EXPERIMENTAL ANALYSIS OF HUMAN BEHAVIOR BULLETIN 2025, 37, 4-22

VARYING INTER-STIMULUS AND INTER-TRIAL INTERVALS DURING A TRANSLATIONAL EXTENSION OF STIMULUS-STIMULUS PAIRING

Patricia Eberhardt[1], April Michele Williams[1], Stephanie P. da Silva[2]

[1] Rollins College
[2] Columbus State University

Stimulus-stimulus pairing (SSP) is a respondent conditioning procedure used to elicit vocalizations in children with language delays. Unfortunately, many of the studies published fail to maximize its success, most likely because specific parameters that generate better outcomes are unknown. One factor that may impact SSP efficacy is the relative duration of the inter-trial interval (ITI) and/or the inter-stimulus interval (ISI). To investigate this, the present study varied the ITI from 20 to 60 s while also varying the ISI to keep it proportional to the ITI. Nine typically developing children, aged 15 to 21 months, were randomly assigned to one of three experimental groups. The study alternated between pairing (sound model, followed by food delivery) and control (no programmed pairing between the sound model and food delivery) trials for all subjects in a trace conditioning procedure. The pairing condition led to a slightly higher percentage of trials with vocalizations across all subjects. Subjects in the 20-s ITI group had the highest rates of vocalizations, followed by those in the 60-s ITI group and the 30-s ITI group. However, subjects in the 60-second ITI group were most likely to approach the apparatus, as would be expected per the delay reduction hypothesis. In conclusion, the effectiveness of SSP depends, in part, on the relative temporal contiguity of events with longer ITIs (e.g., 60 s) producing more approaches to the sound model (NS/CS), although the shortest ITI group had the highest rates of vocalizations in the current study.

Keywords: Autoshaping; classical conditioning; respondent conditioning: stimulus-stimulus pairing; translational research

Applied behavior analysts have long been aware of the necessity for early verbal behavior acquisition and shaping for the clients we serve. Otherwise, those individuals can fail to effectively access (or actively avoid) potentially desirable items, edibles, and social interactions as they age (Barbera, 2007; Koegel et al., 1992; Rosales-Ruiz & Baer, 1997; Sundberg et al., 1996). Along those lines, significant language impairment at age 4 is positively correlated with similar delays in academics, social skills, and adaptive functioning and negatively correlated with the occurrence of challenging behavior (Sundberg, 2008; Sundberg & Partington, 1998; Tomblin et al., 1997; Petursdottir & Mellor, 2017). It is clear, then, the important role language acquisition plays in enriching one's life.

As Skinner (1957) first discussed, functional language typically begins developing in infancy through a combination of social and automatic reinforcement (Shillingsburg et al., 2015; Skinner, 1957; Smith et al., 1996; Sundberg & Michael, 2011). That is, social interactions with caregivers when the infant vocalizes increase the production of topographically similar sounds via respondent conditioning (Gros-Louis et al., 2006; Wu & Gros-Louis, 2017) and come under operant control soon after. The infant's own vocalizations also serve as automatic reinforcers, strengthening future reproduction of sounds even without others present to provide social or other classes of reinforcers (Skinner, 1957; Smith et al., 1996; Vaughan & Michael, 1982). Repeated vocalizations ultimately increase the variety of sound production via automatic reinforcement; in fact, the closer the sound produced matches the caregiver sounds, the more effective the infant-produced sounds function as automatic reinforcers (Smith et al., 1996).

As noted in that typically occurring scenario, infant vocalizations first are elicited via respondent conditioning before being shaped into sounds and words which make up the native language. For example, sounds emitted by caregivers (sound model; neutral stimulus; NS) become conditioned stimuli (CS) by repeated pairing of sound model with unconditioned stimuli (US; e.g., food, warmth, rocking, diaper changes, etc.; Shillingsburg et al., 2015; Sundberg & Michael, 2011; Sundberg et al., 1996). Eventually, the conditioned response (CR; sound production by the infant) occurs in the presence of the sounds produced by the caregivers. Some infants produce very few, if any, initial vocalizations through this common caregiving procedure. For example, nearly 40% of individuals with a diagnosis of autism fail to develop vocal communication skills and remain nonvocal for the rest of their lives (National Autism Association, 2018). Thus, interventions aimed at teaching early language and communication skills are essential to prevent cumulative adverse consequences or skill limitations for such children (Sundberg & Partington, 1998; Sundberg, 2008). *Stimulus-stimulus pairing* (SSP) is a procedure used by applied behavior analysts to do exactly this. Specifically, to facilitate the occurrence and rate of vocalizations. Like what was described previously with caregivers, the SSP process involves pairing an NS (the sound model, typically, produced by an adult) with a preferred item, such as food, tickles, smiles, or a favorite toy. Such "pairing" continues until the NS becomes a CS, indicated by increased vocalizations in the presence of the sound model (CS; Smith et al., 1996).

Note that, as this is a respondent rather than operant procedure, the individual is not required to make any vocal response (Sundberg et al., 1996). Therefore, SSP can help develop vocalizations in language-delayed children without the use of operant procedures that require sounds already existing in the repertoire (e.g., echoic training; Yoon & Bennett, 2000). Specifically, echoic training (i.e., verbal imitation intervention protocols) implemented with nearly non-verbal children has the potential for creating avoidance and escape behavior due to infrequently emitting the target operant and rarely contacts reinforcers as a result. One issue hindering the advancement of the use of SSP procedures for eliciting vocalizations in clinical situations appears to be these procedures have not yet been standardized (da Silva & Williams, 2020; Madden et al., 2023). Documented implementation of SSP has yielded inconsistent outcomes, sometimes leading to increased vocalizations (e.g., Barry et al., 2019; Esch et al., 2009; Rader et al., 2014) and sometimes not (Carroll & Klatt, 2008; Miguel et al., 2002; Stock et al., 2008; Yoon & Feliciano, 2007). Some reports include insufficient details to allow development of standard procedures.

Researchers are attempting to pinpoint the variables believed to play a role in the efficacy of SSP for increasing vocalizations in nearly (or completely) non-vocal children. For instance, the number of sound presentations, type of conditioning used (e.g., delay, trace, or simultaneous), number of trials per pairing, inter-trial interval (ITI) duration, frequency of preference assessments to be administered, elimination of pre-session exposure to target sounds or not, the optimal number of pairings between sound model (NS) and US, whether to use observing prompts, alternating pairing trials with control trials, and using pre-recorded sounds all have been examined as independent variables (Petursdottir et al., 2011; Shillingsburg et al., 2015). Unfortunately, these explorations have not identified the most effective SSP procedures for increasing vocalizations, leaving practitioners to estimate or use trial-and-error approaches in SSP application. The most effective procedure involves delay conditioning, where the reinforcing stimulus (US) is presented while the sound (CS) is still ongoing (e.g., Carroll & Klatt, 2008; Miguel et al., 2002; Lepper et al., 2013; Miliotis et al., 2012; Normand & Knoll, 2006) and trace conditioning, where the US is presented at the end of the CS or afterward (e.g., Barry et al., 2019; Esch et al., 2009; Smith et al., 1996; Sundberg et al., 1996; Rader et al., 2014). Shillingsburg et al. (2015) reported that, across all the studies, delay conditioning produced stronger effects (i.e., higher rates of vocalizations following SSP) compared to trace and simultaneous conditioning.

The same respondent procedure described as SSP has long been performed by researchers in basic laboratories (da Silva & Williams, 2020) to elicit responses other than vocalizations. Based on that research, Madden et al. (2023) provided some guidance for effective implementation of SSP. One example is the use of a larger *C/T ratio*, where *C* is the average time between US presentations and *T* is the duration of the NS onset to US delivery. This *C/T* ratio benefits respondent conditioning outcomes by making the sound model (NS) more salient (due to longer trial spacing). The sound model (NS) is also presented alone for a briefer duration in a larger *C/T* ratio, signaling the US is soon to follow. This predictability helps develop the CR more rapidly than shorter *C/T* ratios. Procedures with longer *C/T* ratios align with the delay reduction hypothesis, explaining why shorter sound model (NS) durations elicit more conditioned responding For example, *autoshaping* (Brown & Jenkins, 1968) occurs when a NS such as a sound or light is presented in the operant chamber just prior to delivery of food (US), typically using either trace or delayed conditioning (i.e., with or without a temporal gap between NS and US). Repeated pairings of NS and US result in the occurrence of the CR to the NS alone, indicating it has become a conditioned stimulus (CS).

Similar to Madden et al. (2023), da Silva and Williams (2020) also discussed some of the variables relevant to SSP by describing basic research outcomes of autoshaping. For example, in autoshaping there often is an inverse correlation between the NS duration and its effectiveness (Brown & Jenkins, 1968). According to Ricci (1973), when 30-s key lights preceded food, pigeons acquired responding (i.e., keypecking) more quickly than when 120-s key lights were presented. Alternately, Brown and Jenkins (1968) found that 8-s key light presentations were more effective comparatively at acquiring responding than 3-s key light presentations. da Silva and Williams (2020) theorized that there appears to be an optimal duration of NS that influences autoshaping and that duration depends in part on the inter-trial interval (ITI), or time between the end of US delivery and beginning of the next NS presentation. This is a very similar approach to Madden et al.'s discussion of the *C/T ratio*. They emphasized the results of basic research showing longer "Cycles" (i.e., *C*), defined as the time between US presentations, leading to longer ITIs.

A prior assessment of ITI impacts was Kaplan (1984), who measured approach responses by pigeons to the NS (i.e., keylight) and withdrawal responses from the NS to evaluate the effects of the relative duration of the ISI to ITI. In Experiment 1, a key light (NS) was presented for 12 s, followed by a 12-s ISI. The duration of the ITI was then systematically varied between 15 and 240 s across five groups of pigeons. It was expected that pigeons would display different rates of acquisition, and that indeed was the case. Kaplan observed that excitatory conditioning (or approach to the NS) occurred when the ITI was longer (e.g., 240 s); inversely, inhibitory conditioning (withdrawal from the NS occurred with shorter ITIs (e.g., 15 s). Overall results indicated that the pigeons tended to approach the NS when the ITI was greater than 60 s, in which they engaged in withdrawal responses when the ITI was less than 60 s. Taken together, the success or failure of SSP might depend on the relationship between the duration of the sound presentations and ITI durations. It then behooves us to further examine these variables, in that shorter sessions with fewer trials might be more desirable than the opposite (Madden et al., 2023).

METHOD

Subjects and Assessments

Nine typically developing children between ages 15 and 21 months participated. The children's parents reported that the subjects were reaching their developmental milestones on time and did not have any health concerns. The experimenter selected such subjects to rule out the influence of developmental delays, neurodevelopmental diagnoses, or any subject characteristics that have not been controlled in previous research (Petursdottir & Lepper, 2015). The children included in the study exhibited the following behaviors as expected for their age: eye contact, social smiles, imitating gestures like pointing, clapping, waving, responding to their

Table 1. VB-MAPP Results Across Subjects.

Verbal Operant	Matt	Tess	Sam	Nelson	Ben	Nolan	Ave	Ken	Carla
Mand	4	3	1	0	0	3	2	1	4
Tact	2	4	1	0	0	3	2	0	5
LR	5	5	5	5	5	5	5	5	5
Echoic	2	5	1	0	4	2.5	4	1	4
Total	13	17	8	5	9	13.5	13	7	18

Note: Individual skill scores were obtained by scoring the subject response based on the criteria identified in each section of the specific milestone scoring form. A response was scored based on three options: 0, ½, or 1. Then, total scores were obtained by adding up all the points acquired by the subject for each skill area.

name, showing enjoyment during simple games like peek-a-boo, shaking their head to mean no, waving bye-bye, and pointing to items or to get the parent's attention (AutismStep, 2022). However, subjects that demonstrated vocalizations under echoic control were excluded from the study. To determine the children's verbal repertoire, the experimenter used the Verbal Behavior Milestones Assessment and Placement Program (VB-MAPP; Sundberg, 2008) and the Early Echoic Skills Assessment (EESA) contained within (see Table 1). The children's verbal repertoire evaluated were mand, tact, listener, and echoic responses. Additionally, the children were observed during one 30-min free-play session, the frequency and topography of their vocalizations were recorded (see Table 2 for the results of the initial vocalization levels during the observation sessions).

Stimulus Preference Assessment

We used a paired-stimulus preference assessment (Fisher et al., 1992) to determine preferred food items for each child (see Table 3). The food items suggested by the parents were presented in pairs, and the child was asked to choose one within 5 s. If no choice was made, the child was prompted. If no choice after the prompt, the food items were removed, and both were scored as not chosen. Each food item was presented three times (allowing for all possible combinations of items to be presented), and the percent of times each item was chosen was calculated. The most preferred food item for each child was used in the pairing and control conditions. The food items used were crackers (Carla, Ken, and Nolan), cereal (Ben, Tess, and Nelson), yogurt melts (Eva), and cookies (Sam). No reinforcer assessment was conducted.

Settings and Materials

Sessions were held in locations that met the needs of the subjects and the preferences of their parents. For instance, some sessions were conducted in the children's bedrooms (Nolan, Carla, and Eva), family rooms (Matt and Tess), and the psychology department's laboratory space on a small liberal arts college campus (Ken, Ben, and Nelson). Parents were invited to attend the sessions but were not directly involved in any experimental procedures. Sam's sessions were held in the home of the first author at the request of the parent, who

Table 2. Subject Characteristics and Vocalizations Observed

Group	Subject	Age	Vocalizations per min	Examples of Vocalizations
20-s ITI	Carla	18 m.o.	0.33	"Mermaid," "Baby," "Bottle," "Star," and "Glub-glub" for fish as tact. "Please" (mand).
	Ken	19 m.o.	2.13	"Eee," and "Ook"
	Matt	18 m.o.	1.06	"Oh!", "Yay," and "Ouch" (mand). "Kitty cat," (tact), "Meow", "Woof", and "Roar" (intraverbal).
30-s ITI	Eva	18 m.o.	1.17	"Uh-oh" (mand), "Ee-ee," "Ah," and "Oiaa."
	Nolan	21 m.o.	1.27	"Ah", "Oh," "Uh-Oh,". "Horse," "Sheep," "Dog," "Book," "Ed" for red, "Fog" (for frog) as tact; and "Ball" and "Book" as mand.
	Tess	17 m.o.	0.6	"Star," "Shoe" and "Ook (for book), "Oat" (for boat), "Keys," "Kup" (for cup), and "Nana" (for banana) as tacts. "More," "Bubbles," "Wawa" (for water), and "Ook" (for book) as mands
60-s ITI	Ben	16 m.o.	2.07	"Asit," "Uh-oh," "Eech," "See Ya," "Chututu." "Shoe" and "Da-da" (tact).
	Nelson	21 m.o.	0.43	"Na-Na"
	Sam	15 m.o.	0.43	"Ee-ee" and "What?" (noncontextual)

Note: Table 2 represents vocalizations per minute that each subject emitted during the 30-min free play observation and scores.

Table 3. Subjects' Stimulus Preference Assessment Outcomes

Subject	Highest Preferred Stimulus	Moderately Preferred Stimulus	Least Preferred Stimulus	No selection
Carla	Bunny crackers (50%)	Goldfish crackers (33%)	Puff cereal (17%)	Cheerios (0%)
Ben	Puff cereal (50%)	Yogurt melts (33%)	Bunny crackers (17%)	Goldfish crackers (0%)
Eva	Peach yogurt melts (67%)	Vanilla yogurt melts (33%)		Puff cereal and bunny crackers (0%)
Ken	Bunny crackers (50%)	Goldfish crackers (33%)	Cheerios (17%)	Puff cereal (0%)
Sam	Oreo cookies (50%)	Puff cereal (33%)	Bunny crackers (17%)	Goldfish crackers (0%)
Matt	Fruit gummies (100%)	*Goldfish crackers		Popcorn and Puff cereal (0%)
Tess	Puff cereal (50%)	Bunny crackers (33%)	Cheerios (17%)	Goldfish crackers (0%)
Nolan	Cheerios (50%)	Goldfish, bunny, and cheese crackers (17% respectively)		
Nelson	Cheerios (50%)	Goldfish cracker (33%)	Puff cereal (17%)	Yogurt melts (0%)

Note: Selection percentages rounded to the nearest whole number are shown below the stimulus within the parentheses. Highest preferred stimulus was used as the unconditioned stimulus during pairing and control conditions. Moderately preferred stimulus was used as preferred item during adaption. For Matt, *Goldfish crackers were arbitrarily selected as most preferred based on parental suggestion.

preferred a location close to their residence but did not consent to having the sessions conducted in their own home. The experimenter utilized a small room that was separated from the main living areas. The parent was provided with comfortable seating and access to a private bathroom. Throughout the sessions, the parent always accompanied the child. Toys (e.g., puzzles, ring stacker, shape form, trains and cars, and books) were available to all subjects across all sessions. These toys were intended for the entertainment of the children during sessions, and they were not assessed or otherwise identified as preferred items. All sessions were recorded with a video camera 10-Sony HDR-CX675 High-Definition Camcorder set in a fixed location of the room to provide the best view.

The experimental setup included a 60 cm by 60 cm folding table, an Insignia-Wave 2 portable Bluetooth speaker, and a Homdox automatic food dispenser with a 10 cm by 15 cm black-and-white picture of a woman's face placed on the folding table (see Figure 1). Before a session, the child's preferred food was added to the top portion of the apparatus, and food was concealed by a plastic cover. During specified times according to the protocols for pairing and control trials, the experimenter pressed a button

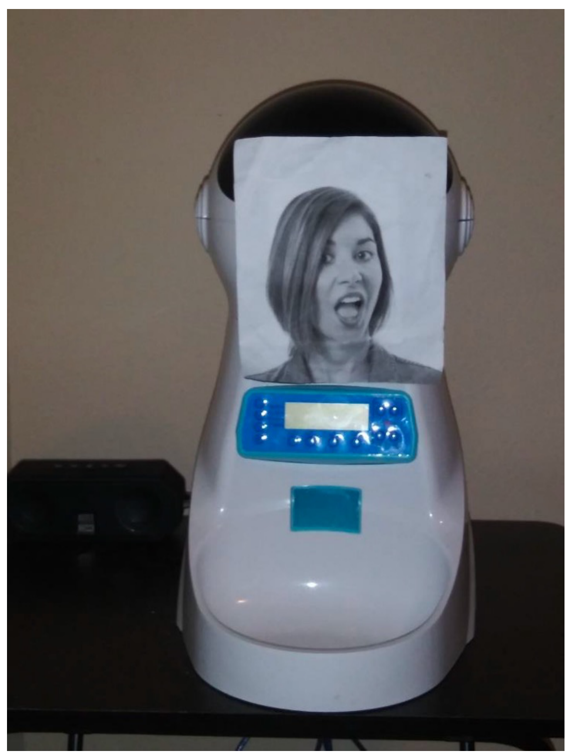

Figure 1. Apparatus Used in Pairing and Control Conditions

Table 4. Distribution of Subjects per Group and Selected Vocal Responses

Group	Subject	Vocal Response Type	Target Vocal Response
20-s ITI	Carla	Novel	Hungry
	Ken	Novel	Grandpa
	Matt	Novel	Apple
30-s ITI	Eva	Novel	Up
	Nolan	Low frequency	Eat
	Tess	Novel	Want
60-s ITI	Ben	Novel	Tissue
	Nelson	Novel	Banana
	Sam	Novel	Exit

to dispense a small portion of the child's preferred food into a tray for the child to consume.

The sound model (NS) was pre-recorded for consistent presentation across trials in both the pairing and control conditions. It was played from a computer to the speaker at 70 dB (Vouloumanos et al., 2009). The experimenter pre-recorded the auditory stimulus (like Petursdottir et al., 2011) by melodically repeating the target word in an exaggerated manner (i.e., motherese modeling) for a predetermined interval (e.g., 3, 6, or 12 s) to fill the ISI. The same sound model (NS) was used for the same participant across conditions to standardize the trace conditioning procedure (i.e., the pairing, or lack thereof, distinguished the trials).

Selection of Target Vocalizations

Table 4 shows the distribution of subjects per group and selected targets. Target vocalizations were either novel or low frequency. This selection offers two benefits: 1) Novelty enhances the salience of the sound model (NS), and, therefore, the subject would be more likely to attend to the sound, and 2) a novel or low-frequency word does not have an extensive learning history that could negatively impact SSP learning (da Silva & Williams, 2020). Novel target vocalizations were words that were known to be new (as per parents reports) or did not occur in the free-play observation period (Shillingsburg et al., 2015; Sundberg et al., 1996) or selected from combining vocalizations that occurred about 10%-25% of the time during the free-play session but that weren't evoked by echoic prompts (Esch et al., 2009). Low-frequency targets were those vocalizations the subject spoke infrequently during the assessment (less than 10% of the time). For instance, Nolan's word was "Eat," which was an infrequent word as he said it once during the assessment. This response seemed to be prompted by a nonverbal stimulus (a photo of a cake) and was categorized as a tact (by function). "Eat" did not occur when prompted verbally or when food was offered. Ben's word was "Tissue," derived from combining the initially observed vocalizations ("t" and "shoe"). Ken did not emit any intelligible words during the initial observation. Thus, his parents suggested a novel word, "Grandpa." Nelson's word was "Banana," selected based on the observation, in which he emitted "na-na."

Similar methods were used to select sound models (NS) for Carla, Eva, Sam, Tess, and Matt (see Table 4) but no attempt was made to equate sound models (NS) across subjects due to the difficulty of choosing words each one could produce.

Response Definition

The main dependent variables were number of target vocalizations per min, percentage of trials with target vocalizations, and percentage of trials with approach to the apparatus. All sessions were video recorded for later data coding and analysis.

Number of Target Vocalizations per Min

A target vocalization was defined as any sound or word produced by the subject that matched (e.g., target vocal response is "apple," and the child says "/ap/ple/") or shared similar acoustical features as the target auditory stimulus (e.g., target vocal response is "apple," and the subject says2 /a/pooh/). Any other emitted vocal responses (e.g., "oooo" or "aaaa") were not counted as correct. The number of target vocalizations was recorded on each trial, where any vocalizations separated by 1 s were recorded as separate responses (Esch et al., 2009). The number of target vocalizations per min was calculated by the sum of target vocalizations produced by the subject divided by the total duration of the session.

Percentage of Trials with Vocalizations

The percentage of trials with vocalizations was calculated by dividing the number of trials in which target vocalizations occurred by the total number of trials.

Percentage of Trials with Approach to the Apparatus

If the subject turned their head or walked toward the apparatus within 15 s of the onset of a pre-recorded target sound model (NS), an approach response was scored. The percentage of trials on which an approach response occurred was calculated by dividing the number of trials on which an approach response occurred by the total number of trials.

Interobserver Agreement

Two independent observers collected data on the dependent variables after sessions from video recordings. Inter-observer agreement (IOA) was assessed for 50% of sessions for all subjects by using exact count-per-interval IOA (Cooper et al., 2020) and was calculated by dividing the number of intervals of 100% agreement by the total number of intervals for each session and multiplying by 100. Agreements were defined as both observers indicating the same number of occurrences or nonoccurrence of behavior in each interval, whereas a disagreement occurred when only one of the two observers recorded the behavior in a corresponding interval. Mean IOA was 93% (range, 86% to 100%) for Eva, 98% (range, 95% to 100%) for Nolan, 94% (range, 88% to 100%) for Carla, 98% (range, 95% to 100%) for Ben and Sam, and 100% for Ken, Nelson, Tess, and Matt.

Design and Procedure

Nine subjects were randomly assigned to one of three experimental groups that varied according to ITI length as a systematic replication of Kaplan (1984). Unlike previous studies in SSP (Carroll & Klatt, 2008; Esch et al., 2009; Miliotis et al., 2012), this study did not include a baseline phase to reduce the exposure of the sound model (NS) before pairing procedure, as recommended by da Silva and Williams (2020) and Madden et al. (2023). In respondent conditioning, repeated exposure to a neutral stimulus can interfere with the conditioning of that stimulus during, a phenomenon known as latent inhibition. This can also weaken sign-tracking behavior, or attending to a visual stimulus which signals potential availability of rewards (cf. Boughner & Papini, 2008; da Silva & Williams, 2020; Madden et al., 2023).

A pairwise design (Iwata et al., 1994) without baseline was used to compare the effect of pairing (i.e., experimental) trials to control trials by sequentially alternating pairing and control conditions across sessions. Each session contained only pairing trials or control trials. Both test and control sessions were conducted on the same day, but there were occasions where only one condition (i.e., either pairing or control) was conducted due to time constraints, a subject's illness, or technical issues (e.g., if the video camera stopped working). If both sessions could not occur on the same day, the experimenter arranged to make up the missed session that same week or the next.

Sessions were arranged close to or before any scheduled snacks or meals, if possible. A session was suspended by 10 min if a subject acted tired or began to cry. If the child continued to be irritable beyond the 10-min break, the session was terminated and a make-up session was scheduled. During the experiment, the parents were invited to be present in the same room but were instructed to avoid any verbal interaction with their child. Parents were encouraged to ask for the assistance of the experimenter when needed or to stop the session at their discretion. Subject-experimenter interactions were limited across experimental conditions to control influence of familiar/unfamiliar pairing agent (Petursdottir & Lepper, 2015).

In the 20-s and 30-s ITI groups, the experimenter divided sets of pairing and control sessions into two days, and each experimental session contained 30 trials per pairing and 30 trials per control. In the 60-s ITI group, sets of pairing and control sessions were divided into three days, in which each session had 20 trials per session. Session lengths varied but never exceeded 35 min. When variable 20-s ITIs were used, average session durations were 22 min for pairing and 21 min for control conditions; when 30-s ITIs were used, average session durations were 26 min for pairing and 25 min for control conditions; when 60-s ITIs were used, average session durations were 31 min for pairing and 27 min for control conditions.

Experimental Conditions

Adaptation. Each subject was trained to pick up the edibles from the apparatus tray. The experimenter placed three pieces of the previously identified preferred food in the apparatus tray and waited 15 s for the subject to consume the edibles. If there was no response, a verbal (e.g., "There is food here") and gestural prompt (e.g., pointing to the tray) were provided to initiate the response. After the subject consumed the food, the experimenter set a timer and dispensed edibles into the tray on a variable time (VT) 30-s schedule. During these sessions, the subject was otherwise free to walk around the room and play with toys without restriction. The adaptation phase ended when the subject picked up the food from the tray within 5 s of food presentation across five successful trials. The subjects completed this phase in an average of 5 min. Following this phase, the subject received a 5-min break, and the experimenter arranged the apparatus for the first session.

Pairing. The experimenter began pairing sessions by presenting a pre-recorded sound model (NS) selected for the child for the entire pre-determined ISI. Upon termination of the sound model (NS), the experimenter then pressed a button on the apparatus to dispense pieces of the preferred edibles (US) into the tray for the child's consumption. If the subject did not consume the food within 5 s of US presentation (as per Esch et al., 2009), the experimenter removed it from the tray. Then, the experimenter set a silent digital timer for the predetermined ITI duration before presenting the sound model (NS)/CS and beginning the next trial. Across groups, the average ITI duration was varied, with mean values of 20 s (range: 15-25 s), 30 s (range: 20-40 s), and 60 s (range: 40-80 s). ISI duration was determined according to a 1:5, or 20%, ratio with the designated ITI value for the variable 30-s and variable 60-s ITIs, and a 3:20, or 15%, ratio for the variable 20-s ITI. We chose these ratios based on Kaplan (1984). ISIs and ITIs for each subject are depicted in Tables 5-7.

Control. Control sessions began by setting two independent silent digital timers (VT 20 s, 30 s, and 60 s) that cued the presentation of the sound model (NS) and food (US) randomly and independently (i.e., random control; Rescorla, 1967), which led to a 50/50 distribution of stimulus presentations. The experimenter arranged the stimulus presentations independently from the occurrence of the subject vocalization or consumption of the food. If the subject did not consume the food within 5 s of its presentation, the experimenter scooped the food from the tray. One important outcome of Kaplan's (1984) study is that the failure to observe CR (i.e., target sound or word) is *not* explained by a failure of the procedure itself. Instead, failure can be explained due to excitatory and inhibitory conditioning of equal strength canceling out each other. In short, the application of Kaplan's findings to SSP research can aid in identifying the optimal ISI for conditioning target vocalizations. Past studies support the utility of varying and lengthening ITI duration in teaching vocalizations reported by Barry et al. (2019), Esch et al. (2009), and Rader et al. (2014) when they varied ITIs between 5-30 s using SSP. In this study, we evaluated three durations of ITIs for eliciting human vocalizations in an SSP paradigm, based on previous research with pigeons by Kaplan (1984) and the theoretical suggestions of da Silva and Williams (2020).

RESULTS

Figure 2 displays the number of vocalizations per min made by each subject in three experimental groups. The left panel shows the number of vocalizations per min for Carla, Ken, and Matt in the 20-s ITI group. For Carla (top panel), the mean increases from control to SSP were <0.1 vocalizations per min. Initially, Carla's vocalizations increased in both control and pairing but later decreased to zero in both conditions. However, the rate of vocalizations was slightly higher in the pairing conditions, at about 0.3 vocalizations per min, compared to the control condition, which was <0.1 vocalizations per min. For Ken (middle panel), the mean increase in vocalizations from the control to the pairing condition was approximately 1.5 vocalizations per min. Ken's vocal responses greatly increased during the first pairing session, reaching about 2.6 vocalizations per min, but then decreased during the second pairing session to about 0.3 responses per min. For Matt (bottom panel), the mean increase from the control to the pairing condition was <0.5 vocalizations per min. Matt's vocal responses also increased in both the pairing and control conditions, but the rate of responses was slightly higher in the SSP condition (0.7 vocalizations per min) compared to the control condition (0.2 vocalizations per min).

The middle panel represents the total number of vocalizations per minute for Eva, Nolan, and Tess in the 30-s ITI group. For Eva (top panel), the mean increase from control sessions to pairing sessions was almost none, slightly less than 0.1 vocalization per min. Eva's vocal responses significantly increased across the pairing sessions, with an increase of about 0.1 vocalization per min in Session 2 and Session 4. By contrast, there were no increases in her vocal responses during the control sessions. For Nolan (middle panel), the mean increase from control to pairing was also about 0.1 vocalization per min. Nolan's vocal responses increased to 0.3 vocalizations per min in the first pairing session but then decreased to <0.1 vocalizations per min in the second pairing session. Interestingly, during the control sessions, his vocal responses also increased by 0.2 vocalizations per min. This indicates that, as the rates of vocal responses decreased in the pairing sessions, they increased in the control sessions. For Tess (bottom panel), the mean increase from control to pairing was 0.3 vocalizations per min. Tess's vocal responses slightly increased in both pairing sessions, showing increases of 0.2 vocalizations per min in Session 1 and 0.3 vocalizations per min in Session 3. There were no vocal responses recorded during the control sessions.

The right panel displays the total number of vocalizations per min for Ben, Nelson, and Sam in the 60-s ITI group. For Ben (top panel), the mean increase in vocal response from the control to the pairing phase was about 0.2 vocalizations per min. Although Ben's vocal responses increased initially, they eventually dropped to zero across all sessions. In Nelson's case (middle panel), the mean increase from control to pairing was 0.3 vocalizations per min. Nelson's vocal responses showed a progressive increase throughout the pairing sessions, with rates of 0.1, 0.4, and 0.5 vocalizations per min, respectively. Notably, there were no vocal responses recorded during the control sessions. For Sam (bottom panel), the mean increase from

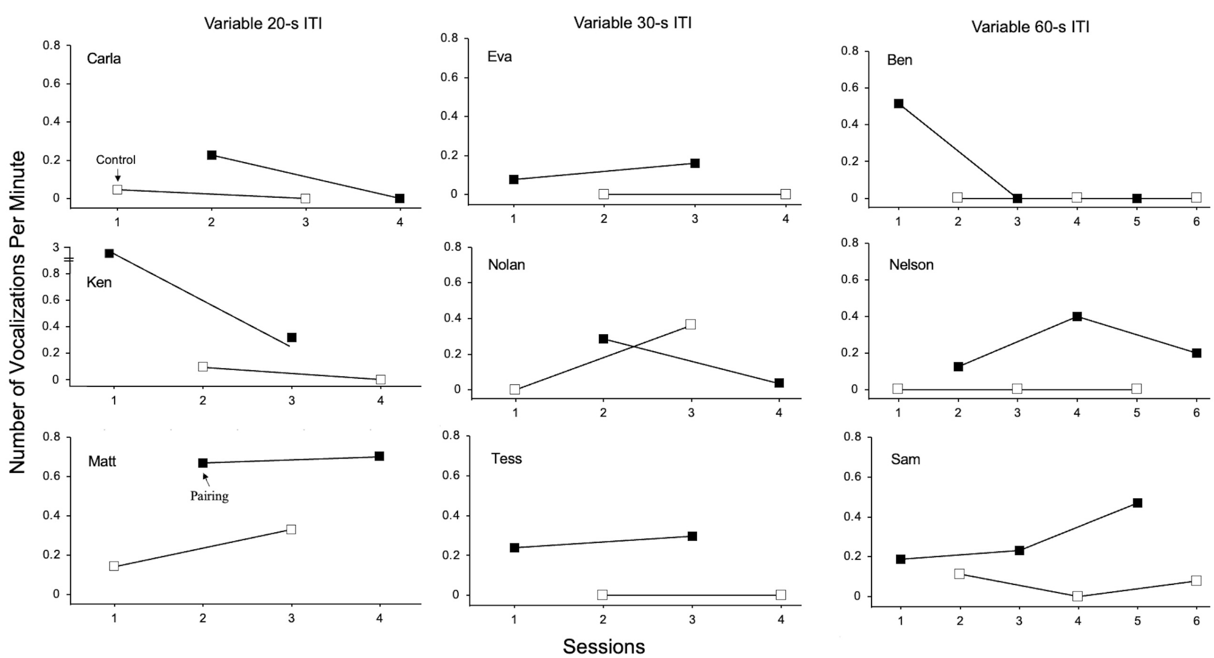

Figure 2. Vocalizations per Min

control to pairing was 0.3 vocalizations per min. Sam's vocal responses also increased initially to 0.5 vocalizations per min but subsequently decreased and stabilized at similar levels across pairing Sessions 4 and 6, with rates of 0.2 responses per min, respectively. In the control sessions, vocal responses remained low, recorded at <0.1, 0, and <0.1 vocalizations per min. These results showed that all subjects had the highest rates of vocalizations in the pairing condition and the lowest rates of vocalizations in the control condition. These results are consistent with previous studies (Barry et al., 2019; Esch et al., 2009; Rader et al., 2014). The implications of these findings could be a beginning of functional vocalizations, as subjects were able to learn produce formerly novel sounds/words without the need for contingent reinforcement, echoic training, or prompting (Sundberg et al., 1996).

In Figure 3, the percentage of trials with vocalizations produced by each subject in three experimental groups is represented for 20-s, 30-s, and 60-s ITIs. In general, it was observed that the percentage of trials with vocalizations was lower in the second pairing sessions for both 20-s and 30-s ITIs; but for 60-s ITI group, the percentage of trials with vocalizations tended to increase in the second and third pairing sessions for two of three subjects. These findings suggest that SSP may have an immediate effect on some subjects' vocalizations, regardless of the type of variable ITI (Esch et al., 2009). In addition, one can speculate that the subjects in group 60-s ITI behaved as expected because waiting times were longer when trials were further spaced out. This outcome aligns with respondent conditioning research in autoshaping (Jenkens et al., 1981) and was predicted by da Silva and Williams (2020) and Madden et al. (2023).

The left panel of Figure 3 displays the percentage of trials with vocalizations. Ken and Matt had higher percentages of vocalizations in the first pairing sessions, with 27% and 20% respectively, compared to Carla's 13%. However, in the control sessions, all three subjects had very low percentages of trials with vocalizations, averaging from 2% to 8%. Both Carla and Ken had a low percentage of trials with vocalizations in the control sessions, at 2%;

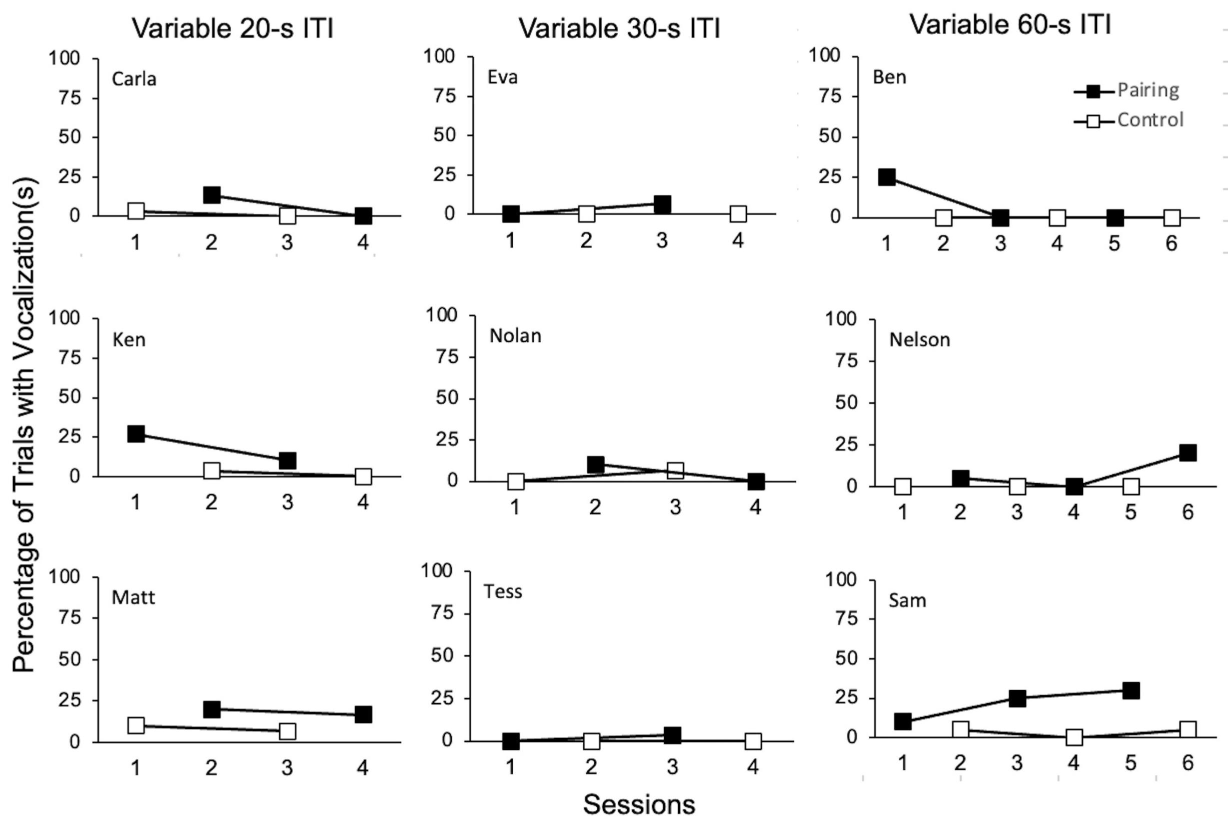

Figure 3. Percentage of Trials with Vocalizations

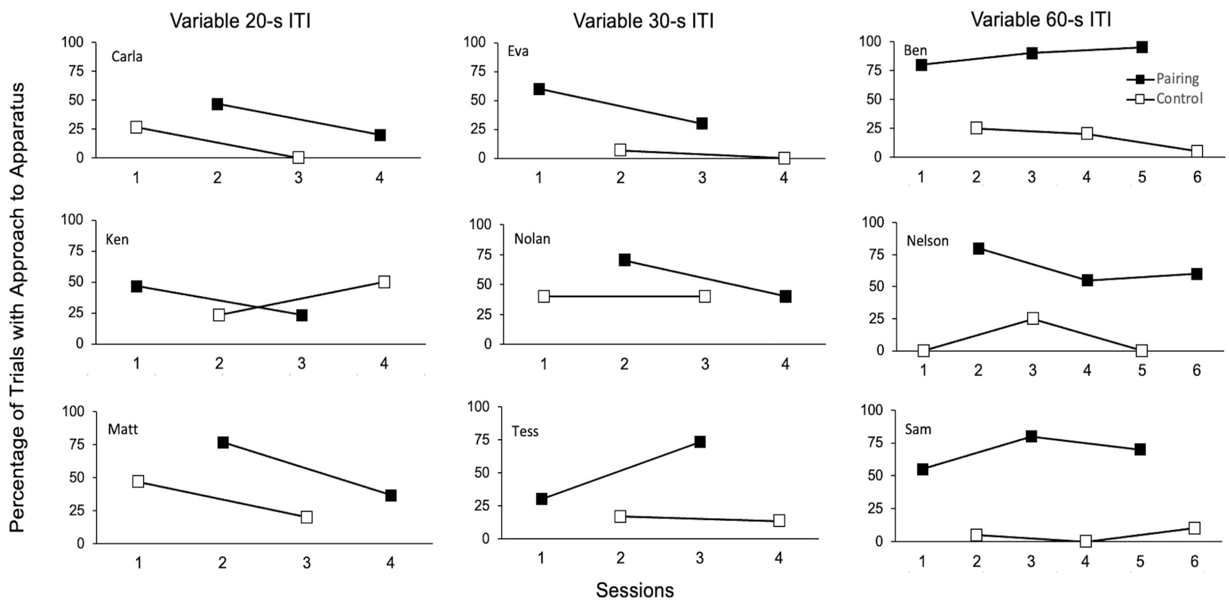

Figure 4. Percentage of Trials with Approach to Apparatus

contrarily, Matt showed 8% of trials with vocalizations, but it continued to be a lower percentage of trials in the control session than in pairing sessions. In summary, all three subjects in 20-s ITI group had the lowest percentage of trials with vocalizations in the control sessions and lower percentages of trials with vocalizations in the second pairing sessions. In the middle panel, Eva and Tess did not exhibit any vocalizations in the first pairing sessions, while Nolan had the highest percentage of trials with vocalizations at 10%. In the right panel, Nelson and Sam showed higher percentages of trials with vocalizations in the third pairing session at 20% and 30% respectively, compared to Ben, who only had 25% of trials with vocalizations in the first pairing session and none in subsequent pairing sessions

Figure 4 shows the percentage of trials in which each subject approached the apparatus in three experimental groups. In the left panel, Carla, Ken, and Matt from the 20-s ITI group had the highest percentage of apparatus approaches in the first pairing sessions, with Carla and Ken at 47% and Matt at 77%. Interestingly, all three subjects also showed increased approach levels ranging from 25% to 50% in the control condition. In the middle panel, Eva and Nolan approached the apparatus in about 60% of the trials in the first pairing sessions, while Tess's approach behavior increased from 30% to 73% in the second pairing session. The right panel shows all subjects from the 60-s ITI group consistently approached the apparatus across pairing sessions, representing the highest percentage of approaches among the three experimental groups. Ben had an average approach response of 88% of trials, followed by Sam at 68%, and Nelson at an average of 65%. All three subjects approach the apparatus less often during the control condition.

DISCUSSION

The results of this study support and extend the findings of previous research in SSP by showing that young children with various levels of language skills can show modest improvements in novel vocal responding through SSP, a procedure akin to autoshaping (da Silva & Williams, 2020) of non-human responding. As a result of a trace conditioning respondent procedure, which resulted in brief durations between sound model (NS) onset and US but relatively long ITIs in comparison (Madden et al., 2023), all subjects showed slightly higher rates of vocalizations in the pairing than the control conditions. Nelson, who had never spoken any words before the training, was able to produce the target response on the first pairing session and responding levels continued to increase in subsequent pairing sessions (but never in the control condition).

These initial findings could begin a new approach to understanding the relevance of SSP as a clinical procedure that can aid nonvocal children who would not benefit from vocal shaping or echoic training due to their lack of production or imitation of initial sounds, by examining recommendations from basic research (da Silva & Williams, 2020). Moreover, prior research (Carroll & Klatt, 2008; Miguel et al., 2002) speculated that the type of vocalizations (i.e., novel versus existing vocalizations) might have differential effects on SSP. In fact, Carroll and Klatt (2008) reported it might be easier to increase existing rather than novel vocalizations, which might account for the modest improvements in our study.

The purpose of the current study was to examine the effects of varying the duration of the ITI and the ISI in of human vocalizations, based on recommendations from respondent research in autoshaping (Brown & Jenkins, 1968; Kaplan, 1984). Previous research in SSP (Barry et al., 2019; Esch et al., 2009; Miliotis et al., 2012; Rader et al., 2014) found success using variable ITIs, which seems important because the variability eliminates the temporal predictability of the presentation of preferred items (i.e., the US is unexpected) and is consistent with results from autoshaping research (da Silva & Williams, 2020). We hypothesized that the relative duration of ITI compared to ISI might have a critical impact on the conditioning of human vocalizations, as has been observed in the conditioning of other species' responses (e.g., Jenkins & Moore, 1973; Sidman & Fletcher, 1968). To date, no SSP studies have evaluated the isolated effect of these variables (Madden et al., 2023) but, rather, have focused on other factors (e.g., subjects' age and diagnosis, type of pairing, number of sound presentations, type of reinforcer, and control procedures).

Somewhat in agreement with the findings of Kaplan's (1984) study and as predicted by Madden et al. (2023), human subjects were sensitive to the relative values of ISI and ITI, and they also showed an approach response in the longer ITI condition. Longer ITI durations might be preferable due to increasing the chances for subjects to acquire sign-tracking behavior (i.e., orientating to the sound model, or NS). However, subjects in the 30-s ITI group demonstrated mild approach response to the sound model (NS) and did not perform as expected, which indicates that this condition can also be appropriate if shorter pairing trials are desired. Additionally, subjects in the 20-s ITI condition failed to acquire any clear tendency to approach the sound model (NS), and these subjects performed similarly to the nonhuman subjects in Kaplan's 60-s ITI condition. Future research should measure whether observed approaches occur before or after delivery of the US and also continue to tease apart some of the additional factors contributing to these mixed findings.

Our results are relevant to explain SSP failure to condition vocalizations for some or all participants in previous SSP studies (Carroll & Klatt, 2008; Esch et al., 2005; Miguel et al., 2002; Normand & Knoll, 2006; Stock et al., 2008) in which ITI duration was shorter than 30 s (e.g., 0-20 s). Nevertheless, Ken (a subject in Group 1, the shortest ITI group) showed the highest rates of vocalizations per min in the pairing conditions of all subjects in all groups. This difference could be due to the subject's individual history of reinforcement. Another possible contributing factor was the use of motherese modeling (i.e., melodic voice) and novelty to enhance the sound model (NS). Shillingsburg et al. (2015) reported motherese modeling and novel sounds can be effective in increasing rates of vocal responses. An interesting aspect of this study is that some subjects emitted newly paired vocalizations sporadically throughout the day outside of training sessions, as reported by their parents. This finding appears to be consistent with results reported in previous studies (e.g., Smith et al., 1996; Sundberg et al., 1996). As discussed previously, vocalizations can become more sensitive to automatic reinforcement when the child hears themself produce sounds or words that share similar acoustic features with the word they heard (Petursdottir & Mellor, 2017; Shillingsburg et al., 2015; Smith et al., 1996).

In a departure from many prior studies, we implemented a truly random control procedure (Rescorla, 1967) as opposed to evaluating the utility of an omission control procedure (i.e., correction delay) to control for adventitious reinforcement of responses. Previous SSP studies that included omission control procedure obtained mixed results (Carroll & Klatt, 2008; Miguel et al., 2002; Normand & Knoll, 2006; Stock et al., 2008). Rescorla (1967), da Silva and Williams (2020), and Madden et al. (2023) each pointed out that this control procedure not only removes the contingency between the sound model (NS) and US but also introduces a new one: the US cannot follow the sound model (NS) during the delay. Consequently, the sound model (NS) signals the absence of the US instead of its presence, contrary to the learning that should be facilitated in SSP. Therefore, failure to produce target sounds in Norman and Knoll (2006) study could be attributed, at least in part, to this effect instead of SSP inefficacy. The truly random control procedure seemed more adequate than the omission control procedure because the latter is based on the idea that the occurrence of the sound model (NS) gives no information about the occurrence of the US. In this study, seven of nine subjects showed differentiated low to zero rates of responding in the control condition compared to the pairing condition, and all acquired the target vocalizations. However, Matt and Nolan also exhibited moderate rates of vocalizations in the control condition. First, we speculate these subjects might have failed to discriminate changes in the conditions; consequently, future research should include stimulus control procedures (e.g., colored cards) to signal to the subject the shift to the next condition. Second, Matt and Nolan might have exhibited target vocalizations because of the establishing operation for food (i.e., when the delivery of the US no longer followed the sound model, or NS), and such an EO could have evoked a mand response (as we observed the children waiting in front of the apparatus to receive the food while they emitted the words). Future research should apply SSP with other forms of reinforcement (e.g., physical touch, smiles, eye contact) and conduct reinforcer assessments of these to further evaluate the effects of SSP for establishing reinforcers.

An additional level of control considered was using an apparatus to present stimuli independent from a human trainer. We hypothesized the apparatus might have helped to standardize the delivery of the US and control other sources of social reinforcement (described above) that are mediated by a practitioner. A limitation noted was that the sounds from the apparatus when dispensing food and the food hitting the tray were not masked. We hypothesized that these extraneous noises could have influenced the sound model (NS)/CS value to elicit target vocalizations. Future research should aim to mask these noises, perhaps by using white noise (Kaplan, 1984) or foam to absorb the sounds. Another limitation in this study was that motherese modeling was included as a procedure to enhance CS salience, but its isolated effects were not measured nor controlled. Future research should examine the variation of sounds during pairing and control trials (Shillingsburg et al., 2015). Unfortunately, treatment integrity data were not collected, although the automation of the apparatus and model vocalizations were beneficial in that regard, reducing the potential for treatment errors.

An additional constraint was the duration of training sessions. On average, training sessions were about 20-30 min. Some subjects acted tired most often as the training session went along, and it appeared that conducting 20-30 trials per session (a total of 40-60 trials per day) was unpractical with these young subjects. It is still a relevant question to determine the optimal number of trials to observe the effect of SSP. Gallistel and Papachristos (2020) found that the rate of acquisition was faster when mice completed 2.5 trials per session. Therefore, the SSP practitioner might consider conducting shorter sessions with fewer trials spaced out across days. This practice might be more suitable for young learners and parents as implementers. In addition, establishing a learning criterion can also guide practitioners' decisions to begin other procedures (e.g., mand or echoic training) to capture emerging vocalizations. Future research should evaluate ways to increase the efficiency of SSP procedure in clinical settings.

The current study attests that SSP effectiveness relies, at least in part, on the relative temporal contiguity of events for conditioning to occur. However, as Madden et al. (2023) made clear, contiguity between stimuli alone is insufficient. Varying the duration of the ISI with respect to the ITI produced slightly positive outcomes among subjects across groups but showed that longer ITIs (e.g., 30-60 s) are preferable for acquisition of novel vocalizations. In conclusion, the study's findings provide empirical support for the effectiveness of varying the time between trials in SSP procedures. The study also sheds light on the importance of considering the duration of ITIs and the use of control procedures in research involving human vocalizations.

REFERENCES

AutismStep. (2022). Developmental Milestones Fact Sheet. Retrieved from: https://www.stepautism.com//developmental-milestones/

Barbera, M. L. (2007). *The verbal behavior approach: How to teach children with autism and related disorders.* Jessica Kingsley Publishers.

Barry, L., Holloway, J., & Gunning, C. (2019). An investigation of the effects of a parent delivered stimulus-stimulus pairing intervention on vocalizations of two children with Autism Spectrum Disorder. *The Analysis of Verbal Behavior*, *35*(1), 57-73. https://doi.org/10.1007/s40616-018-0094-1

Boughner, R. L., & Papini, M.R. (2008). Assessing the relationship between latent inhibition and the partial reinforcement extinction effect in autoshaping with rats. *Pharmacology, Biochemistry, & Behavior*, *89*(3), 432-443. https://doi.org/10.1016/j.pbb.2008.01.019

Brown, P. L., & Jenkins, H. M. (1968). Auto-shaping of the pigeon's key-peck. *Journal of the Experimental Analysis of Behavior*, *11*(1), 1-8. http://doi.org/10.1901/jeab.1968.11-1

Carroll, R. A., & Klatt, K. P. (2008). Using stimulus-stimulus pairing and direct reinforcement to teach verbal behavior to young children with autism. *The Analysis of Verbal Behavior*, *24*(1), 135-146. https://doi.org/10.1007/BF03393062

Cooper, J. O., Heron, T. E., & Heward, W. L. (2020). *Applied behavior analysis* (3rd ed.). Pearson.

da Silva S. P., & Williams A. M. (2020). Translations in stimulus-stimulus pairing: Autoshaping of learner vocalizations. *Perspectives on Behavioral Science. 43*(1), 57-103. https://doi.org/10.1007/s40614-019-00228-9

Esch, B. E., Carr, J. E., & Grow, L. L. (2009). Evaluation of an enhanced stimulus-stimulus pairing procedure to increase early vocalizations of children with autism. *Journal of Applied Behavior Analysis*, *42*(2), 225-241. https://doi.org/10.1901/jaba.2009.42-225

Esch, B. E., Carr, J. E., & Michael, J. (2005). Evaluating stimulus-stimulus pairing and direct reinforcement in the establishment of an echoic repertoire of children diagnosed with autism. *The Analysis of Verbal Behavior*, *21*(1), 43-58. https://doi.org/10.1007/BF03393009

Fisher, W., Piazza, C. C., Bowman, L. G., Hagopian, L. P., Ownes, J. C., & Slevin, I. (1992). A comparison of two approaches for identifying reinforcers for persons with severe and profound disabilities. *Journal of Applied Behavior Analysis*, *25*(2), 491-498. https://doi.org/10.1901/jaba.1992.25-491

Gallistel, C.R., & Papachristos, E.B. (2019). Number and time in *acquisition, extinction and recovery. Journal of the Experimental Analysis of Behavior*, *113*(1), 15-36, https://doi.org/10.1002/jeab.571

Gros-Louis, J., West, M. J., Goldstein, M. H., & King, A. P. (2006). Mothers provide differential feedback to infants' prelinguistic sounds. *International Journal of Behavioral Development*, *30*(6), 509-516. https://doi.org/10.1177/0165025406071914

Jenkins, H.M., Barnes, R.A., & Barrera, F.J. (1981). Why autoshaping depends on trial spacing. In C.M. Locurto, H.S. Terrace, & J. Gibbon (Eds.), *Autoshaping and conditioning Theory (pp. 255-284).* Academic Press.

Jenkins, H. M., & Moore, B. R. (1973). The form of the auto-shaped response with food or water reinforcers. *Journal of the Experimental Analysis of Behavior*, *20*(2), 163-181. https://doi.org/10.1901/jeab.1973.20-163

Kaplan, P. S. (1984). Importance of relative temporal parameters in trace autoshaping: From excitation to inhibition. *Journal of Experimental Psychology: Animal Behavior Processes*, *10*(2), 113-126. http://dx.doi.org/10.1037/0097-7403.10.2.113

Koegel, L. K., Koegel, R. L., Hurley, C., & Frea, W. D. (1992). Improving social skills and disruptive behavior in children with autism through self-management. *Journal of Applied Behavior Analysis*, *25*(2). https://doi.org/10.1901/jaba.1992.25-341

Iwata, B. A., Duncan, B. A., Zarcone, J. R., Lerman, D. C., & Shore, B. A. (1994). A sequential test-control methodology for conducting functional analyses of self-Injurious behavior. *Behavior Modification*, *18*(3), 289-306. https://doi.org/10.1177/01454455940183003

Lepper, T. L., Petursdottir, A. I., & Esch, B. E. (2013). Effects of operant discrimination training on the vocalizations of nonverbal children with autism. *Journal of Applied Behavior Analysis*, *46*(3), 656-661. https://doi.org/10.1002/jaba.55

Madden, G. J., Mahmoudi, S., & Brown, K. (2023). Pavlovian learning and conditioned reinforcement. *Journal of Applied Behavior Analysis*, *56*(3) 1-23. https://doi.org/10.1002/jaba.1004

Miguel, C. F., Carr, J. E., & Michael, J. (2002). The effects of a stimulus-stimulus pairing procedure on the vocal behavior of children diagnosed with autism. *The Analysis of Verbal Behavior*, *18*(1), 3-13. https://doi.org/10.1007/BF03392967

Miliotis, A., Sidener, T. M., Reeve, K. F., Carbone, V., Sidener, D. W., Rader, L., & Delmolino, L. (2012). An evaluation of the number of presentations of target sounds during stimulus-stimulus pairing trials. *Journal of Applied Behavior Analysis*, *45*(4), 809-813. https://doi.org/10.9101/jaba.2012.45-809

National Autism Association. (2018). Autism Fact Sheet. Retrieved from: https://nationalautismassociation.org/resources/autism-fact-sheet/

Normand, M. P., & Knoll, M. L. (2006). The effects of a stimulus-stimulus pairing procedure on the unprompted vocalizations of a young child diagnosed with autism. *The Analysis of Verbal Behavior*, *22*(1), 81-85. https://doi.org/10.1007/BF03393028

Petursdottir, A. I., Carp, C. L., Matthies, D. W., & Esch, B. E. (2011). Analyzing stimulus-stimulus pairing effects on preferences for speech sounds. *The Analysis of Verbal Behavior*, *27*(1), 45-60. https://doi.org/10.1007/BF03393091

Petursdottir, A. I., & Lepper, T. I. (2015). Inducing novel vocalizations by conditioning speech sounds as reinforcers. *Behavior Analysis in Practice*, *8*(2), 223-232. https://doi.org/10.1007/s40617-015-0088-6

Petursdottir, A. J. & Mellor, J. R. (2017). Reinforcement contingencies in language acquisition: Implications for language intervention. *Behavioral and Brain Sciences*, *4*(1), 25-32. https://doi.org/10.1177/2372732216686083

Rader, L., Sidener, T. M., Reeve, K. F., Sidener, D. W., Delmolino, L., Miliotis, A., & Carbone, V. (2014). Stimulus-stimulus pairing of vocalizations: A systematic replication. *The Analysis of Verbal Behavior*, *30*(1), 69-74. https://doi.org/10.1007/s40616-014-0012-0

Ricci, J. (1973). Key pecking under response-independent food presentation after long simple and compound stimuli. *Journal of the Experimental Analysis of Behavior*, *19*(3), 509-516. https://doi.org/10.1901/jeab.1973.19-509

Rescorla, R. A. (1967). Pavlovian conditioning and its proper control procedures. *Psychological Review*, *74*(1), 71–80. https://doi.org/10.1037/h0024109

Rosales-Ruiz, J., & Baer, D. M. (1997). Behavioral cusps: A developmental and pragmatic concept for behavior analysis. *Journal of Applied Analysis of Behavior*, *30*(3). https://doi.org/10.1901/jaba.1997.30-533

Sidman, M., & Fletcher, F. G. (1968). A demonstration of auto-shaping with monkeys. *Journal of the Experimental Analysis of Behavior*, *11*(3), 307-309. https://doi.org/10.1901/jeab.1968.11-307

Shillingsburg, M. A., Hollander, D. L., Yosick, R. N., Bowen, C., & Muskat, L. R. (2015). Stimulus-stimulus pairing to increase vocalizations in children with language delays: A review. *The Analysis of Verbal Behavior*, *31*(1), 215-235. https://doi.org/10.1007/s40616-015-0042-2

Skinner, B. F. (1957). *Verbal behavior.* AVB Press.

Smith, R., Michael, J., & Sundberg, M. L. (1996). Automatic reinforcement and automatic punishment in infant vocal behavior. *The Analysis of Verbal Behavior*, *13*(1), 39-48. http://dx.doi.org/10.1007/BF03392905

Stock, R. A., Schulze, K. A., & Mirenda, P. (2008). A comparison of stimulus-stimulus pairing, standard echoic training, and control procedures on the vocal behavior of children with autism. *The Analysis of Verbal Behavior*, *24*(1), 123-133. https://dx.doi.org/10.1007/BF03393061

Sundberg, M. L. (2008). *VB-MAPP: Verbal behavior milestones assessment and placement program.* AVB Press.

Sundberg, M. L., & Michael, J. (2001). The benefits of Skinner's analysis of verbal behavior for children with autism. *Behavior Modification*, *25*(5), 698-724. https://doi.org/10.1177/0145445501255003

Sundberg, M. L., Michael, J., Partington, J. W., & Sundberg, C. A. (1996). The role of automatic reinforcement in early language acquisition. *The Analysis of Verbal Behavior*, *13*(1), 21-37. http://dx.doi.org/10.1007/BF03392905

Sundberg, M. L., & Partington, J. W. (1998). *Teaching language to children with autism or other developmental disabilities.* AVB Press.

Tomblin, J. B., Records, N. L., Buckwaiter, P., Zhang, X., Smith, E., & O'Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *Journal of Speech, Language, and Hearing Research*, *40*(6), 1245-1260. https://doi.org/10.1044/jslhr.4006.1245

Vouloumanos, A., Druhen, M. J., Hauser, M. D. & Huizink, A. T. (2009). Five-month-old infants' identification of the sources of vocalizations. *Psychological and Cognitive Sciences*, *106*(4), 18867-18872. https://doi.org/10.1073/pnas.0906049106

Wilcove, W. G., & Miller, J. C. (1974). CS-UCS presentations and a lever: Human autoshaping. *Journal of Experimental Psychology*, *103*(5), 868-877. https://doi.org/10.1037/h0037388

Wu, Z., & Gros-Louis, J. (2017). The value of vocalizing: 10-month-olds' vocal usage relates to language outcomes at 15 months. *Infancy*, *22*(1), 23-41. https://doi.org/10.1111/infa.12150

Yoon, S., & Bennett, G. M. (2000). Effects of a stimulus-stimulus pairing procedure on conditioning vocal sounds as reinforcers. *The Analysis of Verbal Behavior*, *17*(1), 75-88. https://doi.org/10.1007/BF03392957

Yoon, S., & Feliciano, G. M. (2007). Stimulus-stimulus pairing and subsequent mand acquisition of children with various levels of verbal repertoires. *The Analysis of Verbal Behavior*, *23*(1), 3-16. https://doi.org/10.1007/BF03393043

Table A1.

Average ITI and ISI Values Across Pairing Trials and Sessions for Participants in the 20-s ITI Group

Trial	ISI	ITI	Trial	ISI	ITI
1	3	15	31	5	25
2	5	25	32	4	20
3	4	20	33	3	15
4	4	20	34	5	20
5	5	25	35	4	20
6	5	25	36	3	15
7	5	25	37	5	25
8	4	20	38	5	25
9	4	20	39	5	25
10	5	25	40	4	20
11	3	15	41	3	15
12	5	25	42	3	15
13	4	20	43	5	25
14	5	25	44	4	20
15	5	25	45	3	15
16	3	15	46	5	25
17	4	20	47	5	25
18	5	25	48	4	20
19	3	15	49	5	25
20	3	15	50	3	15
21	4	20	51	5	25
22	5	25	52	4	20
23	5	25	53	4	20
24	5	25	54	5	25
25	3	15	55	5	25
26	4	20	56	5	25
27	5	25	57	4	20
28	3	15	58	4	20
29	4	20	59	5	25
30	5	25	60	3	15

Note: ITI duration averaged 20 s (range: 15-25 s) and the ISI duration averaged 3 s ISI (range: 3-5 s).

Table A2.

Average ITI and ISI Values Across Pairing Trials and Sessions for Participants in the 30-s ITI Group

Trial	ISI	ITI	Trial	ISI	ITI
1	6	30	31	6	30
2	5	25	32	5	25
3	6	30	33	6	30
4	7	35	34	7	35
5	7	35	35	6	30
6	4	20	36	7	35
7	6	30	37	4	20
8	4	20	38	5	25
9	7	35	39	5	25
10	8	40	40	7	35
11	4	20	41	7	35
12	7	35	42	5	25
13	6	30	43	7	35
14	4	20	44	5	25
15	5	25	45	6	30
16	6	30	46	5	25
17	5	25	47	4	20
18	7	35	48	6	30
19	5	25	49	7	35
20	7	35	50	4	20
21	7	35	51	8	40
22	5	25	52	7	35
23	5	25	53	4	20
24	4	20	54	6	30
25	7	35	55	4	20
26	6	30	56	7	35
27	7	35	57	7	35
28	6	30	58	6	30
29	5	25	59	5	25
30	6	30	60	6	30

Note: ITI duration averaged 30 s (range: 20-40 s) and the ISI duration average was 6 s (range: 4-8 s).

Table A3.

Average ITI and ISI Values Across Pairing Trials and Sessions for Participants in the 60-s ITI Group

Trial	ISI	ITI	Trial	ISI	ITI	Trial	ISI	ITI
1	10	50	21	16	80	41	10	50
2	13	65	22	15	75	42	13	65
3	15	75	23	13	65	43	15	75
4	10	50	24	10	50	44	10	50
5	11	55	25	12	60	45	11	55
6	13	65	26	13	65	46	13	65
7	10	50	27	11	55	47	10	50
8	16	80	28	10	50	48	16	80
9	11	55	29	11	55	49	11	55
10	10	50	30	16	80	50	10	50
11	16	80	31	10	50	51	16	80
12	11	55	32	11	55	52	11	55
13	10	50	33	16	80	53	10	50
14	11	55	34	10	50	54	11	55
15	13	65	35	13	65	55	13	65
16	12	60	36	11	55	56	12	60
17	10	50	37	10	50	57	10	50
18	13	65	38	15	75	58	13	65
19	15	75	39	13	65	59	15	75
20	16	80	40	10	50	60	16	80

Note: ITI duration average was 60 s (range: 40-80 s) and the ISI duration average was 12 s (range: 8-16 s).