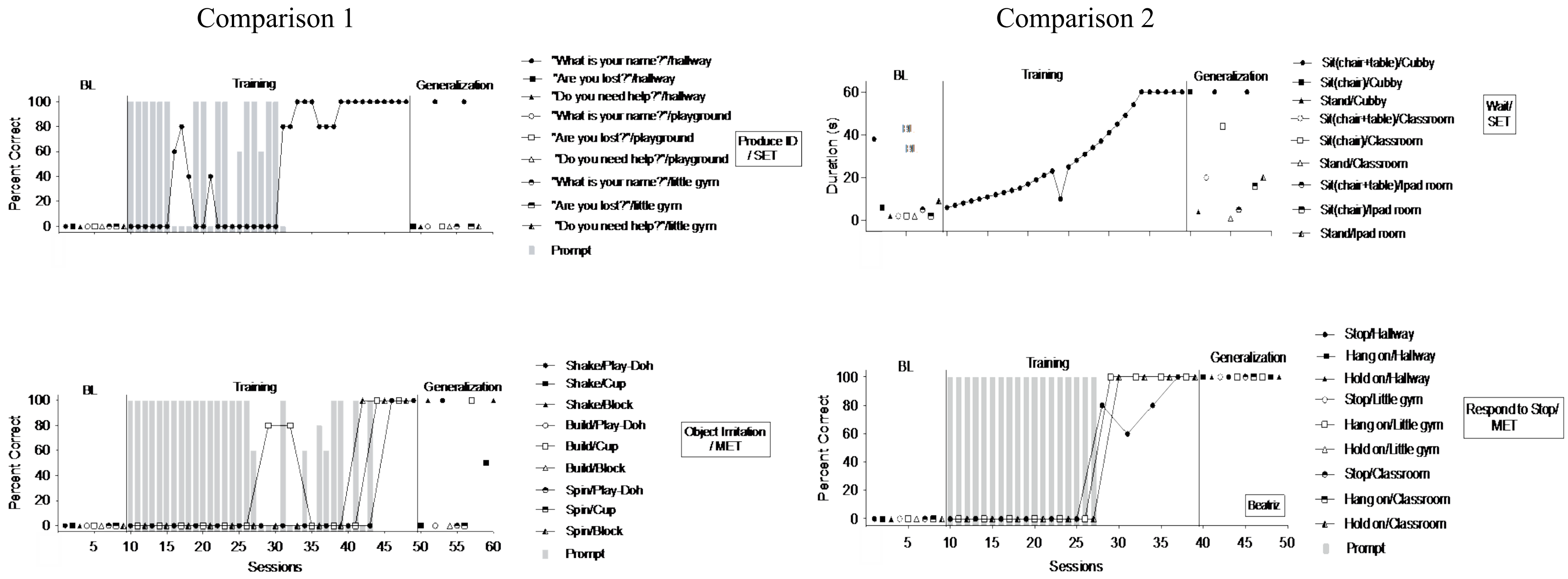
Cumulative Number of Trained and Untrained Contexts with Correct Responding in Study 1



Note. The grey portion shows the number of contexts trained; the black portion shows the number of untrained contexts during the generalization phase. The top panel displays Rafael's results, the middle panel displays Lucas' results, and the bottom panel displays Felipe's results.

Figure 5

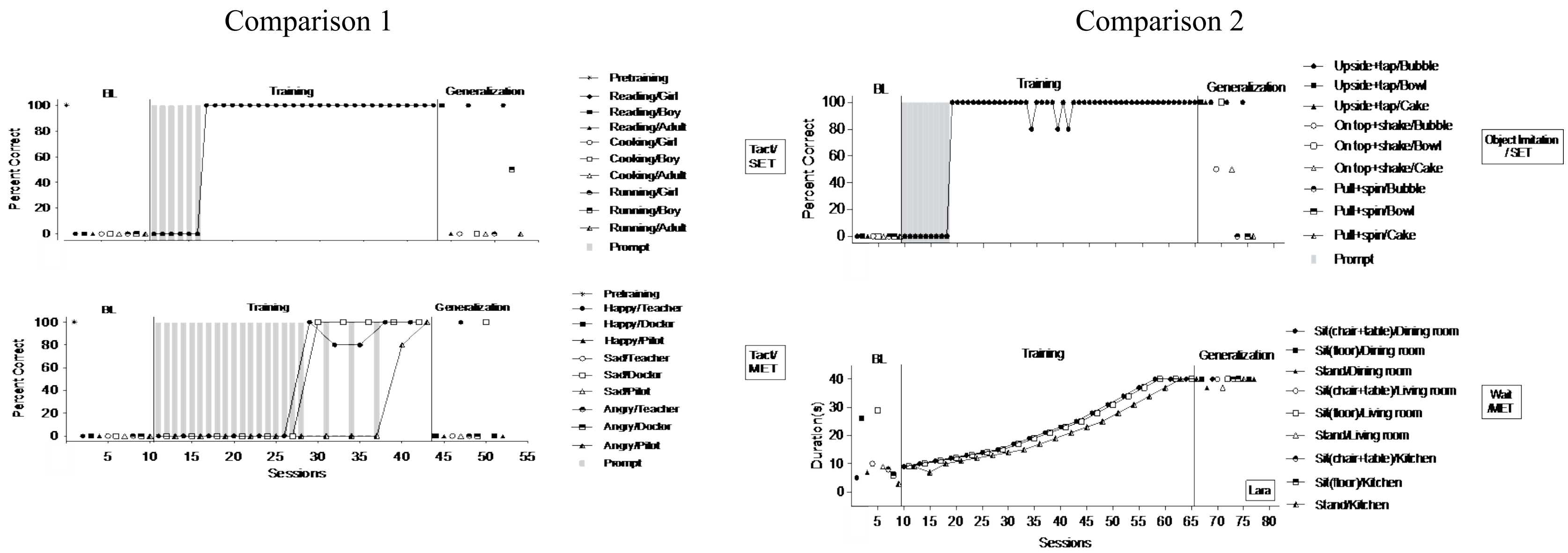
Percentage of Correct Responses Across Phases and Each Condition for Beatriz in Study 2



Note. SET = Single Exemplar Training; MET = Multiple Exemplar Training. The left two panels to display the first comparison (SET the first graph and MET the second graph), and the right two display the second comparison (SET the first graph and MET the second graph). The grey bars show the percentage of correct prompted responses.

Figure 6

Percentage of Correct Responses Across Phases and Each Condition for Lara in Study 2

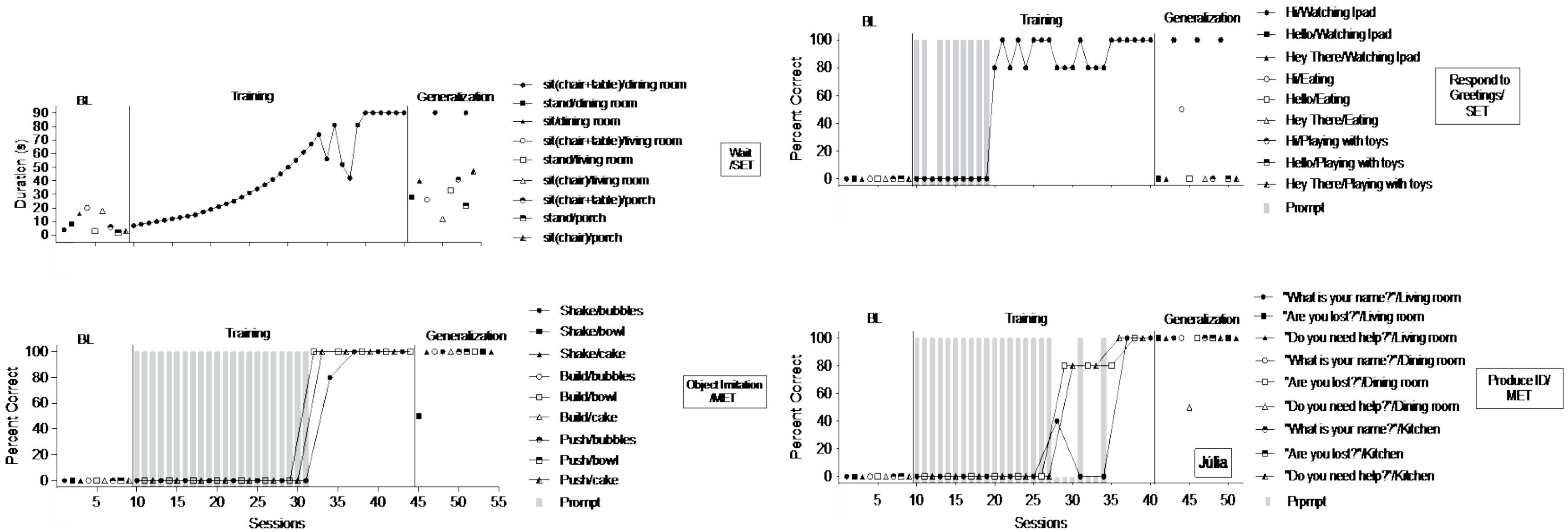


Note. SET = Single Exemplar Training; MET = Multiple Exemplar Training. The left two panels display the first comparison (SET the first graph and MET the second graph), and the right two display the second comparison (SET the first graph and MET the second graph). The grey bars show the percentage of correct prompted responses.

Figure 7

Percentage of Correct Responses Across Phases and Each Condition for Júlia in Study 2

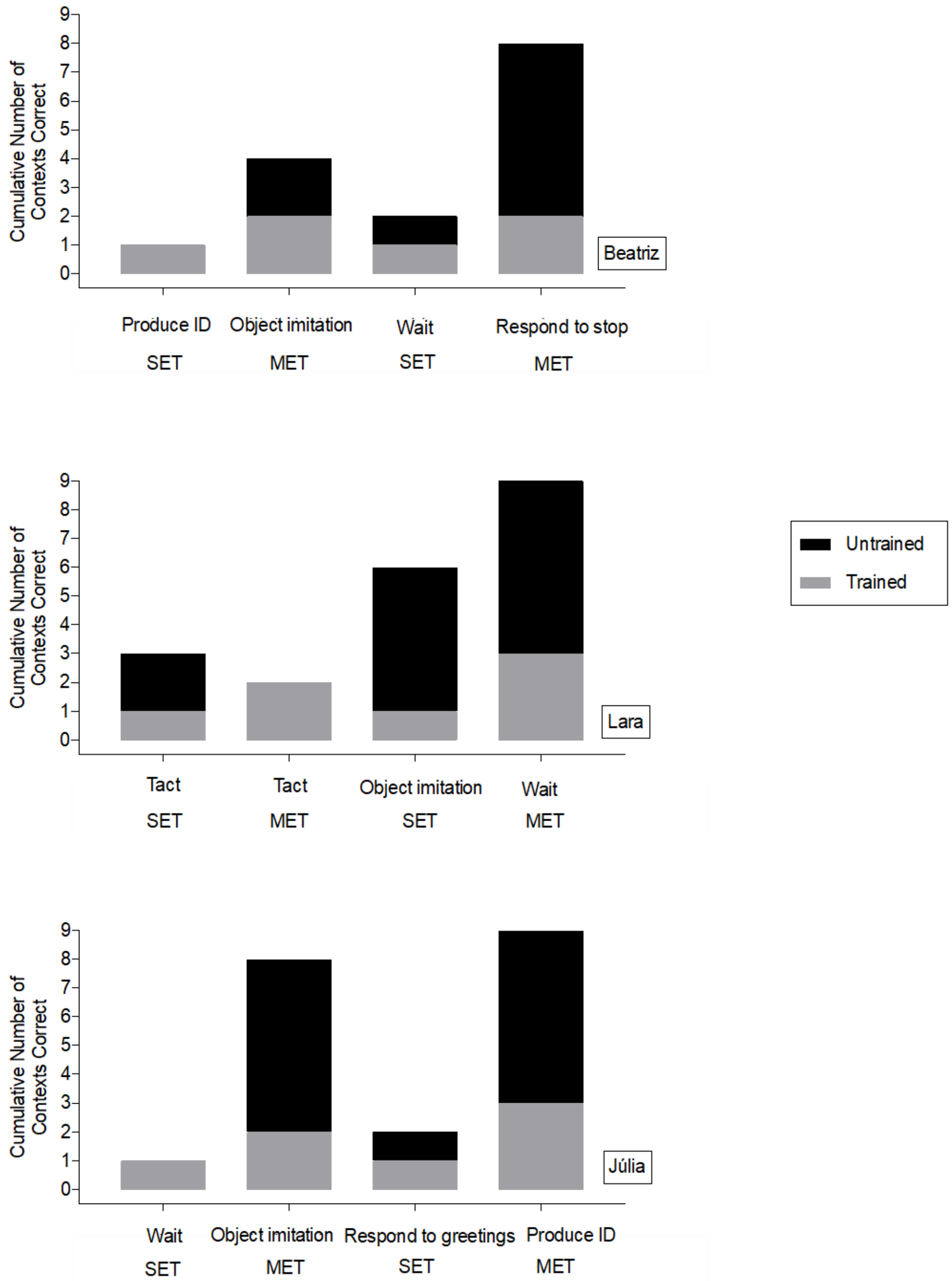
Comparison 1 Comparison 2



Note. SET = Single Exemplar Training; MET = Multiple Exemplar Training. The left two panels display the first comparison (SET the first graph and MET the second graph), and the right two display the second comparison (SET the first graph and MET the second graph). The grey bars show the percentage of correct prompted responses.

Figure 8

Cumulative Number of Trained and Untrained Contexts with Correct Responding in Study 2



Note. The grey portion shows the number of contexts trained; the black portion shows the number of untrained contexts during the generalization phase. The top panel displays Beatriz's results, the middle panel displays Lara's results, and the bottom panel displays Júlia's results.

Appendix A

Social Validity Questionnaire

- Your opinion is very important for my study!
- Only answer the questions associated with the video and graph you viewed.

Program Goals

1. Do you agree that the skills taught are relevant and acceptable for them?

1	2	3	4	5	6	7
Strongly Disagree			No Opinion			Strongly Agree

2. Would you consider generalization to be essential for students with autism?

1	2	3	4	5	6	7
Strongly Disagree			No Opinion			Strongly Agree

After Viewing ***Video 1***

3. Based on what you just saw, are you satisfied with the student's performance?

1	2	3	4	5	6	7
No			Unsure			Yes

4. Why or why not?

After Viewing ***Video 2***

5. Based on what you just saw, are you satisfied with the student's performance?

1	2	3	4	5	6	7
No			Unsure			Yes

6. Why or why not?

After Debriefing

7. Would you agree the instruction provided to the student could be carried out by their teachers?

1	2	3	4	5	6	7
No			Unsure			Yes

8. Would you recommend Multiple Exemplar Training using the instructional matrix to others?

1	2	3	4	5	6	7
No			Unsure			Yes

9. Is there anything else you would like to add that may make the intervention more valuable? Are there any changes you would like to see or any concerns you would like addressed?

After Viewing ***the graph***

10. Based on what you just saw, are you satisfied with the amount of improvement in skill acquisition observed?

1	2	3	4	5	6	7
No			Unsure			Yes

11. Why or why not?

12. Based on what you just saw, do you think Single Exemplar Training was better than Multiple Exemplar Training for generalization?

1	2	3	4	5	6	7
No			Unsure			Yes

13. Why or why not?

14. Based on what you just saw, do you think Multiple Exemplar Training was better than Single Exemplar Training for generalization?

1	2	3	4	5	6	7
No			Unsure			Yes

15. Why or why not?

16. Are you satisfied with the amount of improvement observed in skill generalization?

1	2	3	4	5	6	7
No			Unsure			Yes

17. Why or why not?

Appendix B

Social Validity Script

- ❖ First, I would like to thank you for your participation.
- ❖ This first sheet tells you which skills were taught to your student/your child.

1. Please go ahead and fill out the first two questions of this package (program goals) questions 1& 2.
2. We will watch a video, and after viewing the video, you will answer questions 3 & 4.
3. Now, we will watch another video, and you will answer questions 5 & 6.

4. **Debriefing**

- ➢ You watched two videos showing your student/your child performing a skill in two moments, the pretest and the posttest.
- ➢ The procedures used to teach the skill are called multiple exemplar training using the instructional matrix.
- ➢ Please refer to the sheet in front of you. Each square of the table represents one context (a combination of one instruction and one setting). The procedures consist of teaching a skill using three exemplars (three combinations of instructions and settings). After they learn the skill, tests were conducted in the other combinations that were not directly taught.

5. **Graph**

- ➢ You received a bar graph with the results for each skill learned. On the y-axis is showing the number of combinations (contexts) with correct responding. On the x-axis is showing each skill and the method used to teach the skill (SET – Single Exemplar Training and MET – Multiple Exemplar Training). On the right, there is a legend. The grey portion represents the contexts trained (e.g., one for SET and three for MET). The black portion represents the untrained contexts in which your student/child responded correctly without directly being taught. Please answer questions 10 – 17.

- ❖ When you are done answering the questions, please put your questionnaire in this envelope.
- ❖ Thank you!

Appendix C

Social Validity Questionnaire Open-Ended Answers

After Viewing Pretest

Based on what you just saw, why or why not are you satisfied with the student's performance?	
Parent	1. "Because she did not stop" 2. "I am not sure given this scenario that she could tell someone her proper name and ask for assistance when needed. "I don't like that she said 'Debbie' at first when ask her name'?"
Supervisor	1. "The student did not stop when cued. She paused a few times, but in a parking lot, etc. that would be dangerous" 2. "Does not demonstrate response" 3. "Questions were not answered and vocal verbal behavior not very understandable to novel listener" 4. "Not responding to the experimenter's question or in any way identifying herself, but her eye contact was okay and she did emit same response. Needs to learn a functional response" 5. "The student did not accurately respond to any of the target questions"
Teacher	1. "Student was unable to respond to any SD in any environment" 2. "She did not respond to teacher in any environment" 3. "Student was not responding to any verbal cues" 4. "The student does not stop when directed to do so" 5. "They did not wait in any of the trials" 6. "She does not wait as specified on the dot or otherwise tries to engage in food stealing instead of waiting" 7. "She did not stay in the designated area, attempted to grab the items, and engaged in more interfering behavior" 8. "She occasionally was able to wait longer times, but mostly unable to wait or progress in amount of time waiting" 9. "She could not consistently wait for 40 seconds" 10. "She needs practice waiting! But I know that she will be able to learn" 11. "The student did not wait when instructed to do so" 12. "The student was just repeating vocalizations from the teacher" 13. "Cannot provide self ID if lost"

After Viewing Posttest

Based on what you just saw, why or why not are you satisfied with the student's performance?	
Parent	1. "Because she listened and did what was expected of her" 2. "It shows that Julia has the essential skills to tell someone her name and advise them she is lost. Hugely important from her safety"
Supervisor	1. "She stopped when cue plus looked up to therapist" 2. "The SD seems generally strong" 3. "You get more bang for your buck (MET). SET didn´t give anything for free." 4. "Allows time to teach other things" 5. "Higher number of contexts correct and higher percentage of contexts acquired "for free" without explicit training" 6. "Following training in only three contexts, both skills taught using MET generalized to six untrained contexts". 7. "Student reliably pointed to identification" 8. "Response to the experimenter's question was different from the other more stereotypic responses she was emitting, but if the experimenter hadn't preceded 'Are you lost?' With some other simple instructions, I'm not confident that she would have responded to the question." 9. "I would like to see the student respond to other adults. However, the student only pointed to the self-ID stimulus when asked 1 of the 3 target questions and not when asked to engage in other behavior"
Teacher	1. "Student was able to respond to the SD in any environment!" 2. "She responded each time in different areas of the school" 3. "The student was responding to stop in different environments and with different SDs" 4. "The student consistently stops when directed to do so" 5. "The student waited appropriately in all (but one) trials. She waited so nicely! Very impressed!" 6. "She waited with no challenging behavior in all videos" 7. "She was able to stay the majority of trials, progressed nicely" 8. Most of the times in the post-test she was able to wait for 40sec" 9. "Such progress! Nice job!" 10. "The student waited when instructed to do so" 11. "She was efficiently responding according to self-identification training" 12. "Pointed to ID and can receive help if lost"

After Debriefing

Is there anything else you would like to add that may make the intervention more valuable? Are there any changes you would like to see or any concerns you would like addressed?	
Parent	1. "What would happen under a scenario when Julia did not have her bracelet?" 2. "Julia's stereotypy is problematic"
Supervisor	1. "No, very efficient!" 2. "Would be a good thing to test in trials to mastery with future students (blocks of trials and interspersed trials)" 3. "I would suggest a verbal script for each skill. This means the teacher training the verbal interaction plus the target skill" 4. "I would love for her to raise arm up each time. A novel citizen may not recognize pointing to a wrist as a declaring of information" 5. "Either build into this training matrix or as a 'next level' skill set, I would like to see more naturalistic settings (such as in a store, mall, playground) and more natural ways of approaching the child (such as a stranger walking up and asking her, or even trying to engage in small talk first, realizing she has limited communication, and then asking her the three trained questions)." 6. "I would like to see some way assess and teach the skill with respect to novel unfamiliar adults. I would like to see settings outside of the home as the home itself may function as an SD"
Teacher	1. "I would like to see generalization of these skills in other contexts! Also, across teachers! This could be taught across caregivers and out in the community" 2. "I would like to see the intervention done across additional setting (community, park, stores, etc)" 3. "I would like to see the skills generalize to other teachers. Maybe future studies can test to see if the skills generalize across teachers, parents, or other caregivers. Also adding one of the context at home" 4. "N/A" 5. "I don't think any changes are needed. Awesome intervention!" 6. Generalization across people would also beneficial, I bet her parents/guardians would love to have this skill generalized" 7. "Possibly seeing if the skill of waiting occurs with other people (parents, teachers)" 8. "Teach more verbal instructions beyond 'wait'" 9. "It would be interesting to see if this generalization worked across people (other teachers/parents) as well as these contexts. If so this would make the skill very valuable for the participant" 10. "You did a wonderful job!" 11. "N/A" 12. "Test/probe in community setting"

After Viewing the graph

Based on what you just saw, why or why are you not satisfied with the amount of improvement in skill acquisition observed?	
Parent	1. "Because she learned the skills and performed them well" 2. "She seemed to make progress"
Supervisor	1. "Very impressed with MET!" 2. "Generalized skills without training; acquires more contexts" 3. "Yes for object imitation and produce ID, no for waiting and respond to greetings. There is definitely room for improvement!" 4. "Significant increase in skills taught using MET; little to no increase in skills taught using SET (waiting and respond to greetings did not improve)"
Teacher	1. "The student learned the skill and was responding independently" 2. "They acquired skills in all different skills types" 3. "It seems like the graph shows great correct responding" 4. "Both training methods led to an increase in the number of correct responses in untrained contexts" 5. "She made more progress and performed without much challenge or distraction" 6. "Yes with wait definitely, with tact not as much" 7. "The student showed huge improvement" 8. "It is a progress, but only 2/4 skill were generalized" 9. "Acquired all skills for producing ID"
Based on what you just saw, why or why not do you think Single Exemplar Training was better than Multiple Exemplar Training for generalization?	
Parent	1. "Because the instructions would have to be so specific" 2. "multiple is better because it is across different locations"
Supervisor	1. "SET didn´t give anything for free" 2. "Didn't acquired more contexts" 3. "Both skills that used SET only had one context mastered. And only 1 out of 2 had any generalization to untrained contexts, with a mostly 1 context 4. "Little to no generalization to untrained contexts for the SET skills" 5. "Single exemplar training did not result in generalization to untrained contexts"
Teacher	1. "I think it is better to have multiple to be able to learn across people and environments" 2. "Less trained and untrained responses occurred in SET than MET" 3. "SeT showed correct responding for untrained skills in object imitation, but more acquisition occurred for MET in waiting" 4. "Need more exposure in environments" 5. "Multiple worked very well from what I saw" 6. "It is a skill the student should have generalized"

	7. "It shows less skills mastered under trained and untrained conditions" 8. "Lower skill acquisition"
Based on what you just saw, why or why not do you think Multiple Exemplar Training was better than Single Exemplar Training for generalization?	
Parent	1. "Because it teaches the skill in multiple contexts without having to be so specific" 2. "I think that if Julia can generalize in multiple locations she will retain skills and knowledge in all settings" 3. "I like that multiple practice the skill in different locations"
Supervisor	1. "You get more bang for your buck (MET)" 2. "Allows time to teaching other things" 3. "Trained skills were mastered and untrained. Overall better performance seen with MET" 4. "Higher number of contexts correct and higher percentage of contexts acquired 'for free' without explicit training" 5. "Follows training in only 3 contexts. Both skills taught using MET generalized to 6 untrained contexts"
Teacher	1. "Better to have multiple exemplar for generalizing across people and environment" 2. "More responding in MET sessions" 3. "More skills were acquires using MET for waiting" 4. "More exposure" 5. "It generalized really well" 6. "It is a skill the student should have in many different areas" 7. "More context were mastered on training and more context were generalized without training" 8. "Trained skills were mastered and untrained. Overall better performance seen with MET" 9. "Significantly higher skill acquisition"
Why or why not are you satisfied with the amount of improvement observed in skill generalization?	
Parent	1. "Because it worked well" 2. "Because she answered with her name"
Supervisor	1. "This is great for this student! Would have loved more progress with SET" 2. "Allows time to teaching other things" 3. "I love to see both skills using MET had all contexts mastered. This is an important part of training – can also be the most difficult to produce" 4. "yes, but I would continue working towards more ecologically-relevant generalization as well" 5. "Object imitation and produce ID generalized to 6 contexts, and required no training – this is a significant number!"
Teacher	1. "Student was able to generalize the skill in every environment" 2. "Was able to have skills across different environments"

	3. “The student emitted the correct response in different environments" 4. “Generalization occurred for all skills except tacting with MET" 5. “Yes, ¾ skills generalized to untrained settings" 6. “She improved nicely" 7. “There was a clear difference that she generalized and could complete the skill in the posttest. 8. “Yes! The student had huge improvement" 9. It is definitely a progress, but waiting and respond to greeting would be exposed to MET" 10. “The skill generalized to all generalization setting"