DISCUSSION AND REVIEW PAPER

Inducing Novel Vocalizations by Conditioning Speech Sounds as Reinforcers

Anna Ingeborg Petursdottir[1] · Tracy L. Lepper[2]

Abstract Automatic reinforcement by response-produced auditory stimuli has long been hypothesized to play a role in the vocal development of infants. Clinicians and researchers have translated this idea into stimulus pairing interventions intended to increase novel vocalizations of nonverbal children with autism and other developmental disabilities by conditioning speech sounds as reinforcers. A number of studies have demonstrated positive effects of stimulus pairing procedures, but negative results have also been reported in the literature. This article provides a brief review of the existing literature on stimulus pairing procedures and a discussion of alternative procedures that may serve to establish speech sounds as reinforcers. Directions for future research are discussed and recommendations provided to clinicians.

Keywords Autism · Conditioned reinforcement · Stimulus-stimulus pairing · Vocalizations

Communication impairments are among the defining features of autism spectrum disorder (ASD; American Psychiatric Association 2013). The degree of impairment is variable across individuals (Kjelgaard & Tager-Flusberg 2001), but at one extreme, vocal language may be absent and other forms of communication minimal. Language interventions for these individuals often make use of alternative communication systems such as the Picture Exchange Communication System (PECS; Bondy and Frost 1994), which does not require vocal production. Alongside the use of such systems, efforts may be made to establish beginning vocal speech, especially for young children. However, this goal can be very difficult to accomplish for children who do not readily echo the speech of others and perhaps do not vocalize much in general.

The purpose of this article is to discuss the role that the reinforcing value of speech sounds may play in behavioral language interventions for nonverbal children with ASD. We will provide an overview of research on stimulus pairing procedures that have been investigated as a potential tool for inducing new vocalizations by conditioning speech sounds as reinforcers and discuss related alternative procedures that may produce the same outcome, perhaps more reliably. Finally, we will discuss issues of clinical relevance that remain to be investigated and provide tentative recommendations for clinicians who use or have considered using stimulus pairing procedures. Although the literature has not yet reached a sufficient stage of maturity to afford firm best-practice recommendations, stimulus pairing procedures have been disseminated in early intervention texts (e.g., Sundberg and Partington 1998) and workshops for quite some time. As a result, clinicians may be able to benefit from an overview of what is known and what remains to be investigated.

The Role of Reinforcement in Vocal Development

Children with ASD show delays in vocal development as infants. A major milestone in typical vocal development is the onset of canonical babbling. Canonical babbling refers to the emission of syllables that contain both consonants and

✉ Anna Ingeborg Petursdottir
a.petursdottir@tcu.edu

[1] Department of Psychology, Texas Christian University, Box 298920, Fort Worth, TX 76129, USA

[2] McNeese State University, 4205 Ryan St, Lake Charles, LA 70605, USA

vowels, with rapid transition between consonant and vowel as in adult speech. Retrospective analyses of home videos of infants and toddlers later diagnosed with ASD reveal lower rates of vocalizations and lower rates of canonical babbling than in typically developing infants (Patten et al. 2014). For children who do not communicate vocally, low rates of speech sound production may be an obstacle to successful intervention, as the clinician may have difficulty identifying specific speech sounds that occur with sufficient regularity to be reinforced and brought under stimulus control.

For typically developing infants, the production of canonical syllables appears sensitive to reinforcement by attention and social interaction (e.g., Goldstein, Schwade, and Bornstein 2009; Goldstein, West, and King 2003; Pelaez, Virues-Ortega, and Gewirtz 2011; Rheingold, Gewirtz, and Ross 1959). It is possible that lack of sensitivity to social attention as a reinforcer is, in part, responsible for delays in the vocal development of children with ASD. However, social attention may not be the only type of reinforcement involved in vocal development. It has long been hypothesized (e.g., Skinner 1957) that an automatic reinforcement process also operates in the acquisition of human babbling. Specifically, it has been proposed that the speech sounds of caregivers function as reinforcers for infants, such that when infants hear themselves produce sounds that resemble the sounds that caregivers make, the production of those sounds is reinforced. If so, it might explain why infants babble more when they are playing by themselves than when they are interacting with adults (Harold and Barlow 2013), why delays in babbling among hearing-impaired infants are inversely related to how well they hear with amplification (e.g., Bass-Ringdahl 2010; Hapsburg and Davis 2006), and why infants model the types of sounds that they produce after the types of sounds that their parents make when responding to them (Goldstein and Schwade 2008). It may be speculated, then, that in ASD, failure of speech sounds to function as reinforcers is another potential source of delayed vocal development.

This analysis suggests two possible ways to increase speech sound production in nonverbal children with ASD. First, if social attention does not suffice to reinforce speech sounds, perhaps other reinforcers do. Using highly preferred tangible items to reinforce speech sounds is, of course, an approach that has long been employed in efforts to establish vocal communication. For example, Lovaas (2003) described a seven-phase program for establishing vocal imitation, in which the first phase simply consists of delivering edible or other powerful reinforcers immediately following all vocalizations, whether they be speech sounds or other types of vocalizations, such as grunting or laughing. Once the student vocalizes with sufficient frequency, the second phase focuses on bringing vocalizations under the control of antecedent stimuli (e.g., therapist vocalizations). In the third phase, prompts and shaping are used to promote a match between

the antecedent stimulus and the response (i.e., establishing an echoic). As Lovaas acknowledged, however, a commonly encountered problem is that the child vocalizes at such low rates to begin with that opportunities to reinforce are few and far between, making it difficult to progress through these initial phases. In addition, vocal behavior is difficult to prompt, and shaping requires a great deal of therapist skill.

If a part of the problem is that speech sounds do not function as reinforcers for the child, such that automatic reinforcement does not occur when the child hears herself vocalize, a second approach to increasing speech sound production may be to condition speech sounds as reinforcers. This idea has been promoted in the early intervention literature, for example, by Sundberg and Partington (1998), and investigated in a number of research studies involving a procedure often referred to as stimulus-stimulus pairing (e.g., Miguel, Carr, and Michael 2002). This procedure is easy to implement and, if it is successful, results in the child beginning to emit new speech sounds without a need for prompting or shaping procedures.

Pairing Speech Sounds with Reinforcers

In stimulus-stimulus pairing, an adult repeatedly vocalizes a specific speech sound, followed immediately by the presentation of a preferred item or activity. For example, the adult says "bah" and then immediately delivers a highly preferred edible to the child. No response is required by the child in order to receive the preferred item, except that an orienting response (e.g., looking at the therapist) may be required before the therapist begins to present the target speech sound. Procedurally, stimulus-stimulus pairing resembles a respondent conditioning procedure (delay or trace conditioning). However, the desired outcome is operant reinforcement, specifically, automatic reinforcement of child vocalizations that resemble the paired speech sound. The pairing between the speech sound and the food item is intended to establish the speech sound as a conditioned reinforcer. If the sound "bah" functions as a conditioned reinforcer, then child vocalizations that produce sounds similar to "bah" should increase in frequency. The idea is that when specific speech sounds increase in frequency as a result of the pairing procedure, the clinician can more easily catch and reinforce them and ultimately bring them under appropriate stimulus control as functional vocalizations (e.g., echoics or mands).

Several variations of the basic pairing procedure exist. A comprehensive review is beyond the scope of this article, but a few procedural variations will be mentioned here and are summarized in Table 1 (for more comprehensive reviews, see Shillingsburg, Hollander, Yosick, Bowen, and Muskat in press; Stock, Schulze, and Mirenda 2008). First, the number of times that the target speech sound has been presented each time that it is paired with the preferred item has varied across

Table 1 Published studies on stimulus-stimulus pairing, success rate, and examples of procedural variations

Study	N/positive outcome	Age (years)	Diagnosis	Sound presentations per trial	ITI	Vocal delivery	Types of preferred items	Use of preferred items
Carroll and Klatt (2008)	2/1	<2[a]	ASD	5	0 s	Not described	Food and toys	Mary: same two items in all sessions; Max: one participant-selected items (out of 5) per session
Esch et al. (2009)	3/3	2–5	ASD	3	5–30 s	Exaggerated prosody	Food and toys	Three items per participant; one to three items rotated across trials within each session
Esch et al. (2005)	3/0	6–8	ASD	3	0 s	Not described	Food and toys	Five items per participant; one to five items rotated across trials within each session
Lepper et al. (2013)	3/3	2–4	ASD	3	10–15 s	Not described	Food and toys	One item per participant; same item used in all sessions
Miguel et al. (2002)	3/2	3–5	ASD	5	0 s	Not described	Food	One participant-selected (out of 5) per session
Miliotis et al. (2012)	2/2	6–8	ASD	1 or 3[b]	5–20 s	Exaggerated prosody	Food and toys	Three participant-selected items (out of 6) rotated across trials within each session
Normand and Knoll (2006)	1/0	3	ASD	7	Not stated	Not described	Not stated	Selected by participant prior to each session
Rader et al. (2014)	3/2	4–7	ASD	3	5–30 s	Exaggerated prosody	Not stated	One participant-selected (out of 3) per session
Sundberg et al. (1996)	5/5	2–4	ASD (2), other DD[a] (2), TD[b] (1),	1	Not stated (15 trials per min)[c]	Varied pitch and intonation	Social stimuli, e.g., tickles, praise, physical play	Varied across trials
Smith et al. (1996)	2/2	<2[a]	TD	1	Not stated	Not described	Bubbles, tickles	Not described
Stock et al. (2008)	3/1	2–4	ASD	5	20 s	Monotone presentation	Food	One participant-selected (out of 5) per session
Yoon and Bennett (2000)	3/3	3–4	Severe developmental delays	1	Not stated (12 trials per min)	Not described	Physical interaction	Physical interaction used in all sessions; details on variations not provided
Yoon and Feliciano (2007)	6/3	2–5	ED	1	Not stated (12 trials per min)	Not described	Physical interaction, activities, tangible items, food	Details not provided

ASD autism spectrum disorder, *TD* typically developing, *DD* developmental disability, *ED* classified with educational disabilities, *ITI* intertrial interval; defined here as the period between the end of reinforcer consumption until the next sound presentation.

[a] 22 and 23 months in Carroll and Klatt (2008); 11 and 13 months in Smith et al. (1996).

[b] Effect was greater with one presentation than three

[c] Refers to experiment 1; variable densities were used with one participant in experiment 2

studies. In some studies (e.g., Sundberg, Michael, Partington, and Sundberg 1996; Smith, Michael, and Sundberg 1996; Yoon and Bennett 2000), it has been presented only once (e.g., "bah"), whereas in others (e.g., Miguel, Carr, and Michael 2002; Esch, Carr, and Michael 2005; Esch, Carr, and Grow 2009; Normand and Knoll 2006), it has been presented repeatedly (e.g., "bah-bah-bah" followed by the delivery of the preferred item). Likewise, the density of stimulus presentations (i.e., the length of the intertrial interval) has varied across studies. In some studies (Esch et al. 2009; Miliotis et al. 2012; Rader et al. 2014), the therapist has presented all speech sounds with exaggerated prosody (i.e., "motherese"). By contrast, Sundberg et al. (1996) reported varying pitch and intonation across stimulus presentations, and Stock, Schulze, and Mirenda (2008) presented speech sounds in "a monotone fashion with no facial expression, emotional affect, or voice inflection" (p. 127). Some studies have used edible reinforcers exclusively (e.g., Miguel et al. 2002), others have used both edible and nonedible tangible reinforcers (e.g., Esch et al. 2009), and yet, others have included social reinforcers, such as tickles and praise (e.g., Sundberg et al. 1996). A single highly preferred item has been used for all pairings with speech sounds in some studies (Lepper, Petursdottir, and Esch 2013), whereas in others, the preferred item used has varied across sessions as a result of pre-session preference assessments (e.g., Esch et al. 2009; Miguel et al. 2002; Rader et al. 2014), and in yet others, multiple preferred items have been rotated across stimulus presentations within sessions (Esch et al. 2005). Most of these procedural variations have not been evaluated systematically (for an exception, see Miliotis et al. 2012; discussed below). Finally, there has been variation across studies in terms of experimental design and control for variables other than sound-reinforcer pairings that may affect child vocalizations.

The rationale for using stimulus-stimulus pairing to induce new vocalizations is conceptually sound, but the results reported in this literature have been mixed. A number of studies have reported positive effects on the vocalizations of children with and without ASD diagnoses (Carroll and Klatt 2008; Esch et al. 2009; Lepper et al. 2013; Miguel et al. 2002; Miliotis et al. 2012; Rader et al. 2014; Sundberg et al. 1996; Smith et al. 1996; Stock et al. 2008; Yoon and Bennett 2000; Yoon and Feliciano 2007). Some of these studies, however, have also reported negative results for one or more participants diagnosed with ASD or other disabilities (Carroll and Klatt 2008; Miguel et al. 2002; Rader et al. 2014; Stock et al. 2008; Yoon and Feliciano 2007), and two studies with a total of four participants with ASD have completely failed to find an effect (Esch, Carr, and Michael 2005; Normand and Knoll 2006). Overall, in 13 published studies (see Table 1), stimulus-stimulus pairing has produced an effect on the vocalizations of 27 of 39 participants (69 %), including 16 out of 25 (64 %) participants with ASD diagnoses. Thus, about a third

of participants with ASD have failed to respond to the intervention.

At this time, it is unclear if successful applications are associated with particular participant characteristics or any of the procedural variations that have appeared in the literature. Based on a review of the literature to date, Stock et al. (2008) concluded that stimulus-stimulus pairing might be most likely to produce an effect for very young children (2–3 years old, as opposed to 5 years or older), when the density of stimulus pairings was high, when the target sound was presented only once per pairing, and when social reinforcers were used. However, these conclusions were based on a limited number of studies, and it is unclear if they still hold in light of later additions to the empirical database. Rader et al. obtained positive results for two participants over the age of 5 years and negative results for a younger participant, and Miliotis et al. (2012) obtained clear effects for an 8-year-old and a 6-year-old. Uniformly positive results have been obtained in at least two studies in which intertrial intervals ranged from 5 to 30 s (Esch et al. 2009) and 10 to 15 s (Lepper et al. 2013), producing a relatively low density of pairings, and in which no social reinforcers were used. Miliotis et al. (2012) found support for presenting the target sound only once per trial compared to presenting it three times, but interpretation of these data may be complicated by unequal baseline rates of sounds assigned to different experimental conditions.

Stimulus pairing may fail to produce an effect on vocalizations because it fails to establish the paired speech sounds as reinforcers or because of failures of vocal production (i.e., the participants may be unable to articulate sounds that resemble the paired sounds). Petursdottir, Carp, Matthies, and Esch (2011) evaluated the effects of stimulus-stimulus pairing on the rate of button pressing of children with ASD diagnoses. One button produced a target sound that had just been paired with preferred items, and another button produced a nontarget sound that had been presented repeatedly by itself. By allowing the participants to produce speech sounds by pressing buttons instead of producing them vocally, it was possible to rule out articulation problems as a reason for conditioning failures. No participant showed a preference for pressing the button that produced the target sound over the nontarget sound, and none of several procedural modifications produced a reliable effect. The results were limited in that only a few procedural modifications were evaluated, and two of the three participants already had some vocal speech. Nevertheless, they suggest that stimulus pairing may fail to increase children's speech sound production even when they are not required to vocalize to produce the sounds. Thus, it seems possible that the treatment failures reported in the literature are at least, in part, related to failures of stimulus pairing to establish the target speech sounds as conditioned reinforcers. It is possible that these failures stem from idiosyncratic reasons, such as failures to identify sufficiently powerful primary reinforcers

for particular participants. However, it is also possible that there is a problem with the basic approach of pairing speech sounds with reinforcers without a response requirement. In fact, recent research suggests that other methods for conditioning reinforcers may produce more reliable effects.

Other Procedures for Establishing Conditioned Reinforcers

One way to establish a conditioned reinforcer is through discrimination training that establishes a previously neutral stimulus as a discriminative stimulus (S^D) for reinforcement. When the S^D is later presented contingent on a response, a reinforcement effect is seen. This approach was used in some early laboratory studies on conditioned reinforcement with children with and without disabilities (e.g., Lovaas et al. 1966; Sidowski, Kass, and Wilson 1965), and its effects have recently been demonstrated with children diagnosed with ASD (Holth et al. 2009; Isaksen and Holth 2009; Taylor-Santa, Sidener, Carr, and Reeve 2014). Isaksen and Holth (2009) compared the rate of arbitrary responses that produced either the S^D from a discrimination training procedure or the paired stimulus from a stimulus pairing procedure. Seven of eight participants responded at higher rates for the S^D, suggesting that discrimination training was more successful than the pairing procedure.

The results of Isaksen and Holth (2009) led Lepper et al. (2013) to compare the effects of discrimination training with the effects of stimulus pairing on the vocalizations of three nonverbal 4-year-old children diagnosed with ASD. In discrimination training, when a target speech sound that served as the S^D was presented, the participant had to make an arbitrary motor response (arm raise) in order to receive access to a preferred item. Arm raising that followed presentations of a nontarget speech sound or occurred at other times did not produce the preferred item. In the stimulus pairing condition, a target sound was paired with the delivery of a preferred item and a nontarget sound was presented by itself. Both procedures produced an effect on the vocalizations of all participants, as they vocalized target sounds at greater rates than nontarget sounds and also at greater rates than sounds presented in a control condition. However, there was no advantage of the discrimination training procedure, as both procedures produced approximately equal rates of vocalizations. All participants preferred the discrimination training procedure to the stimulus pairing procedure, which might speak in favor of the former. However, discrimination training was more cumbersome to implement due to the need for prompting and prompt fading to bring the arm-raising response under the control of the S^D. Future research might address whether discrimination training produces an effect on the vocalizations of

children who fail to vocalize as a result of stimulus-stimulus pairing.

In early research on conditioned reinforcement in humans, when stimulus pairing procedures were used instead of discrimination training procedures, pairings of neutral stimuli with reinforcers were commonly contingent on a particular response by the participant. For example, in a study by Myers and Myers (1963), pressing the nose of a toy clown was followed by the presentation of a light or a buzzer (the neutral stimuli) and M&Ms®. As another example, Dorow (1980) paired verbal approval and music (the neutral stimuli) with food contingent on complying with an instruction to hand a bead to the experimenter. A difference between the two studies is that Myers and Myers used a free-operant procedure in which the participants could respond at any time to obtain the paired stimuli, whereas Dorow used a discrete trial procedure in which opportunities to respond were initiated by the experimenter. However, both can be considered examples of what will here be termed *response-contingent* stimulus pairing, in which the participant needs to make a particular response to access the paired stimuli. As a more recent example involving speech as the initially neutral stimulus, participants in a study by Greer, Pistoljevic, Cahill, and Du (2011) pressed a button that produced the voice of a familiar person telling a story. Whenever the button was depressed and the voice recordings were playing, reinforcers were delivered. Among other findings, participants were more likely to opt to listen to stories in a different setting following the conditioning procedure than before.

In contrast to these studies, participants in most studies on stimulus pairing effects on vocalizations have not been required to make any kind of a response to produce presentations of speech sounds and reinforcers; thus, stimulus pairings have been *response-independent*. Possible exceptions are three studies (Esch et al. 2009; Miliotis et al. 2012; Rader et al. 2014) in which pairings between speech sounds and reinforcers were presented only after the participant responded to a prompt (termed an observing prompt) to orient toward the experimenter. Although these studies found positive effects on vocalizations for all but one participant, it is unknown if these effects can in any part be attributed to the observing prompt and the subsequent orienting response.

Dozier et al. (2012) evaluated the effects of a response-independent pairing procedure and a response-contingent pairing procedure (termed stimulus-response pairing in that study) on the conditioned reinforcing value of praise for individuals with intellectual disabilities. Response-independent pairing did not successfully establish praise as a reinforcer for any of the four participants in the first experiment, whereas response-contingent pairing was effective for four of eight participants in a second experiment. This finding brings up the possibility that response-contingent pairing procedures involving speech sounds and reinforcers may produce greater

effects on vocalizations than response-independent procedures.

In the three published studies to date in which pairings between speech sounds and reinforcers were contingent on the child looking at the experimenter (Esch et al. 2009; Miliotis et al. 2012; Rader et al. 2014), the timing of stimulus presentations was nevertheless controlled by the experimenter, because the observing prompt (e.g., "look") initiated each opportunity for reinforcement. Recently, Lepper and Petursdottir (under review) evaluated the effects of a response-contingent pairing procedure in which the participants could initiate stimulus presentations by pressing a button. Pressing the button sometimes produced a target speech sound that was paired with a reinforcer and, sometimes, a nontarget speech sound that was not followed by reinforcer delivery. The button was available throughout the entire session with the exception of a 10-s period that followed each speech sound presentation and consumption of the reinforcer, if applicable. In experiment 1, this procedure was compared to a response-independent pairing procedure that, nevertheless, included a prompt ("look") to orient toward the experimenter before speech sounds were presented. The response-independent condition also involved a target sound that was paired with a reinforcer (the same reinforcer as in the response-contingent condition) and a nontarget sound that was not followed by reinforcer delivery. The timing of all sound presentations in response-independent sessions was yoked to a previous response-contingent pairing session, as the two types of sessions were alternated in a multielement design. Three boys diagnosed with ASD, between the ages of 4 and 6 years, participated. In the response-contingent pairing condition, all three participants vocalized the target speech sounds at higher rates than the nontarget sounds. By contrast, an effect of the response-independent pairing condition was seen for only one of the three participants, and in that case, the effect was much smaller than the effect of response-contingent pairing. Subsequently, response-contingent pairing was applied to the target speech sounds that had previously been included in response-independent pairing condition, and response-contingent pairing produced increases in the rate of these sounds for all three participants. These data suggest that giving the child an opportunity to initiate the pairing between a speech sound and a reinforcer may produce more reliable effects on vocalizations than when the therapist initiates the pairing.

Future Directions

Researchers should continue to explore the effects of discrimination training and response-contingent pairing as alternatives to the response-independent stimulus pairing procedure that have predominated in the literature on establishing speech

sounds as reinforcers. More research is also needed on several aspects of these procedures in general; for example, the selection of speech sounds to establish as conditioned reinforcers, the selection of reinforcers to pair them with, and the number of repetitions of the speech sound per presentation. In addition, the clinical utility of inducing speech sounds through these procedures needs further investigation. To date, only a few studies have included demonstrations that the newly induced speech sounds can be "caught" and strengthened further via direct reinforcement (Esch et al. 2009; Lepper and Petursdottir under review) or brought under stimulus control as functional vocalizations (Carroll and Klatt 2008; Yoon and Feliciano 2007).

Future research might also address the possibility of better matching stimulus pairing procedures to the clinical goals of the intervention. In the empirical literature to date, it has been typical to conduct repeated pairings of one specific sound with reinforcers and, when an effect is seen, start pairing another sound. Typically, different speech sounds have been paired with the same reinforcer or the same set of reinforcers. Although the existing data show that this arrangement may produce an effect on vocalizations, other alternatives exist that might be more relevant to the clinical goals of stimulus pairing but have not yet been evaluated empirically. The clinical goals of stimulus pairing may differ across individuals, as illustrated in the following two examples.

First, a 2-year-old child has just started an early intervention program. The child is nonverbal and communicates primarily by crying and tugging at people. He engages in highly variable vocal play but does not say any recognizable words and does not attempt to echo speech or other sounds. For this child, establishing a rudimentary mand repertoire would be a priority. The high level of vocalizations suggests that it may be feasible to establish vocal mands, but the lack of an echoic repertoire is an impediment, so a stimulus pairing procedure might be considered to induce specific vocalizations that could be reinforced as the child's initial mands. In the mand relation, a specific response form is reinforced with a specific consequence, so it might make sense to anticipate that relation during the pairing procedure by pairing specific speech sounds with specific reinforces. For example, "mmm" might be paired with milk, "kuh" with a cookie, and "buh" with a ball. However, while this arrangement may sound intuitive, it has not been evaluated empirically.

In the second example, a 4-year-old child has been enrolled in an early intervention program for a year. He has acquired a fairly extensive mand and tact repertoire using PECS but does not communicate vocally. Levels of noncommunicative vocalizations are low and include few recognizable phonemes, and he echoes only one or two specific sounds. If stimulus pairing were considered for this child, its immediate goal might not be to establish vocal mands, as the child already has a mand repertoire. Instead, the goal of stimulus pairing

might be to increase the range of speech sounds that the child produces in preparation of efforts to establish echoic responding. Here, it seems possible that instead of pairing only one specific speech sounds at a time, it might be worthwhile to explore pairing multiple exemplars of particular syllable structures. For example, the therapist might pair consonant-vowel combinations with reinforcers but vary the specific consonant-vowel combination across pairings (e.g., "bah", "moo," "lee") in hopes of observing a general increase in the child's emission of consonant-vowel combinations (see Goldstein and Schwade 2008). Again, this is an idea that has not yet been evaluated empirically but could potentially have clinical utility.

Recommendations for Practitioners

What can clinicians who work with nonverbal clients take from the research that has been described in this article? Some clinicians may be skeptical about the utility of stimulus pairing procedures, given the mixed results reported in the literature. However, the existing data clearly suggest that procedures intended to establish speech sounds as reinforcers have the potential to produce meaningful positive outcomes for at least some children who have not been able to communicate vocally. At this time, little is known about the conditions under which these procedures will produce reliable effects. As a result, we do not recommend investing much time in this approach if initial efforts fail, at least until more data are available. However, it may be at least worth trying one or more of the procedural variations described in previous research.

As noted earlier, more research is needed on how to select target speech sounds and how to select reinforcers to pair them with. It is probably important to select target sounds that fall within the range of the child's vocal developmental level. For example, for a child who primarily emits vowel sounds and single-syllable consonant-vowel combinations, a two-syllable target sound might be unlikely to emerge as a result of a stimulus pairing procedure. One approach is to select sounds that the child already has been observed to emit, but very infrequently (Esch et al. 2009); for the child in the previous example, this might include selecting vowels or consonant-vowel syllables that occur at very low rates. Another approach may be to select sounds that the child has not been observed to emit, but the minimal units of which are clearly in the child's vocal repertoire. For example, Lepper et al. (2013) and Lepper and Petursdottir (under review) targeted consonant-vowel combinations that met the criterion that (a) the target consonant-vowel combination never occurred during extended free-play observations, but (b) the component consonant and the component vowel occurred within other syllables. For example, if a participant was observed to emit the syllables "ma" and "boo" during observation, but never emitted the

syllable moo, then "moo" could be selected for inclusion in the experimental evaluation. This approach was apparently successful as all participants ultimately started producing each of two target syllables. Another consideration to keep in mind is that if the goal is to induce speech sounds that will subsequently be reinforced as mands, it may be useful to select target sounds that approximate the conventional names of the child's preferred items. For example, if a child's most highly preferred items include chips, syllables like "cha" or "dip" might be chosen over irrelevant syllables like "ga."

Validated preference assessments are recommended for selecting reinforcing items to pair with target speech sounds. Positive effects of stimulus pairing procedures have been demonstrated when speech sounds have been paired with highly preferred food items (e.g., Miguel et al. 2002), toys (e.g., Lepper et al. 2013), and social reinforcers (e.g., Sundberg et al. 1996), so the type of reinforcer may not be crucial as long as it is highly preferred. As noted earlier, more research is needed on other aspects of reinforcer selection. In some of our own studies (Lepper et al. 2013; Lepper and Petursdottir under review), the same reinforcer was used in all sessions and paired with at least two different speech sounds. This approach, however, was used primarily for experimental control purposes. That is, the purpose of these studies was to compare two or more procedures, and it was important to equate the quality and quantity of reinforcement across conditions. Clinically, one would likely be more inclined to conduct a preference assessment before each session to select a preferred item for use in that session, an approach that has been used in a number of successful studies (e.g., Esch et al. 2009; Rader et al. 2014). Alternatively, as noted earlier, it may be worthwhile to explore the pairing of specific speech sounds with specific reinforcers, especially if the immediate clinical goal is to establish vocal mands. However, this practice has not yet been evaluated empirically.

Once target speech sounds and reinforcers have been identified, the next step is to select a conditioning procedure. In this article, it has been suggested that response-independent pairing procedures may not be the most effective means of imbuing speech sounds with reinforcing properties. However, research on alternative procedures is still in its early stages. The results of Lepper and Petursdottir (under review) may tentatively suggest a benefit of making contiguous presentations of speech sounds and reinforcers contingent on particular responses initiated by the child, but more research is needed on this procedure, including the selection of an appropriate response. The response used in Lepper and Petursdottir (under review) was a button press, but this might be seen as an awkward response to use clinically. Other possibilities might include presenting stimuli contingent on unsolicited eye contact with the experimenter or engaging in a nonvocal communication response. Although these responses have not been evaluated directly in the context of the research

discussed in this article, it may be noted that in several of studies on mand instruction using alternative communication systems, an increase in vocalizations has accompanied the acquisition of PECS (e.g., Charlop-Christy, Carpenter, Le, LeBlanc, and Kellet 2002; Ganz and Simpson 2004; Greenberg, Tomaino, and Charlop 2014; Tincani, Crozier, and Alazetta 2006) or signed mands (Tincani 2004). It may be speculated that a potential source of this effect is the pairing between adult speech and a reinforcer that typically occurs when a mand is reinforced (e.g., the child hands a picture of a cookie to the instructor, who responds by saying "cookie" while delivering a piece of cookie to the child). As such, it may exemplify a type of response-contingent stimulus pairing. However, this effect has not occurred in all studies (e.g., Ganz, Simpson, and Corbin-Newsome 2008; Howlin, Gordon, Pasco, Wade, and Charman 2007) and more research is needed on the conditions under which it occurs and how it may be related to the stimulus pairing component of the instructional procedures.

Although the effects of response-independent stimulus pairing procedures have been inconsistent across studies, positive effects have nevertheless been demonstrated repeatedly, so clinicians might consider choosing this procedure, given the dearth of research on response-contingent procedures. When using response-independent pairing, a tentative recommendation can be provided to prompt an orientation response before stimulus presentation, as described in several recent studies (Esch et al. 2009; Miliotis et al. 2012; Rader et al. 2014). If response-independent pairing does not produce an effect, clinicians might consider switching to either response-contingent pairing or a discrimination training procedure (Lepper et al. 2013).

Once the conditioning procedure has been selected, how frequently and how many times does the target sound need to be paired with a reinforcer? When response-contingent pairing is used, the density of pairings is determined by the child's responses; however, when a response-independent procedure or discrimination training is used, the clinician must determine the appropriate spacing of pairings. The literature offers little guidance in this respect due to a lack of systematic evaluations. However, most studies have paired the target sound with reinforcers several times per minute for a total of ten or more pairings per session. As for how long to continue, an effect usually appears to be evident within one to five sessions (Esch et al. 2009; Lepper et al. 2013; Miguel et al. 2002; Miliotis et al. 2012; Rader et al. 2014) so it may not be fruitful to persist much longer than that if vocalizations do not increase. Another question concerns whether to present each sound only once (e.g., "bah") or repeatedly (e.g., "bah-bah-bah") each time it is paired with a reinforcer. As noted previously, the results of Miliotis et al. (2012) may tentatively speak in favor of only one presentation per pairing, but additional research is needed.

Finally, clinicians who familiarize themselves with the stimulus-stimulus pairing literature may observe that some research studies on this topic employ procedures that appear either countertherapeutic or clinically irrelevant. Potentially countertherapeutic practices include that child vocalizations of target speech sounds are typically not reinforced during stimulus pairing sessions, and in some studies, they have even produced a delay in the scheduled delivery of a highly preferred item. These procedures are necessary in research to demonstrate that the increase in vocalizations is attributable to the pairing procedure and not simply to direct reinforcement. However, they could potentially result in extinction or punishment of the target sounds, which may, in part, explain negative findings in some studies. Clinically, every occurrence of the target speech sounds should be reinforced as soon as it begins to appear. Also, a number of recent, successful studies have interspersed target speech sound presentations with presentations of a nontarget that is not followed by reinforcer delivery (Esch et al. 2009; Lepper et al. 2013; Miliotis et al. 2012; Rader et al. 2014), which at first sight may not seem clinically relevant. These target sounds are, in part, intended to function as control sounds to demonstrate that merely presenting a speech sound by itself is not sufficient to produce an effect. However, it is possible that the inclusion of the nontarget sounds contributes to the effects of stimulus pairing by increasing the salience of the target sound (for a discussion, see Esch et al. 2009). Although this possible role has not yet been verified, clinicians might consider interspersing stimulus pairings with nontarget sound presentations, given that doing so adds little complexity to the procedure.

In summary, attempting to condition speech sounds as reinforcers is a conceptually sound approach to induce novel vocalizations of nonverbal children that has a decent amount of empirical support. The negative results sometimes reported in the literature (e.g., Esch et al. 2005; Normand and Knoll 2006) may have caused clinicians to shy away from using these procedures. However, as procedures are refined and alternatives to response-independent pairing evaluated more fully, it may be possible to produce the effect more reliably and improve guidelines for maximizing clinical impact.

References

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.

Bass-Ringdahl, S. M. (2010). The relationship of audibility and the development of canonical babbling in young children with hearing impairment. *Journal of Deaf Studies and Deaf Education, 15*, 287–310.

Bondy, A. S., & Frost, L. A. (1994). The picture exchange communication system. *Focus on Autistic Behavior, 9*(3), 1–19.

Carroll, R. A., & Klatt, K. P. (2008). Using stimulus-stimulus pairing and direct reinforcement to teach vocal verbal behavior to young children with autism. *The Analysis of Verbal Behavior, 24*, 2008.

Charlop-Christy, M. H., Carpenter, M., Le, L., LeBlanc, L. A., & Kellet, K. (2002). Using the Picture Exchange Communication System (PECS) with children with autism: assessment of PECS acquisition, speech, social-communicative behavior and problem behavior. *Journal of Applied Behavior Analysis, 35*, 213–231.

Dorow, L. G. (1980). Generalization effects of newly conditioned reinforcers. *Education and Training of the Mentally Retarded, 15*, 8–14.

Dozier, C. L., Iwata, B. A., Thomason-Sassi, J., Worsdell, A. S., & Wilson, D. M. (2012). A comparison of two pairing procedures to establish praise as a reinforcer. *Journal of Applied Behavior Analysis, 45*, 721–735.

Esch, B. E., Carr, J. E., & Michael, J. (2005). Evaluation stimulus-stimulus pairing and direct reinforcement in the establishment of an echoic repertoire of children diagnosed with autism. *The Analysis of Verbal Behavior, 21*, 43–58.

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*, 225–241. doi:10.1901/jaba.2009.42-225.

Ganz, J. B., & Simpson, R. L. (2004). Effects on communicative requests and speech development of the picture exchange communication system in children with characteristics of autism. *Journal of Autism and Developmental Disorders, 34*, 395–409.

Ganz, J. B., Simpson, R. L., & Corbin-Newsome, J. (2008). The impact of the picture exchange communication system on requesting and speech development in preschoolers with autism spectrum disorders and similar characteristics. *Research in Autism Spectrum Disorders, 2*, 157–169.

Goldstein, M. H., & Schwade, J. A. (2008). Social feedback to infants' babbling facilitates rapid phonological learning. *Psychological Science, 19*, 515–523.

Goldstein, M. H., West, M. J., & King, A. P. (2003). Social shaping of babbling: a new parallel with bird song. *Proceedings of the National Academy of Sciences, 100*, 8050–8055.

Goldstein, M. H., Schwade, J. A., & Bornstein, M. H. (2009). The value of vocalizing: five-month-old infants associate their own noncry vocalizations with responses from caregivers. *Child Development, 80*, 636–644.

Greenberg, A. L., Tomaino, M. E., & Charlop, M. H. (2014). Adapting the picture exchange communication system to elicit vocalizations in children with autism. *Journal of Developmental and Physical Disabilities, 26*, 35–51.

Greer, R. D., Pistoljevic, N., Cahill, C., & Du, L. (2011). Effects of conditioning voices as reinforcers for listener responses on rate of learning, awareness, and preference for listening to stories in preschoolers with autism. *The Analysis of Verbal Behavior, 27*, 103–124.

Hapsburg, D., & Davis, B. L. (2006). Auditory sensitivity and the prelinguistic vocalizations of early-amplified infants. *Journal of Speech, Language, and Hearing Research, 49*, 809–822.

Harold, M. P., & Barlow, S. M. (2013). Effects of environmental stimulation on infant vocalizations and orofacial dynamics at the onset of canonical babbling. *Infant Behavior and Development, 36*, 84–93.

Holth, P., Vandbakk, M., Finstad, J., Gronnerud, E. M., & Akselsen-Sorensen, J. M. (2009). An operant analysis of joint attention and the establishment of conditioned social reinforcers. *European Journal of Behavior Analysis, 10*, 143–158.

Howlin, P., Gordon, R. K., Pasco, G., Wade, A., & Charman, T. (2007). The effectiveness of Picture Exchange Communication System (PECS) training for teachers of children with autism: a pragmatic, group randomized controlled trial. *Journal of Child Psychology and Psychiatry, 48*, 473–481.

Isaksen, J., & Holth, P. (2009). An operant approach to teaching joint attention skills to children with autism. *Behavioral Interventions, 24*, 215–236.

Kjelgaard, M. M., & Tager-Flusberg, H. (2001). An investigation of language impairment in autism: implications for genetic subgroups. *Language and Cognitive Processes, 16*, 287–308.

Lepper, T. L. & Petursdottir, A. I. (under review). Effects of response-contingent stimulus pairing on vocalizations of nonverbal children with autism.

Lepper, T. L., Petursdottir, A. I., & Esch, B. E. (2013). Effects of an operant discrimination training procedure on the vocalizations of nonverbal children with autism. *Journal of Applied Behavior Analysis, 46*, 656–661.

Lovaas, O. I. (2003). *Teaching individuals with developmental delays. Basic intervention techniques*. Austin: PRO-ED.

Lovaas, O. I., Freitag, G., Kinder, M. I., Rubenstein, B. D., Shaeffer, B., & Simmons, J. Q. (1966). Establishment of social reinforcers in two schizophrenic children on the basis of food. *Journal of Experimental Child Psychology, 4*, 109–125.

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*, 3–13.

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*, 809–813.

Myers, N. A., & Myers, J. L. (1963). Secondary reinforcement in children as a function of conditioning associations, extinction percentages, and stimulus types. *Journal of Experimental Psychology, 65*, 455–459.

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*, 81–85.

Patten, E., Belardi, K., Baranek, G. T., Watson, L. R., Labban, J. D., & Oller, D. K. (2014). Vocal patterns in infants with autism spectrum disorders: canonical babbling status and vocalization frequency. *Journal of Autism and Developmental Disorders, 44*, 2413–2428.

Pelaez, M., Virues-Ortega, J., & Gewirtz, J. L. (2011). Reinforcement of vocalizations through contingent vocal imitation. *Journal of Applied Behavior Analysis, 44*, 33–40.

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*, 45–60.

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*, 69–74.

Rheingold, H. L., Gewirtz, J. L., & Ross, H. W. (1959). Social conditioning of vocalizations in the infant. *Journal of Comparative and Physiological Psychology, 52*, 68–73.

Shillingsburg, A., Hollander, D. L., Yosick, R. N., Bowen, C., & Muskat, L. R. (in press). Stimulus-stimulus pairing to increase vocalizations in children with language delays: a review. *The Analysis of Verbal Behavior*.

Sidowski, J. B., Kass, N., & Wilson, H. (1965). Cue and secondary reinforcement effects with children. *Journal of Experimental Psychology, 69*, 340–342.

Skinner, B. F. (1957). *Verbal behavior*. New York: Appleton-Century-Crofts.

Smith, R., Michael, J., & Sundberg, M. L. (1996). Automatic reinforcement and automatic punishment in infant vocal behavior. *The Analysis of Verbal Behavior, 13*, 39–48.

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*, 123–133.

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*, 21–37.

Sundberg, M. L., & Partington, J. W. (1998). *Teaching language to children with autism or other developmental disabilities*. Pleasant Hill, CA: Behavior Analysts, Inc.

Taylor-Santa, C., Sidener, T. M., Carr, J. E., & Reeve, K. F. (2014). A discrimination training procedure to establish conditioned reinforcers for children with autism. *Behavioral Interventions, 29*, 157–176. doi:10.1002/bin.1384.

Tincani, M. (2004). Comparing the picture exchange communication system and sign language training for children with autism. *Focus on Autism and Other Developmental Disabilities, 19*, 152–163.

Tincani, M., Crozier, S., & Alazetta, L. (2006). The picture exchange communication system: effects on manding and speech development for school-aged children with autism. *Education and Training in Developmental Disabilities, 41*, 177–184.

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*, 75–88.

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*, 3–16.