The Role of Automatic Reinforcement in Early Speech Acquisition

Barbara E. Esch
Western Michigan University

Follow this and additional works at: http://scholarworks.wmich.edu/dissertations

Part of the Child Psychology Commons, Clinical Psychology Commons, and the Developmental Psychology Commons

Recommended Citation
http://scholarworks.wmich.edu/dissertations/856
THE ROLE OF AUTOMATIC REINFORCEMENT
IN EARLY SPEECH ACQUISITION

by

Barbara E. Esch

A Dissertation
Submitted to the
Faculty of The Graduate College
in partial fulfillment of the
requirements for the
Degree of Doctor of Philosophy
Department of Psychology
Adviser: James E. Carr, Ph.D.

Western Michigan University
Kalamazoo, Michigan
June 2007
ACKNOWLEDGMENTS

I am indebted to Dr. Mark Sundberg for his initial work in applying behavioral principles to the acquisition of early speech skills. The line of research he started in stimulus-stimulus pairing offered both hope and clinical direction to those of us searching for more effective ways to help children learn to speak. Mark has been a supportive colleague and friend who has contributed greatly to my professional growth and I am deeply appreciative.

Next, I would like to thank the members of my dissertation committee, Drs. Scott Gaynor, Jack Michael, David Palmer, and most especially my chair, Dr. Jim Carr. It was my extreme good fortune to be guided by individuals with such in-depth knowledge in the subject matter and such generosity of spirit. Special thanks to Dr. Michael, friend and neighbor extraordinaire, for bringing me into the Psychology program as his last graduate student before retirement and for our many Monday night post-retirement chats sharing olives, chocolates, and wine at my bistro table.

Finally, deep love and gratitude to my husband and fellow behavior analyst, Dr. John Esch, who set the best example ever of how to complete a Ph.D. while maintaining a normal life. His words and humor encouraged me all along this journey while his love kept the ground steady underneath.

Barbara E. Esch
<table>
<thead>
<tr>
<th>TABLE OF CONTENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENTS .......................................................... ii</td>
</tr>
<tr>
<td>LIST OF TABLES .................................................................. v</td>
</tr>
<tr>
<td>LIST OF FIGURES ................................................................ vi</td>
</tr>
<tr>
<td>INTRODUCTION ...................................................................... 1</td>
</tr>
<tr>
<td>Theories of Language Acquisition .................................... 1</td>
</tr>
<tr>
<td>Behavioral Interpretation of Language Development ........... 4</td>
</tr>
<tr>
<td>Role of Automatic Reinforcement in Early Speech ............... 8</td>
</tr>
<tr>
<td>Speech as a Conditioned Reinforcer ................................. 9</td>
</tr>
<tr>
<td>Focus of Current Investigation ......................................... 16</td>
</tr>
<tr>
<td>METHOD ............................................................................. 20</td>
</tr>
<tr>
<td>Participants and Setting ................................................. 20</td>
</tr>
<tr>
<td>Pre-experimental Assessments .......................................... 22</td>
</tr>
<tr>
<td>Stimulus Preference Assessment ....................................... 27</td>
</tr>
<tr>
<td>Target Responses and Response Definition ....................... 28</td>
</tr>
<tr>
<td>Interobserver Agreement ................................................. 30</td>
</tr>
<tr>
<td>Procedure ......................................................................... 31</td>
</tr>
<tr>
<td>Experimental design ....................................................... 31</td>
</tr>
<tr>
<td>Baseline ........................................................................... 32</td>
</tr>
<tr>
<td>Stimulus-stimulus pairings .............................................. 33</td>
</tr>
</tbody>
</table>
Table of Contents – Continued

Differential reinforcement ................................................................. 34
Withdrawal of differential reinforcement ........................................... 35
Caregiver training ............................................................................. 36
Independent Variable Integrity ............................................................. 37
RESULTS .............................................................................................. 38
DISCUSSION ....................................................................................... 44
Implications ....................................................................................... 45
Conceptual Issues ............................................................................ 45
Limitations .......................................................................................... 65
Future Research ................................................................................ 66
SUMMARY ............................................................................................ 69
APPENDICES .......................................................................................... 70
A. Pilot Study .................................................................................... 71
B. Research Protocol Clearance ......................................................... 84
REFERENCES ....................................................................................... 86
LIST OF TABLES

1. Speech and Language Assessment Scores ............................................................... 24
2. Stimulus-stimulus Pairing Procedure ....................................................................... 33
LIST OF FIGURES

1. Nonconcurrent multiple baseline for Joshua.......................................................... 39
2. Nonconcurrent multiple baseline for Madison...................................................... 41
3. Nonconcurrent multiple baseline for Daniel.......................................................... 43
INTRODUCTION

One of the most sophisticated and powerful interchanges humans engage in is communication. The intricacy of a particular exchange may not be evident in the formal characteristics of the interaction itself (e.g., hello/hi), but its analysis is fundamental to determining how such intercourse comes about. Across cultures, the linguistic complexities of native languages seem to require very little formal training for effective use (Bijou & Baer, 1965; Moerk, 1990; Schlinger, 1995). As yet, however, there is no consensus about the processes that account for language acquisition, including its apparent adaptivity under changing environmental requirements and the fluency with which a given linguistic response is made.

Theories of Language Acquisition

Some language acquisition theories appeal to internal structures and innate processes that largely disregard the behavior-shaping influence of an individual’s idiosyncratic experiences in a physical world and a verbal community. Chomsky (1959) proposed internal mechanisms, genetically coded, by which linguistic rules are internally derived from samples of speech. These native structures (hence, nativist theories) are theorized to reside in the nervous system, allowing all humans regardless of their particular language system (e.g., French, Russian) to acquire a universal grammar that governs language use. Others (Pinker & Bloom, 1990) have proposed related accounts in which mental life, or faculties, of which language is a sub-set,
occurs as information processing through organs of computation in a mind. Still others describe language development by emphasizing its structural aspects such as mean length of utterance (e.g., Bates, O’Connell, & Shore, 1987) or syntactic form (e.g., Gentner, 1982).

Language theories invoking mental states or a genetic apparatus in lieu of ontogenic effects on behavior offer no advantage over more parsimonious explanations of a particular individual’s experiences as a member of a unique verbal community. That is, language learners are exposed to naturally occurring contingencies whose descriptions (i.e., rules) do not require a priori explication or derivation in order to impose their effects. Our ability to describe any number of grammatical, syntactic, semantic, or environmental arrangements in no way assumes they require, or are a function of, internal states or structures. Furthermore, a technology does not yet exist to identify these hypothesized relations and processes or, more importantly, to measure the purported influence of these mechanisms. Therefore, their utility to explain language acquisition is difficult to ascertain since, even if they could be identified, an analysis of variables responsible for activating these mechanisms would perforce have to be made.

In contrast to cognitive theories of language development, Skinner (1957) provided a selectionist perspective of the apparent phenomenon (i.e., language acquisition) as a function of environmental variables operating according to behavioral processes in which discriminative stimuli, strengthening or weakening
consequences, and relevant motivative variables converge in particular arrangements. Verbal behavior is selected into the repertoire as a function of consequences provided by others who are specifically trained to do so (termed mediated or social reinforcement, Vaughan & Michael, 1982) or through stimuli resulting from one’s own actions (automatic reinforcement; Skinner, 1957, 1969; Vaughan & Michael). Accordingly, words and their meanings are not considered separate entities from the behavior of the speaker but instead are defined functionally according to the contingencies in which verbal responses occur. Moreover, formal aspects (i.e., topographies, grammar) of speaker behavior make no special contribution to its functional analysis. That we understand (in the sense of responding appropriately with respect to particular stimuli) given instances of linguistic constructions can be attributed to contingencies of reinforcement (mediated or automatic) and, in the case of new entities, in part to the behavioral process of generalization (Skinner, 1957) in which stimulus control for particular contingency relations is extended to other relations containing similar stimuli. Thus, a child whose behavioral history includes reinforcement for responding to get, ball, under, and table and who can tact box could respond appropriately to a novel instruction such as get the box that’s under the table even though responding to that instruction had never been reinforced. A behavioral account of language acquisition, therefore, need not credit our verbal behavior with meaning apart from that contained in a functional description of environmental influences that strengthen or weaken it.
Behavioral Interpretation of Language Development

There is already a robust empirical literature\(^1\) that validates the utility of socially mediated (i.e., indirect or *programmed*) reinforcement to establish a wide range of human as well as non-human behavior. Much of this literature reflects efforts to modify nonverbal behavior. In the past 20 years, however, behavioral researchers have increasingly investigated variables that control functional *verbal operants* (e.g., mands, tacts; Skinner, 1957) in an effort to explain acquisition and maintenance of behaviors that comprise complex human language. As with studies of nonverbal behavior, the efficacy of indirect reinforcement to strengthen verbal behavior is strongly demonstrated across a number of important functions, identified by distinct antecedent and consequent controlling variables (see Sautter & LeBlanc, 2006).

However, when one considers how rapidly and effortlessly most children learn their respective native languages, it is implausible that reinforcement of an indirect, socially mediated nature could be the sole explanation for behavioral repertoires acquired in such exponential leaps. Other influences must surely merit some consideration for this robust growth. In fact, the strengthening effect of stimuli that follow an individual’s behavior on that behavior is not restricted to mediation through another (Skinner). It likely occurs at times because particular responses themselves produce stimuli that serve to strengthen their occurrence. In other words, the

---

\(^1\) See *Journal of Experimental Analysis of Behavior* and *Journal of Applied Behavior Analysis* among other sources.
strengthening effect is self-produced; thus, the reinforcement is automatic (Vaughan & Michael, 1982).

Logistically, automatic reinforcement may adequately bear the burden of explaining a good deal of early language acquisition, and perhaps more complex repertoires as well (e.g., Palmer, 1996, 1998), by submitting a well-established theory (i.e., reinforcement) to the empirical process. However, investigations should not expect to identify new process effects exclusive to self-generated consequences. That is, when stimuli arise automatically from a response, their influence on future responding can be described in the same terms applied to analyses of behavior acquired under indirect forms of reinforcement (or punishment). Furthermore, automatic reinforcement neither obviates nor precludes influence on response strength through social forms of reinforcement, although certain classes of (controlling) stimuli may be variously salient at times. Nor is behavior that is strengthened automatically constrained topographically more (or less) than indirectly reinforced responses. For example, when we have an itch, we scratch it and the itch goes away. When the room is too hot, we open a window and a cool breeze wafts through. If the road is steep, we lift our legs higher as we walk and successfully mount the incline. Examples are just as easily offered for the strengthening effect on these same behaviors through mediation by others. The current investigative task, then, is not to explain that behavior can be strengthened by its consequences but in discovering how this occurs with respect to human language, and Skinner's prolific
(Vaughan & Michael, 1982) use of the term furthers the analysis by drawing attention to “...the relevance of reinforcement in cases where it might be easily overlooked” (p. 218).

It is likely that much complex human vocal activity originates from automatically reinforced responses (Bijou & Baer, 1965; Pear, 2001). In terms of early language acquisition, automatic reinforcement may account for much early babbling and vocal play through a process in which a neutral stimulus becomes conditioned as a reinforcer if it has been sufficiently paired with another stimulus that already functions as a reinforcer. For instance, during infant caregiving, parents typically vocalize while delivering reinforcing stimuli (e.g., talking to the baby while feeding, changing diapers, rocking). These caregiver vocalizations (as initially neutral stimuli) acquire reinforcing properties as a result of repeated pairings with unconditioned or conditioned reinforcers related to caregiving (Schlinger, 1995). To the degree the child produces similar sounds (i.e., a close match to those of the parent), these stimuli, the auditory products of vocal musculature movements, may function as automatic reinforcers for those movements (Bijou & Baer; Schlinger). Thus, vocalizations could then be shaped into speech utterances that resemble those of the child’s own verbal community and, subsequently, further generalize through social (Skinner, 1957) and automatic (Palmer, 1996) contingencies of reinforcement operating within this unique community.
Automatically reinforced behavioral repertoires, including speech development, are consistent with a selectionist perspective of acquisition (Palmer & Donahoe, 1992). Susceptibility of an organism to reinforcement from stimuli produced by its own actions would have strong survival value, both phylogenically and ontogenically. Behavior such as orienting to, obtaining, and ingesting food, expelling or avoiding noxious substances, running from danger, approaching warmth, or engaging in sexual activity produce stimuli that benefit the organism and enhance the likelihood of its survival. That certain stimuli are unique to an organism’s particular environment is of minor import in explaining their strengthening or weakening effects on the behavioral repertoire of that organism. It is the contingent relation between those stimuli and behavior that dictates the selection of responses into the repertoire (Catania, 1998). Many behaviors acquired through automatic reinforcement are probably a function of non-social contingencies. Consider, for example, response variations of those already in the repertoire, evoked by requirements of the physical environment (e.g., stepping higher surmounts a larger than usual rock). However, other automatically reinforced behaviors may be predicated, in part at least, on social factors resulting from observing, imitating, and interacting with group members (e.g., running when others run escapes danger). These socio-cultural advantages of observation and imitation are easily extended to vocal behavior. Accurate imitation (of sounds and linguistic forms) that has been made possible by an adaptive and intricate repertoire, initially strengthened through
automatic reinforcement, benefits the individual (or social/cultural group) in ways that, collectively, result in the development of effective language practices.

**Role of Automatic Reinforcement in Early Speech**

By definition, if automatic reinforcement plays a role in early speech acquisition, it does so by strengthening vocal responses that produce auditory stimuli (i.e., speech sounds). That is, human speech sounds become conditioned reinforcers for the responses that produce these stimuli. The process by which stimulus changes acquire reinforcing properties has been hypothesized to occur through their contiguity with already established forms of either unconditioned or conditioned reinforcement (Catania, 1998; Michael, 2004). Although no response is required during the conditioning (i.e., pairing) process, a subsequent response contingency is needed to demonstrate these effects, showing a neutral stimulus has acquired reinforcing value. The hypothesized pairing function is extensively supported in the literature. It has been clearly shown that, with respect to operant\(^2\) conditioning, pairing establishes formerly neutral stimuli as effective reinforcers; further, that conditioned reinforcers affect response strength with both humans and non-humans (Kelleher & Gollub, 1962; Williams, 2002) although requisite variables (i.e., pairing parameters) essential to the conditioning process are the subject of debate (Williams, 1994).

\(^2\) Pairing is also used in respondent conditioning to establish an elicitation function of a formerly neutral stimulus as a conditioned stimulus (CS). This should not be confused with pairing effects in operant conditioning. In the former, responses are elicited by newly conditioned stimuli whereas, in the latter, a response is followed by a (formerly neutral) stimulus that, as a result of previous pairing with an effective reinforcer, now functions as a reinforcer in its own right for that operant response (Michael, 2004).
For typically developing children, naturally occurring pairings play a vital role by providing a launch pad for a wealth of early vocal play whose phonological units, in varied arrangements, can come under further contingencies of indirect and automatic reinforcement. Thus is built a dynamic, complex language repertoire. But what of the child who vocalizes minimally or not at all? Although natural environments may provide the requisite elements for vocal responses to be automatically reinforced (i.e., pairing history), some children may not be affected by these environmental influences. That is, developmental impairment (e.g., autism, mental retardation) may reduce the effects of social stimuli upon one's own utterances (Catania, 1998), thus decreasing the influence of paired stimuli and, ultimately, the opportunity for reinforcement of vocal responses.

Regardless of the available source of reinforcement (self-produced or mediated by another), the weakness of vocal responding severely compromises a child's potential to acquire language since little behavior is emitted for influence by any contingencies. Thus, procedures to establish auditory speech stimuli as reinforcing events would provide an important clinical tool to overcome response deficits and augment the available pool of vocal play responses, ultimately allowing topographies to be strengthened in functional relations.

*Speech as a Conditioned Reinforcer*

Few studies have focused specifically on the pairing aspects of conditioning neutral stimuli as reinforcers in humans, although there are notable exceptions. In a
group study with 70 first-grade typically developing children, Haines (1977) found that pairing a light with candy differentially increased subsequent lever presses that produced the light alone. Birnbrauer (1971) paired nonsense syllables with candy in a study with four mentally impaired adolescent boys to increase button pressing that produced only the nonsense syllable. These early studies are useful in that they demonstrate the utility of stimulus-stimulus pairing (SSP) to establish the value of previously neutral stimuli as conditioned reinforcers for automatically reinforced behavior. The Birnbrauer study, in particular, is important since the response strengthened through SSP (i.e., a button press) specifically produced a speech stimulus. Such behavior is analogous to responses of a vocal tract system (e.g., vocal cords, tongue, lips, lungs) that are strengthened by the auditory stimuli they produce.

However, since vocal responses were not required to produce the conditioned auditory stimulus, nor were they incidentally reported, it is unclear whether vocalizations of the paired stimulus (thus, automatically reinforced) also might have been at strength as a result of this preparation. In addition, since button pressing would not have produced the speech stimulus in the absence of the consequence being deliberately arranged (i.e., programming the button to produce an auditory stimulus), this relation might be interpreted as an example, not of automatic reinforcement, but rather, of indirect, socially mediated reinforcement. This possible interpretation notwithstanding, the Birnbrauer study illustrates that human speech, as
an auditory stimulus, can be established through SSP as a conditioned reinforcer for motor movements that produce it.

Several studies have recently evaluated the effects of SSP on free-operant vocalizations of children with both normal and delayed speech/language skills. Sundberg, Michael, Partington, and Sundberg (1996) paired novel syllables, words, or short phrases by experimenters with reinforcers in a study with 5 preschool-aged children, 4 of whom who demonstrated severe to moderate language delays. Although effects were short-lived, dissipating to pre-intervention rates after approximately 9 min in some cases, all children spontaneously emitted new vocal responses after pairings (although not all targeted sounds were emitted by all children). Because new responses appeared to be acquired without direct reinforcement, prompts, or direct echoic training, Sundberg et al. attributed effects of the pairing procedure to automatic reinforcement.

Smith, Michael, and Sundberg (1996) demonstrated increased post-pairing vocalizations of syllables already existing in the repertoires of two typically developing infants (< 18 mos) when syllables were paired with reinforcer delivery. Pairing other syllables with a mild aversive stimulus (i.e., verbal reprimand) conditioned those stimuli as punishers, evidenced by post-pairing decreases in paired topographies. To address the possibility that observed increases in responding might be attributable to already-established echoic control, Smith et al. included a neutral condition in which the auditory stimulus was presented without reinforcer delivery.
Since this did not result in emission of the target sound, the authors concluded that increased vocalizations following positive (reinforcer) pairings were not under the control of echoic contingencies but, instead, were automatically reinforced.

Other researchers also have shown SSP effects on subsequent vocalizations of previously low-rate responses in preschool-aged children with moderate to severe language delays. In a study using pleasant physical interactions (i.e., tickles) as reinforcers, Yoon and Bennett (2000) reported differential increases (albeit temporary, 3 to 16 min) in post-pairing vocal-verbal responses and concluded these increases were not attributable to echoic control since they were observed only following pairings and not when echoic contingences were in effect (with one exception). Showing less robust effects, Miguel, Carr, and Michael (2002) observed increased post-pairing vocalizations in 2 of 3 participants on at least one target syllable (previously paired with candy delivery). For one participant only, the transient effects of the pairing procedure, reported across minutes in previous studies, were similarly evident across sessions.

In these, and in subsequent studies reporting absent SSP effects (Esch, Carr, & Michael, 2005; Normand & Knoll, 2006), variables responsible for failures or diminished effects remain speculative. Some studies suggest that differential responding may be related to level of preexisting language skills. Yoon and Bennett (2000), for example, reported greater post-pairing increases by a participant with a relatively stronger pre-intervention vocal repertoire. In contrast, Miguel et al. (2002)
found that a child with stronger preexisting language skills demonstrated fewer increases in vocalizations following pairing. It may be that, for some children with more advanced language skills (e.g., mands, intraverbals; Skinner, 1957), reinforcement available through verbal interactions with others, as well as that provided by achieving parity (Palmer, 1996) with linguistic practices of one's verbal community, may supersede the eductive effects of the pairing procedure on non-function-based vocal responding (i.e., vocal play). It should be noted, however, that Sundberg et al. (1996) observed post-pairing increases in vocalizations by children with both strong and weak pre-intervention repertoires whereas Esch et al. (2005) reported no vocalization increases after providing an extensive pairing history for three young children (ages 6 to 8) with little preexisting vocal-verbal behavior; similarly, null findings were reported by Normand and Knoll (2006) using a pairing procedure with a three-year-old boy whose pre-intervention repertoire contained several vocal mands and tacts. Collectively, these findings suggest that failure or success of the pairing procedure to produce effects cannot be attributed, at least solely, to idiosyncratic characteristics of existing language skills, although, given the paucity of empirical literature in this area, careful delineation of these repertoires can only enhance attempts to identify which variables might be most relevant.

Discrepant findings reported in these studies highlight the question of whether the auditory stimuli (i.e., speech syllables) that were paired with reinforcers were successfully conditioned, in fact, as reinforcers themselves. Experimental
studies on infant vocal conditioning with secondary (i.e., conditioned) reinforcers have been concerned with demonstrating the strengthening effect of socially mediated reinforcement (e.g., Poulson, Kyparissos, Andreatos, Kymissis, & Parnes, 2002; Routh, 1969; Sheppard, 1969) in contrast to evaluating antecedent manipulations (i.e., SSP) to condition response-produced stimuli as automatic reinforcers.

To clarify this issue, it is useful to consider factors involved in SSP conditioning that may impinge on a paired neutral stimulus in some way that constrains its sensitivity to this process and the durability of its effects. For example, temporary effects (e.g., Sundberg et al., 1996; Yoon & Bennett, 2000) may indicate decreasing value of the self-produced auditory stimulus as these stimuli undergo unpairing similar to respondent extinction (Miguel et al., 2002). Such effects would occur as a result of the conditioned stimulus occurring repeatedly in the absence of supportive reinforcement (Michael, 2004). This analysis is plausible in investigations where effects are observed, regardless of their magnitude but, if accurate, points to the importance of prompt interventions to broaden the contingent control for vocal-verbal behavior expansion.

Stimulus decrement may impact responding in other ways. When SSP fails to increase responding, it is important to consider whether component laryngeal and articulatory responses are already at sufficient strength in a child’s repertoire such that they produce salient auditory stimuli. Hart and Risley (1999) noted that even before typically developing children said their first “real” words, they were emitting,
on average, 109 non-word utterances per hour. This intensity of unprompted practice allows speech-producing structures (e.g., lungs, vocal cords, tongue) to respond with ease to auditory models encountered in the environment. For children with severely limited speech and language skills, who emit very few vocalizations during the day, response requirements may be too great to emit coordinated movements of the vocal apparatus such that reasonable speech imitations are produced. In cases where speech stimuli have been conditioned through previous pairing, any decrement between the original auditory stimulus and that which is self-produced would result in decreased reinforcing effectiveness (Michael, 2004) of the child’s own speech, thereby reducing the probability of future vocal responses.

Aside from the comparative strength of preexisting vocal-verbal repertoires (discussed above), the number of pairings required to produce an effect may vary. Yoon and Bennett (2000) reported effects with as few as 36 pairings, whereas Esch et al. (2005) were unable to demonstrate increases after several thousand pairings. In addition, it may be difficult to adequately evaluate the conditioned value of a given speech stimulus, even if the number of pairings appears adequate. If stimulus value remains relatively weak, further decreases (in value) may be either negligible or occur so rapidly that effects cannot be observed.

Finally, it also may be that stimuli selected for pairing, and assumed to function as unconditioned or conditioned reinforcers, may not compete successfully with extant stimuli (i.e., other reinforcers) in particular environments. However,
studies in which SSP effects were not demonstrated employed stimulus preference assessment (SPA), a procedure that has been shown to accurately identify items that subsequently function as reinforcers for arbitrary responses (Hagopian, Long, & Rush, 2004). Moreover, participants’ consistent approaches to these stimuli (e.g., Esch et al., 2005) underscore that SPA-identified items had endurance, although perhaps less salience and value according to momentary motivative fluctuations.

**Focus of Current Investigation**

The current study was designed to explicate some of these issues. The primary question was whether novel speech could be conditioned as a reinforcer, evidenced by increased unprompted production of those speech syllables during post-pairing observation periods. Previous research (e.g., Miguel et al., 2002; Yoon & Bennett, 2000) showed inconsistent, temporary effects, a finding that critically limits the utility of SSP if its effects cannot be reliably produced and sustained over even a brief period (e.g., 5 min post-pairing). Perhaps conditioning vocal stimuli occurs only under distinctive conditions that are uniquely related to speech stimuli and/or vocal responses. The behavioral literature is replete with examples (Kelleher & Gollub, 1962; Williams, 2002) in which stimuli have been conditioned through pairing with either unconditioned reinforcers or other conditioned reinforcers and, thus established, have served to increase arbitrary non-vocal responding. However, previous SSP research (e.g., Esch et al., 2005; Miguel et al.) suggests that post-pairing vocal responding in some humans may be differently affected by, or less
sensitive to, stimuli conditioned (or unsuccessfully treated) in this manner. It is possible that the process of conditioning human vocal stimuli, such that they function to strengthen responses that produce them, requires different treatment than those procedures applied in recent investigations.

In light of these possibilities, several modifications to the typical SSP procedure were implemented to maximize obtained effects and inoculate against their disruption to the extent possible. With respect to data collection, earlier investigators typically evaluated responding that occurred during brief observation periods immediately preceding and/or following the delivery of pairings. However, these observation periods often yielded data that indicated weak, temporary, or absent effects of the independent variable. The current study was originally designed to similarly evaluate responding, but when pre/post-pairing observation periods failed to demonstrate SSP effects with the first participant, and concomitantly, experimenters observed responding during pairing sessions, it was decided to report within-session data to more accurately capture the effects of SSP on subsequent vocalizations. Furthermore, analysis of within-session data allowed appropriate data comparison during conditions (SSP/DR) where relevant stimuli were present in contrast to post-session periods in which vocalizations produced the auditory stimulus but the paired reinforcer was absent.

Procedural modifications were made to enhance the salience of the paired stimulus and to augment experimental control. In earlier studies, sessions consisted
of a series of trials in which a syllable was paired with delivery of a putative reinforcer (e.g., *ba-ba-ba* plus candy). To the degree that an individual was affected by these presentations, ability of the speech stimulus to acquire reinforcing properties through pairing would be more or less strong. Evidence from basic experimental research (see Dinsmoor, 1995a, 1995b) has shown that the effects of pairing can be increased by interspersing a stimulus that is not followed by a reinforcing stimulus (i.e., unpaired comparison S-) with trials in which a different stimulus (S+) is followed immediately by such an event. Given contrasting exposure to an interspersed schedule of negative pairing trials, the organism learns to react differentially to the stimulus with the pairing history, which, in its role as a newly conditioned reinforcer, serves to strengthen further responding. In the case of vocal responses, this advantage would obtain for a self-produced auditory stimulus resembling one with a pairing history over that of a non-paired response product. In the current study, interspersing S- trials with those of S+ maximized these pairing effects and provided additional benefit by controlling for elicitation effects of a non-target auditory stimulus.3

Another change to the SSP procedure involved addition of an observing prompt (Dinsmoor, 1995b) prior to initiation of any trial. The purpose of this prompt (e.g., *look*) was to increase the likelihood that succeeding auditory stimuli would be

---

3 This use of the terms S+ and S- is unconventional in the sense that neither stimulus has discriminative properties (see Dinsmoor, 1995a, 1995b). These designations are used here to specify the degree of their potential conditionality in order to provide a contrast for the likelihood of either stimulus acquiring such properties.
more salient as a result of the child’s attending response immediately preceding SSP presentation. This prompt preceded all trials because it was important for S- and S+ to be equally observable. (That S+ might subsequently acquire stronger reinforcing value over S- was not a function of the attending prompt but rather of the interspersal process during pairing.) Next, experimenters used highly salient prosodic patterns (motherese; Falk, 2004) when presenting syllabic models of both S+/S-. This ensured that speech models during sessions were different from non-relevant speech stimuli occurring between formal sessions (e.g., incidental conversation during session breaks) and, to the degree that this was achieved, the child would respond differentially to relevant stimuli (S+/S-).

A second research question considered whether, in cases where vocalizations were increased through pairing, specific reinforcement of responses that produced a previously paired speech syllable (i.e., mand training) would increase the frequency of those responses under relevant motivative conditions. There is conceptual support (Michael, 2004; Skinner, 1957, 1969) for such an effect and the clinical value for communication-impaired individuals is evident (e.g., Shafer, 1994).

A follow-up experiment (see Appendix A) was originally proposed for individuals in this investigation for whom SSP proved ineffective. The purpose of this evaluation was to determine if, as a result of a pairing history with established reinforcers, auditory stimuli functioned as conditioned reinforcers when an analog non-vocal response (i.e., button press) produced those stimuli. However, this
experiment was not conducted as part of the current study since all participants showed increases in target vocal responses, although the observation period that yielded these results was adjusted from between-session to within-session to more adequately capture treatment effects (see Results).

In summary, this study paired adult vocalizations with the delivery of preferred stimuli as consequences (i.e., reinforcers) to evaluate effectiveness of several modifications to the typical SSP procedure to increase subsequent vocalizations in children with delayed speech. Further, it evaluated the extent to which these new speech sounds could be transformed into meaningful communication (i.e., mands).

METHOD

Participants and Setting

Three children with severely delayed speech/language skills and diagnosed with autism spectrum disorder (ASD) participated in the study. Joshua, 2 yrs 4 mos, was not currently attending school, but had been enrolled to start soon in a pre-school classroom that offered applied behavior analysis (ABA) instruction for children with a diagnosis of autism. Although he occasionally vocalized, these vocal responses did not appear to be under the control of social contingencies (i.e., mand, tact, echoic, intraverbal). Madison, 2 yrs 8 mos, had been attending an ABA pre-school classroom for 1 month at the time of her participation in the study. For approximately 6 mos prior to school enrollment, she had received 1 hr per week of educational support at
home from her local school district. She frequently vocalized at play but, as with Joshua, she emitted no forms of verbal behavior with adults/peers. Daniel, 5 yrs 7 mos, had been attending an ABA classroom for 1.5 years at the time of the study and had been enrolled for 1 year previously in a classroom for children with developmental disabilities. Daniel’s verbal behavior consisted of a few mands, tacts, and intraverbals, all of which were emitted only under highly structured training conditions with the classroom teacher. At times, he was observed to echo simple sounds/words from the environment (e.g., TV, electronic toys) but did not usually or consistently do so when asked by teachers or others. Both Madison and Daniel received speech therapy 30-60 min per week at school. None of the children evidenced problem behavior (e.g., aggression, self-injury) or contraindicated sensory loss (i.e., deafness, dysarthria) and all were from homes in which English was the primary language.

Informed consent was obtained prior to the study, according to the requirements of the experimenter’s Institutional Review Board (see Appendix B). The experimenter obtained child assent, written parental consent, and verbal agreement from school administrators and teachers to conduct the study.

Children attended sessions 3-5 days per week. Each session lasted between 5 and 15 min, depending upon the experimental condition; these were typically conducted in a contiguous manner, 3-4 in a row, with brief play periods in between. Sessions were scheduled to maximize the value of putative reinforcers used in the
study; therefore, children did not attend sessions right after lunch, recess, or playtime. All sessions were conducted in unused classrooms or small conference rooms. Rooms were carpeted and equipped with a small table, chairs, and toys that caregivers had nominated as low to moderately preferred. Recording equipment consisted of a video camera mounted on a tripod and various data collection sheets on clipboards. During sessions, edibles and high-preference toys used as reinforcers in the study were kept in closed, opaque containers and were not available to the child except during appropriate sessions.

Prior to the study, the primary experimenter established familiarity with each participant by (1) observing the child in his/her classroom (or with a parent) for at least 15 min, (2) providing the child at least 15 min of free-play time with caregiver-nominated toys/activities in the treatment room, and (3) interacting in this room with the child directly for at least 15 min using caregiver-nominated reinforcers (e.g., toys, edibles). This procedure ensured the child was familiar and comfortable with the experimenter and the experimental setting.

Pre-experimental Assessments

Since clinical significance of the study concerned speech acquisition facility in children with ASD, it was important to evaluate existing verbal skills using tests that could provide information about relevant speech/language repertoires. A series of standardized and/or criterion-referenced assessments (see Table 1) were administered by a qualified speech pathologist (the primary experimenter), or parent/teacher,
depending upon test requirements. The Kaufman Speech Praxis Test (KSPT; Kaufman, 1995) was administered to determine participants’ extant echoic repertoires with single phonemes (e.g., /m/) and more complex syllabic (e.g., /ma/) and multi-syllabic (e.g., /mama/) constructions. Although KSPT is not explicitly described as a measure of echoic verbal function, its assessment tasks require echoic responses (with the exception of subtest 1 which requires non-vocal, oral-motor duplic responses such as moving the tongue from side to side). Additional speech/language information was obtained from informant observations of participants’ verbal repertoires, including the echoic function, using the Behavioral Language Assessment (BLA; Sundberg & Partington, 1998). This 5-point scale queries performance across the elementary verbal operants (Skinner, 1957; Sundberg, 2007) and selected skills related to learning (e.g., cooperation, social interaction). In addition to these assessments, an attempt was made to administer a normed receptive language test to evaluate skills in selecting (pointing to) named pictures. However, no participants met basal level responding on The Peabody Picture Vocabulary Test-III (PPVT; Dunn, Dunn, & Dunn, 1997) so an alternate skill inventory, The Receptive-Expressive Emergent Language Test – Third Edition (REEL-3; Bzoch, League, & Brown, 2003), was used to provide a reference point for comparing derived expressive and receptive language ages of participants. This assessment provides dichotomous (yes/no) data about language performance as observed in natural (non-instructional) settings. Similar to
Table 1

*Speech and Language Assessment Scores*

<table>
<thead>
<tr>
<th>Assessments</th>
<th>KSPT&lt;sup&gt;a&lt;/sup&gt;</th>
<th>PPVT&lt;sup&gt;b&lt;/sup&gt;</th>
<th>REEL-3&lt;sup&gt;c&lt;/sup&gt;</th>
<th>BLA&lt;sup&gt;d&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (months)</td>
<td>Items echoed (24 total)</td>
<td>Basal</td>
<td>Receptive</td>
<td>Expressive</td>
</tr>
<tr>
<td>Joshua</td>
<td>28</td>
<td>0 none</td>
<td>&lt; 6 mos</td>
<td>&lt; 6 mos</td>
</tr>
<tr>
<td>Madison</td>
<td>32</td>
<td>3 none</td>
<td>11 mos</td>
<td>12 mos</td>
</tr>
<tr>
<td>Daniel</td>
<td>67</td>
<td>2 none</td>
<td>12 mos</td>
<td>10 mos</td>
</tr>
</tbody>
</table>

<sup>a</sup>Kaufman Speech Praxis Test (Kaufman, 1995).

<sup>b</sup>Peabody Picture Vocabulary TestIII (Dunn, Dunn, & Dunn, 1997).

<sup>c</sup>Receptive-Expressive Emergent Language Test (Bzoch, League, & Brown, 2003).

<sup>d</sup>Behavioral Language Assessment (Sundberg & Partington, 1998). Verbal scores reflect range obtained for mand, tact, echoic, and intraverbal sections of this assessment.

the BLA, the REEL-3 is based on informant observations, but unlike the BLA, items are not arranged according to verbal function; therefore, specific controlling variables could not be inferred. Moreover, the REEL-3 was normed on non-autistic individuals and test items heavily sample social attention, a skill that, when impaired, is one of the defining features of autism. Thus, results of this test should be viewed conservatively when inferring language repertoires of individuals with skill deficits that could result in widely skewed scores on such measures. In addition to the speech/language tests discussed above, within one week prior to the study, an
inventory was made of typical vocal play consisting of phonemes emitted during a videotaped 30-min free-play period. Recording periods were recommended by parents/caregivers as optimal “talking times.” This inventory was later analyzed by topography and frequency of (any) phoneme occurrence and this information was used to inform target selection.

For Joshua, the BLA indicated low frequency of vocal play, no echoic responses, and no other verbal operants (i.e., mand, tact, intraverbal) in his repertoire. Echoic responses were similarly absent on the KSPT. No receptive language basal could be established on the PPVT-III. On the REEL-3, Joshua’s receptive and expressive language age scores were both below 6 mos. During the 30-min free-play period, he emitted 12 (of 42 possible) English phonemes. In comparison, phonetic transcription of speech samples from 520 typically developing children in California showed that all phonemes were acquired (i.e., emitted accurately) by age 3, with the exception of /l/ and /r/ topographies (Porter & Hodson, 2001). The 12 phonemes Joshua emitted were distributed over less than half (37%) of 30-s recording intervals during the 30-min observation period. Collectively, these speech/language assessments showed Joshua’s vocal-verbal repertoire to be severely delayed.

Madison’s test performance also indicated low verbal skills. She emitted few responses on the KSPT, none of which topographically matched the model. Although

---

4 In clinical practice, identification of phonologic process errors (e.g., final consonant deletion, syllable reduction) provides further information about the phonetic-phonologic context in which these topographies occur (see Khan & Lewis, 2002).
she vocalized frequently in general, the BLA reported no verbal functions; that is, Madison was not observed to emit mand, tact, echoic, or intraverbal responses. However, it should be noted that, in her classroom, Madison occasionally echoed 1-2 words (hi Madison), but this appeared random and inconsistent. The REEL-3 yielded receptive/expressive language age scores of approximately 12 mos and, like Joshua, receptive language responding could not be established during PPVT-III administration. During the vocal play inventory, Madison emitted 29 English phonemes; these responses occurred over 93% of 30-s intervals, indicating frequent and varied free-operant vocalizations. With vocal play her single strength, and absent any verbal operants, Madison’s functional speech and language skills were judged to be substantially impaired.

Daniel’s verbal skills were slightly higher than those of the other two participants. Although he echoed only 2 of 24 items on the KSPT, these responses were topographically accurate. Observations on the BLA indicated he vocalized frequently and could emit a few responses under mand, tact, and/or intraverbal control. As with Joshua and Madison, no basal was established on the PPVT-III and his receptive and expressive language age scores on the REEL-3 indicated severe delay (12 mos and 10 mos, respectively). Despite incipient verbal skills such as mand and tact relations, Daniel demonstrated infrequent vocal behavior during free-play. He emitted 21 English phonemes, but these were distributed over only 32% of 30-s
intervals, a pattern similar to Joshua’s (whose verbal behavior contained no functional operants).

**Stimulus Preference Assessment**

Prior to the study, parents and/or teachers completed a preference assessment survey (Fisher, Piazza, Bowman, & Amari, 1996) that yielded a ranked list of each child’s preferred edibles and toy items. These items were presented in separate assessments (i.e., toys, edibles) to the child three times in a multiple stimulus (without replacement) array (Carr, Nicolson, & Higbee, 2000) to verify preference ranking. The three highest-ranked items in each SPA then were selected for use as putative reinforcers during the study. Although reinforcer assessments were not conducted with these items, the literature (Carr et al.; Hagopian et al., 2004) indicates that items identified through stimulus preference assessments can function as reinforcers during treatment. To address the possibility of day-to-day changes in relative preference for various stimuli, the first daily session began with the experimenter presenting an array of items previously identified as preferred. Any items not touched, reached for, or accompanied by smiles when presented during a 1-min pre-session sampling period were eliminated and remaining items were randomly rotated during that day’s sessions. For simplicity of description, items identified through SPA as *preferred* henceforth will be referred to as *reinforcers.*
Target Responses and Response Definition

Target responses were selected from information obtained on speech/language assessments and from vocalizations observed during the 30-min free-operant period (described above). Targets were selected from phonemes or phoneme combinations that either were weak echoics (i.e., inaccurate or infrequent echoic responses to a model) or were not under echoic control (evaluated through the KSPT and/or BLA), yet were observed to occur in 10%–25% of 30-s intervals during the observation period. This selection procedure ensured that targets (or phoneme units that comprised targets) could be emitted by participants, and were emitted more than rarely, yet qualified as weak responses in terms of being under evocative control of an echoic stimulus. In cases where fewer than 10% of observation intervals contained potential targets or phoneme units, the experimenter selected targets based on available (emitted) topographies. Target syllables were selected that had little likelihood of prior reinforcement in order to avoid experimental confounds in which previous experience may have exerted influence over responding. Targets were beh and oo for Joshua, sheba and aypayk for Madison, and reeklo and tebba for Daniel.

Target responding was defined as the production of any target syllable that matched or was topographically (e.g., pa for ba) or acoustically (e.g., uh for ah) similar to the paired model to be trained (S+). Responses matching single phoneme targets (e.g., /a/) were counted if they contained the same vowel or a similar vowel and did not contain a consonant phoneme (e.g., /p/). Responses matching multiple
phoneme targets (e.g., /ba/) were counted if they contained the same or similar vowel and consonant. Similarity was defined as acoustic and/or phonologic (e.g., articulatory feature plus proximate placement) approximations to a particular phoneme.

A non-target response was defined as any response that was the same or similar to the unpaired comparison model (S-) to be interspersed with S+ for contrast purposes. Vocalizations that did not meet the definition of target or non-target responses were not counted. Non-speech vocalizations (e.g., laughing, burping, screaming, crying, coughing, grunting, gagging, sustained or repetitive humming) also were excluded. Any syllable separated by a 1-s interval from any other syllable was counted as one response. In this case, an interval was defined as the absence of a voiced (e.g., /m/) or unvoiced (e.g., /f/) vocal response. Thus, syllable production that was temporally distorted from the experimenter’s model was counted as either one response (e.g., baaaaa) or several responses (e.g., b-b-b-b-ba) depending upon the inter-response interval.

Participants’ vocal responses were recorded as they were emitted during varied intertrial intervals (Catania, 1998; Gibbon & Balsam, 1981) of baseline and SSP conditions and throughout DR and No DR conditions (see below). By varying the intertrial intervals (ITI), fixed time passage was reduced as a confound to SSP effects in that target responses could more reliably be attributed to the conditioning
procedure by eliminating temporal predictability while holding constant the paired
relation between relevant stimuli.

*Interobserver Agreement*

Two independent observers manually recorded session data during a
minimum of 25% of randomly selected sessions (balanced across conditions) either *in vivo* or from video recordings. Interobserver agreement (IOA) on frequency of target
and non-target vocalizations was calculated separately using the total count method
by dividing the smaller frequency of vocalizations per session by the larger frequency
per session and multiplied by 100 to yield a percentage of agreement per session.
Target IOA and non-target IOA were assessed for sessions in which the respective
relevant data were collected (target: BL, SSP, DR, No DR; non-target: BL, SSP).
Data were excluded for sessions in which events occurred that appeared to have an
unscheduled rate-influencing effect. For Daniel, all data were excluded for a session
during which, when the school’s public address system came on unexpectedly, he
starting yelling and crying and ran from the training area. No data were excluded
from any other sessions.

For Joshua, target-1 IOA was calculated on 81% of sessions; mean IOA
percentage on target occurrence was 89.4% (range, 56% to 100%). Non-target IOA,
calculated on 86% of sessions, was 100%. On target 2, IOA was assessed on 73% of
sessions and averaged 91.8% agreement (range, 80% to 100%). Non-target IOA for
topography 2, calculated on 63% of sessions, was 82.8% (range, 0% to 100%). For Madison, target-1 IOA was assessed across 35% of sessions. Agreement averaged 93.1% (range, 75% to 100%). Non-target IOA on 40% of sessions averaged 86.9% (range, 33% to 100%). Target 2 IOA for Madison was assessed on 35% of sessions; mean agreement was 98.1% (range, 91% to 100%). Non-target IOA was calculated on 38% of sessions and was 100%. For Daniel, target-1 IOA was calculated for 29% of sessions and averaged 96.5% (range, 85% to 100%). Non-target IOA, assessed on 29% of sessions, was 100%. For target 2, IOA was calculated on 38% of sessions, averaging 99.6% (range 94% to 100%). Non-target IOA was assessed on 33% of sessions and averaged 97.6% (range, 83% to 100%).

Procedure

The purpose of this study was to evaluate the effects of an enhanced SSP procedure on self-produced speech sounds as conditioned reinforcers for responses that produced these auditory stimuli and to directly strengthen these responses as functional forms of communication through specific reinforcement (i.e., mand training).

Experimental design. A nonconcurrent multiple-baseline design across phoneme topographies was combined with an ABCA'C design to evaluate the

---

This value represents one observer scoring 1 instance of the behavior and another observer recording no occurrence. It should be noted that Joshua's target 2 was a single vowel phoneme (/u/); thus, all other vowels were counted as non-targets. Although independent observers (college students) were trained to listen for target/non-target sounds, some discriminations (e.g., oo vs. aw) presented a challenge for those without specific (i.e., speech pathology) listener training.
effects of the SSP procedure and subsequent differential reinforcement on the
frequency of within-session vocalizations.

Baseline. Baseline sessions began with a brief period (1-2 min) in which the
child had access to low-moderate preference toys in the session room. The
experimenter then directed the child to the treatment setting (table/chair or floor area),
designed to be visually barren to enhance stimulus control during sessions. Each
baseline session consisted of 20 trials, 10 each of S+ and S-, randomly arranged but
with no more than 2 consecutive trials of either to decrease predictability of a
particular stimulus (Catania, 1998). Trials of S+ consisted of the experimenter saying
the target syllable without its paired stimulus. S- trials had no paired stimulus; thus,
baseline S+ and S- trials were identical except for the auditory characteristics of the
two models presented. All trials were preceded with a prompt to attend (e.g., look).
Syllables were presented at the rate of one syllable per second for 3 s (e.g., S+ ba ba
ba; S- dee dee dee). Trials were separated by an ITI that varied between 5 s and 30 s
after which the next scheduled trial was presented. Between sessions, participants had
access in the session setting to free-play with low-moderate preference toys.
Interactions between participant and experimenter occurred only to the extent
minimally necessary to ensure safety (e.g., preventing the participant from climbing
or leaving a supervised area). Baseline was terminated when visual inspection of data
revealed no consistent separation of target and non-target data paths.
Stimulus-stimulus pairings. Pairing sessions (see Table 2) were identical to those in baseline with two exceptions. Presentation of S+ trials included immediate delivery of a reinforcer following the model. During S+ trials, the child had access to the reinforcer for 10 s or, in the case of edibles, until swallowed. The other difference was inclusion during S+ trials of a 20-s correction delay if the participant emitted the target response between the experimenter’s model and delivery of the reinforcer.

Table 2

<table>
<thead>
<tr>
<th>Trial type</th>
<th>Trial components</th>
<th>Sample Procedure</th>
</tr>
</thead>
<tbody>
<tr>
<td>S+</td>
<td>Speech syllables and reinforcer</td>
<td>la la la tickles ITI</td>
</tr>
<tr>
<td>S-</td>
<td>Speech syllables only</td>
<td>mi mi mi ITI</td>
</tr>
</tbody>
</table>

Note. All trials were preceded by an observing cue (e.g., look!).

During the correction delay, the experimenter did not look at, or interact with, the participant. If a reinforcer was already partially delivered or irretrievable (e.g., the bubble machine was already turned on or the edible was already placed at the child’s lips), reinforcer delivery was completed. Otherwise, reinforcer delivery was withheld for the correction delay period. Any responses emitted during this delay were not counted. The correction procedure controlled for adventitious reinforcement of the target response. If the child emitted any other vocal response in the period between

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
stimulus presentation and reinforcer delivery, no correction delay was imposed since target responding was the variable of experimental interest; moreover, reinforcers only followed S+ models and the likelihood of a non-target response occurring between S+ and reinforcer delivery was not great. Furthermore, even if non-target responses (or any other vocalizations) were emitted between S+ and reinforcer delivery, thus undergoing accidental reinforcement, any subsequent target (over non-target) rate increase would further support SSP effects on the dependent variable. As in baseline, data were collected during ITI and interactions between participants and observers were kept to a minimum. With one exception (see Figure 3, upper panel), the pairing condition was concluded when graphed data showed clear and consistent differences over a minimum of 3 consecutive data points between target and non-target responding.

*Differential reinforcement.* The purpose of the differential reinforcement (DR) condition was to evaluate and promote function transfer of the response product (speech stimulus) from one that maintained automatically reinforced responding to one that produced a specific reinforcing event mediated by others. Increased target responding during DR would suggest either a) maintenance of the automatic reinforcement contingency and/or b) establishment of a mand function through contingencies of differential reinforcement. Although a return to baseline condition would have provided additional basis for interpreting SSP effects on responding, previous research showed these effects to be fragile and temporary; therefore, the
multiple-baseline design provided increased experimental control while avoiding a return to baseline and its associated risk of losing the newly acquired response altogether.

During each 5-min DR session, the experimenter delivered a reinforcer within 5 s after a target vocalization. To maximize the likelihood of obtaining a reinforceable response, each DR session was preceded by (typically less than 5) SSP trials of S+ only, provided at the rate of 1 syllable per second (in triads) every 5-10 s until the child emitted a target response (possibly under at least partial echoic control; Tonneau, 2005). These preliminary pairings were omitted if, when the session began, the child immediately emitted the target response. Data collection for DR sessions began after the first target response and continued for 5 min. During this period, if 1 min elapsed with no target responding, another S+ pairing was provided. The DR condition was concluded when an upward trend in target responding was evident from visual inspection of the data or when a stable data path suggested further changes were unlikely.

Withdrawal of differential reinforcement. The comparative effects of automatic and differential reinforcement on target responding were further evaluated by conducting withdrawal (No DR) sessions in which preferred stimuli were delivered for 5 min on a fixed-time schedule (Cooper, Heron, & Heward, 2007) every 30 s (FT 30-s). As in SSP, a correction delay of 20 s was implemented to control for adventitious reinforcement. If a target response occurred within 5 s of scheduled...
reinforcer delivery on the FT 30-s schedule, a 20-s delay was imposed, followed by reinstatement of the FT schedule.

To the extent responding maintained in the absence of DR during this condition, conclusions could be made regarding the strength of DR as a controlling variable for post-SSP responding and the separate influence of automatic reinforcement on response maintenance in the absence of DR influences. This condition also served to offset somewhat the omission of a return to baseline following SSP (before DR). In other words, because reversal after SSP was undesirable due to (previously demonstrated) fragile effects, the No DR condition was useful in that it enabled interpretation (although not conclusive evidence) of enduring automatic reinforcement effects to the extent responding maintained during this condition. Thus, No DR permitted comparison against baseline of response strength in the absence of the original paired stimulus, yielding information that could potentially clarify the issue of how speech maintains its conditioned value in the absence of the initial conditioning agent (i.e., reinforcing stimulus).

Caregiver training. Since clinical benefit beyond the experimental setting was desired for participants, parents or caregivers received brief training as part of the exit interview held at the conclusion of the experiment. The purpose of this meeting was to demonstrate SSP and differential reinforcement procedures, thus increasing the likelihood that new vocal responses would be evoked and effectively maintained in the child’s natural environment.
Independent Variable Integrity

An independent observer assessed treatment integrity for a minimum of 25% of sessions, balanced across conditions. Observations were made either during sessions or later from videotaped recordings of sessions. Trials were scored as completely correct or incorrect. The number of correct trials, divided by the number of correct plus incorrect trials, and multiplied by 100, yielded an independent variable integrity (IVI) percentage score that was averaged across sessions (mean IVI score). Baseline trials were correct if (a) the orientation prompt preceded each presentation of S+ or S-, (b) 3 scheduled syllables (either S+ or S-) were given within 5 s, and (c) no reinforcers followed S+ or S- presentations. A pairing segment trial was correct if (a) the orientation prompt preceded each presentation of S+ or S-, (b) 3 scheduled syllables (either S+ or S-) were given within 5 s, (c) reinforcers followed S+ presentations within 5 s and no reinforcers followed S- presentations, (d) an inter-trial interval of 5-30 s occurred, and (e) a correction interval of 20 s followed any target response that occurred between presentations of S+ and reinforcer. Mand training trials were correct if a reinforcer was presented within 5 s after the child vocalized any target syllable that was not immediately (within 3 s) preceded by an echoic prompt from the experimenter. Withdrawal trials were correct if (a) reinforcers were delivered on an FT 30-s schedule and (b) a delay interval of 20 s occurred after any target behavior emitted within 5 s prior to scheduled reinforcer delivery. For Joshua, treatment integrity was calculated for 67% (target 1) and 53% (target 2) of sessions.
Mean treatment integrity was 99% for both targets (target 1 range, 94% to 100%; target 2 range, 90% to 100%). IVI was calculated on 35% of all sessions with Madison. Mean IVI percentages for Madison were 99% (range, 95% to 100%) for target 1 and 100% for target 2. Daniel’s IVI percentages were calculated on 29% (target 1) and 38% (target 2) of sessions; mean treatment integrity was 100% for both targets.

RESULTS

Results for Joshua are shown in Figure 1. Rate of target 1 responses (upper panel) during SSP increased slowly over near-zero baseline rates to 4.4, with interim responding stable, but varied, within a frequency range of 0 to 2.8 responses per min (RPM). This variability without trend underscores the difficulties often seen in teaching children with autism where stimulus control can be slow to establish, an issue perhaps related to low stimulus salience as a function of weak or differential (i.e., error) attending. Inclusion of S- in procedural modifications partially addressed this, but the S+/S- contrasting stimulus conditions were slow to demonstrate their effects. When social contingencies (DR condition) were implemented, responding initially decreased, demonstrating sensitivity to the absence of pairing or, perhaps, lack of sensitivity to response-dependent contingencies. However, responding recovered to SSP levels by the 5th DR session, indicating a learning effect, and continued in an upward trend to 7.6 RPM. When reinforcement was withdrawn and reinforcers were available on an FT schedule, target responding immediately
decreased to 1.2 RPM, but then maintained at rates between 3.4 and 6.2 RPM, suggesting that the self-produced speech stimulus was, at least in part, maintained by automatic reinforcement since other (indirect) contingencies were no longer available. Joshua did not emit any non-target responses during the experiment for target 1 providing further support for positive SSP effects. Similar effects were observed with the second topography. Target 2 responding (Figure 1, lower panel) occurred at a mean rate of 0.6 RPM (range, 0 to 1.6) throughout baseline and the first 20 SSP
sessions. Over the next 19 SSP sessions, however, mean rate of response increased to 2.2 (range, 0 to 7.8). When DR was implemented, responding maintained at or above SSP levels, with a mean rate of 2.9 responses per min (range, 1 to 5). Evidence for an automatic reinforcement effect was indicated when responding during No DR maintained, after an initial decrease to zero following DR, with a mean rate of 1.8 RPM (range, 0.4 to 2.6).

Results for Madison are shown in Figure 2. The upper panel shows Madison’s vocalizations of target 1. Non-target vocalizations were not emitted. Baseline target responses were minimal, averaging 0.06 RPM (range, 0 to 0.5). Target responding steadily increased during SSP, showing differential pairing effects, but these appeared modest as response levels did not exceed 2 RPM; mean response rate was 0.9 (range, 0 to 1.7). When target responses were differentially reinforced (DR), vocalizations immediately increased to 8.8 RPM, stabilizing at about 6 RPM. Target overall mean rate during DR was 6.4 RPM (range, 5.4 to 8.8). The magnitude and immediacy of response changes upon DR implementation suggests strong differential influence of indirect reinforcement contingencies. When reinforcers were available on a response-independent basis (No DR), responding decreased to near zero (0.6 RPM) within 3 sessions. Upon reinstating the DR contingency, responding re-established at previously high levels; mean rate of response during the second DR segment was 5.1 (range, 0.5 to 7). For target 2 (Figure 2, lower panel), BL target and non-target topographies were emitted at similar, low frequency rates (except in session 4 where
both responses occurred approximately 5 times per min). In SSP, there was a
differential increase in target over non-target vocalizations (which remained near zero
throughout the pairing condition). Despite steady increases, overall SSP mean target
rate was 1 RPM (range, 0 to 2.3) compared to a BL rate of 0.6 (range, 0 to 4.8).
During DR, target 2 responding was variable which appeared, anecdotaly, to be
related to changes in reinforcer value (i.e., satiation). Nevertheless, the DR condition
showed a moderate upward trend over SSP levels with a mean target rate of 4.1
(range, 0.6 to 9.4) during DR. When reinforcement was withdrawn (No DR condition) and reinforcers were available on a response-independent schedule (FT 30-s), response level immediately decreased, dropping to zero by the 3rd session and continuing at or near zero throughout. Mean target responding during No DR was 0.7 (range, 0 to 3). In order to recover the target vocalization under mand contingencies, DR was reinstated briefly. Responding immediately increased to 3.4 RPM and maintained near this level (mean rate, 3.3 RPM; range, 2.8 to 3.6).

Results for Daniel are shown in Figure 3. During the evaluation of Target 1 (upper panel), no BL responding occurred for either topography (target/non-target). During the first half of the SSP condition (sessions 6 through 16), the differential effect of pairing on targets was evident, with a mean rate of 2.7 (range, 0.3 to 4.7) compared to the non-target mean frequency of 0.5 RPM (range, 0 to 1.1). However, during the second half of SSP (sessions 17 through 28), both target and non-target vocalizations decreased to less than 0.5 RPM, with one exception (target data point 24 value was 1.4). When contingencies of social reinforcement were initiated (DR condition), level of target responding showed an immediate increase (5.8 RPM) then continued above 2.5 RPM (mean rate 4 RPM). When reinforcement was withdrawn and reinforcers were available on an FT schedule, target responding decreased over 3 sessions to 0.6 RPM, suggesting that the self-produced speech stimulus was no longer maintained by automatic reinforcement. Reinforcement contingencies were reinstated, and target responding immediately increased to 5.9 RPM, higher than
response levels during the first DR condition. Mean target frequency during the 2nd DR was 6.6 RPM (range, 5.4 to 7.6).

The lower panel in Figure 3 shows Daniel’s vocalizations during Target 2 training. During BL, after an initial period when both target and non-target topographies were emitted at fairly high rates (approximately 10 RPM during session 1), vocalizations varied in a downward trend between 0 and approximately 4 RPM, and then stabilized at zero. Target responding remained below 1 RPM for the first 4
sessions of SSP but then increased to approximately 3 RPM. At this point, given performance on Target 1, DR was initiated in order to avoid response loss and maximize clinical benefit by strengthening this topography as a mand. The target response immediately responded to the contingency with an initial frequency of 15 RPM, increasing to 19 RPM by the 3rd session of the DR condition. When the FT schedule was implemented (No DR), response level decreased to zero. Reinstatement of the DR contingency again resulted in immediate increases in responding to 5.9 RPM, increasing to a mean overall rate of 8.2 during the 2nd DR condition.

DISCUSSION

This study investigated the separate roles of automatic and indirect (i.e., socially mediated) reinforcement on novel vocal responses of speech-delayed children with a diagnosis of autism spectrum disorder. Of primary interest was the evaluation of an enhanced SSP procedure in establishing auditory speech stimuli as conditioned reinforcers for the vocal responses that produced these stimuli. Secondarily, the study evaluated the effects of indirect reinforcement and its subsequent withdrawal on SSP-induced speech responses as mand operants. Results showed that all children acquired new target vocalizations at acceptable but moderate levels over baseline and in contrast to non-target syllables that did not undergo conditioning through pairing, thus suggesting an automatic reinforcement function for self-produced speech stimuli. Further, in all cases, target topographies were strengthened at or above SSP levels through subsequent socially mediated
reinforcement. When these response-dependent social contingencies were withdrawn, only the participant with the lowest pre-experimental vocal repertoire demonstrated target maintenance.

Implications

The current study offers strong evidence that new speech responses can be induced through SSP. Use of an enhanced pairing procedure that maximized stimulus salience resulted in the most reliable empirical findings in this line of research to date. The relatively modest effects shown for two participants, however, suggest that the absolute conditioned reinforcing value of these vocalizations may play a minor role in response maintenance, particularly in speech-delayed children with at least moderate existing vocal play. Taken together with previous findings, these data suggest that automatic reinforcement of self-produced speech sounds, conditioned through pairing, may briefly strengthen such vocalizations, but that socially mediated reinforcement is required to produce durability. Moreover, the timeframe to initiate this requirement is brief, perhaps only a few minutes.

Conceptual Issues

Inherent in SSP speech studies is the assumption that, for individuals with delayed speech, a greater, perhaps critical (cf. Yoon & Bennett, 2000), mass of vocal play responses would provide increased opportunity for these responses to come under the control of reinforcement contingencies, thus strengthening the speech repertoire; further, that these responses are not presently occurring at adequate
strength precisely due to the lack of such reinforcement histories. Moreover, in the absence of sufficient behavior that can come under the influence of mediated reinforcement, automatic reinforcement might bridge the gap by strengthening weak or newly acquired responses that subsequently can be subjected to social contingencies. Thus, the empirical issue is how speech (i.e., auditory response products) can be conditioned as an automatically reinforcing event for children who fail to emit adequate vocal-verbal behavior at an age when such a repertoire would be expected.

To isolate the specific contribution of automatic reinforcement in the current study, it was important to rule out other possible sources of control. First, speech responses in this investigation were pre-empted from access to indirect reinforcement during SSP by inclusion of a correction delay; thus, although target responding increased in SSP, higher rates could not be attributed to adventitious reinforcement of these responses as mands, echoics, or other verbal operants. Other explanations were excluded because they similarly lacked parsimony. It could be argued that the adult speech model functioned as an arousal stimulus (Tonneau, 2005) since all participants demonstrated responding when these models were present. This is plausible since both target and non-target syllables consisted of phonemes that were at some (low) strength in the pre-experimental repertoire and, therefore, it could be expected they would be emitted under certain stimulus conditions. However, target and non-target response rates, undifferentiated during baseline, showed clear separation during SSP,
making this explanation inadequate since any eliciting effect would not be likely to selectively affect only target responding. Finally, another possible account concerns a stimulus difference between baseline and SSP. During baseline, both speech models (S+, S-) were presented, but reinforcers were absent. In SSP, the same auditory models were presented but S+ was followed by reinforcer delivery. It is possible that the enriched environment (i.e., reinforcer presence) during SSP could have positively influenced response rates, irrespective of contiguity with the speech model. However, if this were the case, non-target speech would have had an equal opportunity to be similarly affected, which it was not, as evidenced by higher levels of target (relative to non-target) responding. Furthermore, previous research (Miguel et al., 2002) provided no support for an elicitation effect when these variables (speech models and reinforcer presentations) were controlled. Thus, it is reasonable to conclude that response rates during SSP were differentially affected by the specific pairing of target syllables with established reinforcers and were not solely due to the presence of preferred stimuli.

The positive effects of pairing reported in this and previous investigations are tempered by their transience. (Indeed, in this study, responding rarely occurred during the 3-min post-pairing intervals initially used for data collection, so observations were replaced by within-session data collection in order to more accurately capture effects.) This type of SSP responding indicates that there may be only the briefest optimal window during which particular stimuli are salient (and valuable) enough to
function as reinforcers for new responses. Furthermore, characteristics of an ideal group of responders are not clear as requisite learning histories and repertoires have not been defined, nor is it apparent how such conditions might contribute to stability and durability of new vocal responses. Improved behavioral procedures such as the enhanced SSP method described herein provide a treatment of choice for severely speech-impaired individuals but highlight the need for further analysis of antecedent and consequent stimulus control over weak vocalizations.

Some relevant antecedent variables are not available to the researcher and practitioner for modification. Existing vocal repertoires and learning histories undoubtedly influence sensitivity to treatment contingencies (i.e., SSP procedure), yet their relative importance in speech acquisition awaits empirical clarification. Certainly, weak SSP effects are not unexpected in a population of children with weak vocal-verbal skills. Indeed, the speech/language delays that partially define autism may be affected negatively even further by another feature of the disorder in which responding (e.g., attending) to relevant stimuli is compromised, often severely so. As a result, children with autism may be minimally impacted by natural (unprogrammed) environments that, for other children, are adequate to stimulate and strengthen early vocal behavior in the form of automatically reinforced vocal play. Thus, we could expect that these individuals experience fewer opportunities to achieve speech fluency at a basic phonation level (i.e., coordinated rapid movements of articulators with simultaneously controlled air pressure and resonance valving) and, more
functionally, verbal competence through acquisition of mands, tacts, and other verbal operants. In effect, these restrictive factors result in few reinforceable vocal responses, and, therefore, children who are already speech-delayed may be further disadvantaged by reinforcement histories that are persistently deficient, thus impeding even more profoundly the acquisition of adequate speech skills.

Esch et al. (2005) speculated that weak SSP effects might be due to the articulatory demands of the vocal task, especially for children with infrequent vocal behavior in general. There would seem to be an obvious advantage to foundation repertoires (i.e., babbling and vocal play) that are more extensive in terms of topographic and frequency parameters and, accordingly, an assumption that SSP effects would obtain for higher repertoires. However, results of the present study suggest that parametric differences in existing vocal play repertoires may actually have an inverse relation to predicted SSP effects. Performance differences were found between Joshua, whose vocal play repertoire was weakest (12 topographies), and Daniel and Madison, both of whom emitted approximately twice as many phoneme topographies (21 and 29, respectively) during pre-experimental observations. Responding for Daniel and Madison, with stronger initial repertoires than Joshua, was more sensitive to mediated reinforcement contingencies, as evidenced by response extinction when these contingencies were withdrawn (i.e., No DR condition). In contrast, Joshua’s responding on both topographies maintained (after an initial decrease) at levels observed in DR despite the removal of reinforcers delivered by the
experiment, indicating that auditory response products themselves served as
(automatic) reinforcers for the response. The point at which the value of speech
stimuli is altered from absolute to relative in terms of control by social contingencies
is speculative, but performance by these participants indicates that fairly small
discrepancies in vocal play may be predictive of susceptibility to these changes in
contingency source. That is, all the children demonstrated severely delayed speech
skills, but only Joshua, who appears to have been differentially influenced by SSP,
had a vocal play repertoire that was restricted in both response topography and
frequency. Thus, perhaps as soon as vocal play repertoire is at particular strength,
social contingencies more effectively compete to strengthen responding.

Despite differences in initial vocal play skills of these participants, any
differential influence on specific target responding was probably negligible. That is, it
is unlikely that any syllable had a selection advantage based upon relative complexity
or prior reinforcement history, since individual targets consisted of nonsense syllables
and were selected from phonemes occurring at low frequencies in each child’s vocal
play. Therefore, although level of initial vocal play may influence overall sensitivity
to automatic reinforcement contingencies, it likely did not have a differential effect
on any particular target topography in this investigation.

In addition to participants’ learning histories and pre-experimental repertoires,
other variables may have constrained SSP effects. Weak or fragile pairing effects
have been more evident in human studies than in those conducted with non-humans

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
(e.g., Dinsmoor, 1995b) and the reasons appear related, not surprisingly, to both stimulus control and response access (i.e., repertoire). Special difficulties arise in arranging effective reinforcement histories for humans where researchers have less latitude in experimental manipulation. With respect to stimulus control, one challenge is in identifying and establishing necessary levels of deprivation (i.e., establishing operations; Michael, 2004) in order to maximize reinforcer values that will support conditioning, a problem more easily managed in applications with non-humans. Although SSP does not require a reinforceable response, effective conditioning of an arbitrary stimulus necessitates that the conditioning stimulus already function as a reinforcer and that it maintains its value throughout the conditioning process (Michael). Since stimulus value exists on a continuum, influenced by dynamic environmental events, the applied researcher must be vigilant to maintain conditions necessary for stimuli to function as effective conditioners since response variability may be related to these environmental stimulus changes. In the current study, it is possible that a more molecular analysis of responding (i.e., momentary fluctuations in motivations) would show variability to be a function of the relative strength or weakness of the particular reinforcing stimulus used during pairing. Certainly both of these factors, variability in responding and changes in reinforcer value, were observed with all participants in this investigation, but their covariance was not determined and it is possible that negatively reinforcing stimuli would have resulted in similar, or greater, SSP effects had relevant establishing operations been arranged.
In terms of response access and the demonstration of SSP effects, there is an important difference between responses targeted for conditioning and acquisition in human SSP research and the dependent variables targeted in basic studies with non-humans. In the latter, the target response is typically one that is already strong in the animal’s repertoire (e.g., pecking); therefore, the task is to bring an already probable response under the control of specific stimulus conditions. In contrast, the response of interest in SSP studies (i.e., talking) is obviously weak in the repertoires of speech-delayed children. In fact, the procedure’s utility as a clinical tool specifically hinges on its ability to establish novel response topographies or strengthen those that have been resistant to acquisition. In either case, components of the human target response (i.e., specific vocalizations) likely have little history of reinforcement regardless of its source. These species differences and the experimental challenges they present should not imply that behavioral procedures that have proven robust in the non-human laboratory are inherently less useful in applied settings or with specific populations, but the difficulties of effectively extending these applications must be resolved.

The relative inflexibility of the foregoing factors in explaining SSP restricts their role in defining both optimal utility of the procedure and those most likely to respond to it. The current investigative task was to identify features or deficiencies in the typical SSP procedure that could be altered (or added) to produce greater clinical benefit. Collectively, the integration of several procedural changes (inclusion of an orienting prompt, interspersals of unpaired/paired syllables, accentuated prosody of
syllable models, and variable ITI) into an enhanced SSP treatment established novel speech syllables as conditioned reinforcers for vocal responses producing those stimuli. As such, either singly or in concert, they represent important variables for analysis in delineating the behavioral processes through which early speech may be acquired in the absence of extensive or specifically programmed contingencies but rather through reinforcement arising from self-delivered stimuli (i.e., automatic reinforcement). Since reinforcement regardless of source (i.e., automatic or socially mediated) offers a powerful clinical tool to help speech-delayed children achieve functional communication, the interplay between these two potential sources of reinforcement is initially examined, followed by a discussion of issues related to specific procedural changes.

Several of the issues in this investigation are related to the multiple sources of stimuli and associated stimulus values that are characteristic of SSP studies. Pairing requires that a stimulus of greater value (i.e., unconditioned or conditioned reinforcer) follow, according to some temporal arrangement (Dinsmoor, 1995a; Williams, 1994), a stimulus of lesser value (arbitrary stimulus; Pear, 2001). The conditioning effect can be disrupted in a number of ways (e.g., blocking, overshadowing; Donahoe & Palmer, 1994) but, in general, there is broad empirical support for the ability of established reinforcers to imbue their value to other stimuli. In speech acquisition studies, proof that the latter has acquired reinforcing value through its repeated pairings with a known reinforcer is evidenced by increases in vocalization responses.
that produce these (arbitrary) auditory stimuli. The current study demonstrated response increases related to all six paired syllables, suggesting increased value of the target stimulus as a result of SSP and thus automatic reinforcement as the source of control. If some stimuli had failed to be conditioned through pairing, an appeal to aforementioned influences (e.g., blocking) might inform the analysis but, since all participants demonstrated acquisition during SSP, issues of concern must be otherwise explained.

The first of these issues is Daniel’s performance on topography 1 (see Figure 3, upper panel); it is unique to that of other participants’ as well as to his later performance on topography 2. Separation of the target and non-target data paths during the first half of SSP indicates a difference in stimulus values between these syllables; thus, a conditioning effect of pairing. However, this effect was lost during the second half of the condition (i.e., while pairing was still occurring) suggesting interruption of pairing, or interference with its effects, through other mechanisms.

In this case, there are at least three possible interpretations for loss of effect during SSP. First, SSP-induced responding may have been highly dependent upon continued pairings with the original paired reinforcer (e.g., cake), since every post-pairing response was essentially an un-pairing (see Michael, 2004; Pear 2001). To the extent that responses occurred uncorrelated with the original paired reinforcing stimulus, the target response would be weakened. (The other possibility of the paired
reinforcer being delivered without the arbitrary paired stimulus was not an issue; treatment integrity for this condition was 100% so this event did not occur.)

Second, another source of control (i.e., past social reinforcement for vocalizations) may have been stronger than that of automatic reinforcement such that Daniel’s responding was brought to strength not as a result of the current pairings, but, instead, by more distal contingencies (i.e., previous mand or echoic training), further occasioned by the instructional setting and the presence of an adult with reinforcers related to establishing operations currently in effect (an interpretation supported by his frequent reaching for reinforcers). Such a history may have prevented control by, or displaced, more immediate automatic reinforcement contingencies. This might be the case if, for example, stimuli related to past reinforcement (e.g., hunger and the sight of cake) were more salient for Daniel than the auditory stimuli related to the pairings, either adult- or self-produced. An EO component may be salient given even a brief or weak DR history, and could have obscured other effects (e.g., blocking) that may have prevented conditioning of the target stimulus in the first place. Thus, a competing DR history could have accounted for both increases (EO) and decreases (extinction) in responding during SSP even though the sound itself had not acquired any conditioned reinforcing value. However, these respective influences are difficult to isolate in speech studies since, reinforcement history notwithstanding, the auditory stimulus produced by the vocal response is not masked and, therefore, is presumably salient to some degree.
A third interpretation for Daniel’s loss of SSP effects is habituation to the auditory stimulus produced with each vocalization. Several issues are pertinent. With respect to initial learning, repeated presentations of the target syllable prior to delivering the paired reinforcer may have compromised the ability of the stimulus to become conditioned (Pear, 2001). That is, Daniel may have stopped attending to the stimulus by the time it was actually paired with the reinforcer; thus the pairing, unobserved, would have no conditioning influence. In terms of their automatic reinforcing value, it is possible that Daniel’s repeated vocalizations of the target syllable were weakened (i.e., not maintained) due to the decreasing strength of the auditory stimulus produced, either through habituation to the stimulus or through measurable decreases in certain parameters (loudness, accuracy) of the acoustic signal. The plausibility of this interpretation is augmented by the fact that Daniel’s speech was weak in general (e.g., low volume, articulation errors). Therefore, it is reasonable that these auditory stimuli, weak to begin with, may have had associated weak effects in terms of their ability to strengthen and maintain novel vocal responses.

Although speculative, the strength of a competing history of differential reinforcement may be further indicated by the rapid decreases in response levels for both Daniel and Madison during the No DR condition when reinforcement was available on a response-independent (FT) basis and, comparatively, by the rapidity with which these responses were brought to strength upon reinstatement of the DR
contingency. Socially mediated responding may decrease when maintaining reinforcers are available noncontingently (e.g., Vollmer, Iwata, Zarcone, Smith, & Mazaleski, 1993) as they were during No DR, but responses maintained by automatic reinforcement should remain at some strength to the degree that their response products (in this case, speech stimuli) have absolute value. The purpose of the No DR condition was to determine response maintenance when DR contingencies were absent and reinforcers were present noncontingently. Responding under these conditions would indicate some mechanism (i.e., automatic reinforcement) other than DR since the only stimulus additionally available through responding would be the auditory stimulus presumably conditioned through pairing. Since responding was not maintained during No DR for Daniel and Madison, apparently the conditioning effects of pairing, initially demonstrated in SSP, dissipated before (or during) No DR such that the auditory stimulus no longer had absolute value as an automatic reinforcer. Previous researchers (e.g., Sundberg et al., 1996; Yoon & Bennett, 2000) also have noted the temporary effects of pairing but the current study allows closer analysis of these effects, as it is the first to demonstrate comparative value of the paired speech stimulus under conditions of automatic, DR, and response-independent contingencies. In doing so, the fragile nature of the newly acquired response is not only clearly illustrated but can be evaluated in terms of specific contingencies most likely to strengthen and maintain the response. It is clear, for example, that all targets, initially brought to strength through pairing (i.e., automatic reinforcement), were
sensitive to non-automatic contingencies and quickly came under their control (i.e., DR, No DR). Thus, practitioners can maximize the likelihood of response maintenance by carefully monitoring acquisition data and implementing appropriate contingencies before pairing effects are lost.

The relative strength of DR over automatic reinforcement, or vice versa, for individuals with weak speech repertoires remains an empirical issue. That is, superiority of particular reinforcement contingencies (i.e., social vs. automatic) cannot be assumed; these arrangements likely exist on a continuum and thus are dynamic. There is some indication that automatic reinforcement, although a transient influence on target responding for two participants in this study, was a strong source of control for on-going (extraneous) vocal behavior with all participants both before and throughout the experiment. All children regularly emitted incidental vocal behaviors for which no DR contingencies could be identified and, thus, responses appeared maintained solely by automatic reinforcement. For example, Madison often repeated double-you double-you when no corresponding stimulus was evident (i.e., the letter/sound “W”) and Daniel’s free-play speech contained frequent repetitive patterns of nonsense vocalizations. Control by automatic reinforcement is similarly indicated by Joshua’s target response maintenance during No DR (see Figure 1) although it is possible that extending his topography 2 SSP condition would have revealed a pattern similar to Daniel’s (i.e., loss of effect) since a downward trend was evident between sessions 33 and 43 (Figure 1, lower panel).
For Joshua, response variability throughout the experiment and the sluggish pattern of acquisition seen in topography 2 SSP, combined with the paucity of his pre-experimental vocal play and lack of verbal behavior, all suggest that speech stimuli in general were not yet highly valuable. Furthermore, even when conditioning effects were shown for one topography (Figure 1, upper panel), these did not generalize easily to another topography undergoing identically conducted pairings (Figure 1, lower panel). Response irregularities such as these are typical of the sporadic and slow-to-establish responses often seen clinically in children with weak overall speech repertoires, underscoring the supposition that a critical mass of speech responses (i.e., rich vocal play) is required for acquisition of more complex vocalizations or, in fact, for varied stimuli (i.e., speech) to function more broadly as effective conditioned reinforcers.

Speech-learning difficulty may be particularly evident with children whose remedial speech programs have intensively targeted operant (e.g., echoic, mand) responses but have lacked commensurate efforts to establish vocal play (as automatically-reinforced behavior). Although this training discrepancy would be irrelevant for individuals with typically developing repertoires, it could be a key omission for those who have failed to acquire integrated articulatory and prosodic patterns (e.g., Joshua), or who lack spontaneous language learning and only speak under tightly controlled, and highly structured conditions in which contingencies are tightly managed (e.g., Daniel and Madison). With such children, any delay to
reinforce early vocal attempts may negatively impact subsequent response efforts.

Daniel's response loss during topography 1 SSP (and later, during No DR) highlights this possibility. In SSP, each target presentation was followed by reinforcer delivery, but only if no response occurred between these events. Echoing the model imposed a correction delay. Similarly, during No DR, responding just before a scheduled reinforcer delayed its delivery. Thus, for Daniel and to some extent Madison, both of whom were sensitive to social reinforcement contingencies evidenced by a few (albeit weak) verbal operants, such delay or removal of reinforcers could have functioned as aversive stimuli that punished subsequent responding. Practitioners, therefore, need to carefully engineer and quickly capture the separate or combined effects of automatic and socially mediated reinforcement contingencies for early speech learners and recognize that the influence of either may not be discretionary but, in fact, may be required for the other to exert optimal effect.

A few additional observations with respect to stimulus value and salience are noteworthy. In terms of the paired reinforcer, it is possible that, in an attempt to avoid satiation by using a variety of reinforcers during SSP, the pairing effect was weakened. This could occur if EOs (i.e., reinforcer value) varied during SSP but stimuli were not changed to quickly capture these fluctuations or, conversely, if motivation remained high for a particular item but alternative stimuli were substituted during pairing. However, Joshua's sustained target performance during No DR suggests that paired stimuli were indeed effective to condition novel speech syllables.
as conditioned reinforcers, thus supporting the notion that, at least for this participant, paired stimuli maintained their value.

A separate consideration must be made for the value of the conditioned (arbitrary) stimulus; that is, the paired speech syllable. For Madison and Daniel, there is little evidence that these stimuli alone were sufficient to maintain responding after pairing ceased. Increased responding during DR may have been partially a function of the automatic reinforcing effect of a newly conditioned auditory stimulus, but data analysis from the No DR condition makes this an unlikely possibility. In the absence of DR contingencies, responding for both participants decreased dramatically, returning to near-zero levels within a few sessions. Thus, it appears that the automatic reinforcing value of the previously paired stimulus was not adequate to maintain target vocalizations absent other consequent stimuli of presumably greater value.

Another problem related to value and salience is that the stimulus to be conditioned through pairing is not available as a conditioned reinforcer to the same degree when it is a product of the child's response. Undoubtedly there is a stimulus decrement occurring from the difference between the adult-produced auditory signal and the child's self-produced signal (in terms of prosody, intonation, resonance, and other characteristics; see Moerk, 1990). Any inhibitory effect of this decrement on conditioning is ultimately an empirical issue, but it may be an important one in understanding SSP more fully, and speech acquisition with this population in particular.
This study implemented a number of procedural modifications that were not included in previous SSP investigations with speech-delayed children. As such, they represent the first empirical attempt to specifically enhance stimulus salience for greater SSP treatment effect. Initially it was important to maximize the likelihood that stimuli during pairing would impact receptors. An orienting prompt (e.g., *look*) preceded each trial during baseline and SSP. Although all participants emitted observable orienting responses to this stimulus, they did not do so consistently, thus bringing into question whether subsequent stimuli (pairings) were sufficiently salient for conditioning to occur. It is interesting to note, however, that Madison nearly always emitted the orienting prompt as part of each target response thus indicating that not only did the subsequent pairing stimuli have salience but also that the observing prompt was equally obvious and came to function as an integral part of the conditioned unit. Because Madison imitated this vocal prompt, it was replaced with a clicker noise, an alteration that proved effective at eliminating the prompt’s inclusion into subsequent target responses. In addition, since the clicker was a non-vocal signal, it may have further distinguished the target model, as the more relevant (speech) stimulus, from the orienting prompt.

To emphasize the correlated stimuli used during pairing (S+) trials and to differentially establish paired syllables as conditioned reinforcers, non-paired syllables (S-) were randomly, but equally, interspersed during baseline and SSP sessions. This maximized salience of the paired stimulus by changing the ratio of
reinforcer\(^6\) probability (Gibbon & Balsam, 1981) and, thus, increased the likelihood that important stimuli would be more potent (Dinsmoor, 1995a) and participants would disregard, relative to S+, the unimportant stimuli (i.e., those that were not followed by reinforcers). This modification proved effective, as all participants demonstrated increased target over non-target responding during SSP sessions, although, initially, S+ and S- may have been equally salient for all participants (see baseline for Daniel topography 1, Figure 3 and Joshua and Madison topography 2, Figures 1 and 2, respectively). Furthermore, more frequent non-target responding by Daniel, and to a lesser extent Madison, suggests that stimulus salience may have been different for these participants than for Joshua, who rarely emitted the non-target response. This may have been due to his relatively weaker vocal skills compared to Madison and Daniel such that fewer responses in general were available in his repertoire and, of these, non-target responses were even less likely. However, for all participants, both target and non-target syllables were comprised of phonemes evident in the pre-experimental repertoire. Therefore, although non-target syllable emission may have been unlikely for Joshua, it was not impossible, given accessibility of these responses in his recently demonstrated repertoire.

Even though all participants clearly attended to experimental stimuli, the source of stimulus control was difficult to determine in some instances. This is

\(^6\) It should be noted that the term *reinforcer*, in this instance, refers to the paired stimulus and is not used in the typical sense since no response is required for this stimulus to be delivered.
illustrated by the triadic pattern of responding (e.g., beh beh beh) that children generally emitted. Given that pairings were presented in triads, there is no reason to expect an altered (single) response since the pattern, as originally paired, would comprise the conditioned reinforcer. It is noteworthy, however, that the same response form persisted into DR even when reinforcer delivery, to the extent possible, followed a single response (e.g., beh). Perhaps the logistic difficulty of delivering some types of reinforcers (e.g., bubbles) between one response and the next contributed to this pattern, thereby precluding clear discrimination of the minimal response-reinforcer unit. Extending the DR condition and/or allocating reinforcer delivery more precisely may have informed this issue. It is not unusual with this population of children that responding often appears to be under idiosyncratic stimulus control, but this merely points to the instructional imperative for practitioners to manage the influence of secondary stimuli that may be salient in the learning environment and to highlight those that are most immediately relevant.

A final modification to the typical SSP procedure involved use of a variable intertrial interval (ITI), implemented to increase attention to unpredictable stimuli (Gibbon & Balsam, 1981). ITI values varied from 5 s to 30 s, but this may not have been long enough to optimize stimulus salience. In the case of edible reinforcers, for instance, it is possible that even with maximum ITI values, food taste dissipated slowly and remained available as a reinforcer during non-paired trials, making less observable the difference between S+ and S-. Edibles were used in varying degrees
across participants and it may be that performances varied according to (compromised) pairing effects related to this issue. For Madison, edibles were the reinforcer of choice. In contrast, edibles were rarely used with Joshua and, during Daniel’s sessions, edibles were selected about half the time. Since daily sessions were preceded by brief preference sampling, it is likely that relevant EOs were effectively addressed, thus maximizing pairing effects regardless of reinforcer type. However, it is possible that, in sessions when edibles were utilized, less discrimination was achieved than during sessions in which reinforcer presence and absence could be more discretely controlled.

Limitations

Research that seeks to evaluate treatment effects on specific speech responses may have limited generality due, in part, to the difficulty of providing standard vocal models and accurately identifying and reinforcing phoneme response topographies with minimal acoustic differences. To the degree that researchers can make these fine discriminations, treatment effects may be expected with other populations or response topographies. This study required extensive training to standardize prosodic features (e.g., pitch, loudness, intonation) of SSP presentations among various individuals administering treatment and to ensure both adequate IOA and treatment integrity.

A further limitation of this study concerns the initial vocal play observation period. Existing vocal repertoires were assessed in a 30-min free-play setting for each child. This may not have produced a representative sample of overall vocal play
skills. In addition, participants were observed at solitary play (except for a short period in which Joshua's mother was present) so little information was available from this source about existing verbal operant repertoires. Longer sampling periods may yield different values than those obtained in this investigation in terms of topographic frequency and diversity and such information may provide a greater pool from which to select effective pairing targets.

In addition, the delivery of reinforcers used in the experiment was not controlled outside treatment sessions. Since access to reinforcers absent the paired auditory stimulus would have weakened the conditioning effect of SSP by eliminating the contiguous and conditional relation between these two stimuli, researchers requested that caregivers and parents eliminate use of paired reinforcers for the duration of the study. However, no monitoring was implemented to determine adherence to this request although compliance was verbally reported to researchers.

**Future Research**

Most children with ASD require speech instruction yet mastery remains elusive for many. Perhaps speech as an environmental stimulus is too obscure, even when self-produced, and thus not easily conditioned as a reinforcer for individuals who, in general, unreliably attend to changes in the environment. Attending to relevant stimuli is an essential behavioral response for conditioning to occur, for establishing operations to exert influence, and for stimuli to acquire discriminative functions. That is, regardless of how reinforcing any particular stimulus might be, it
cannot exert influence over responding if not actively observed (Michael, 2004). In terms of speech acquisition, richly varied and frequent vocal play, upon which more complex verbal behavior could develop, may be less likely for individuals with selective or inconsistent attending. Therefore, it would be important to identify variables that increase stimulus control related to attending responses and subsequent vocal responses to facilitate acquisition of functional speech. In addition to the current line of research, others have begun to identify language-related variables (e.g., Charlop & Milstein, 1989; Halle, Baer, & Spradlin, 1981; Koegel, Dunlap, & Dyer, 1980) but the task is not complete and much of this work has addressed only receptive language skills (e.g., Grindle & Remington, 2002) or sign language (Clarke, Remington, & Light, 1986) in contrast to vocal play and speech.

Another area for future research is the issue of isolating SSP’s effects on the absolute value of a speech stimulus as a conditioned reinforcer. Using a strongly enhanced SSP procedure, this study demonstrated the highest and most reliable session rates to date, yet newly acquired targets were so fragile that they did not continue beyond the experimental setting even though other vocal behavior that appeared automatically maintained was evident. Furthermore, in some cases, response variability compromised a straightforward conclusion that the conditioned stimulus functioned as a strong reinforcer for vocal responses. This issue might be addressed by evaluating SSP effects via analog responding in which a non-laryngeal response (e.g., button press) produces the same auditory stimuli as those paired with
reinforcers during SSP. Although reinforcement technically might be considered mediated, in the sense of the button being programmed, clear conditioning effects would support an interpretation of absolute value for the auditory stimulus as an automatic reinforcer.

However, if non-vocal responses, like responses produced by the vocal musculature, are not easily brought under the control of speech stimuli as conditioned reinforcers, this may elucidate the susceptibility of speech vs. other stimulus forms (e.g., visual) as stimulus classes conditionable through SSP. Given equivalent conditioning (pairing) histories in a human organism with intact receptor systems, it is difficult to see how one stimulus class would obtain over another except in the context of relevant establishing operations. It is possible, though, that for some learners, repertoires are established more easily given particular sensory input (i.e., auditory vs. visual). This possibility, though, would not simplify the speech-teaching task since, in all cases, speech-related auditory stimuli must achieve strong conditioned value, if the responses that produce them are to be selected into the repertoire in any meaningful (i.e., functional) way.

Optimal conditioning procedures are not the only unknown variables in augmenting speech acquisition when severe delays exist. Prerequisites for the repertoires that comprise speech are unclear, apart from normal physical systems of hearing and speech production. The current study suggested differential SSP effects based upon existing vocal play strength, but this relation is not well defined.
Furthermore, other contributing requisite repertoires were not considered here. In the recent line of research, SSP has been investigated with a total of 18 children, yet only 3 were described as having normal speech/language repertoires for their age. Therefore, future studies could extend the analysis by including participants with more clearly discrepant repertoires (i.e., severely impaired vs. typically developing) in vocal play, motor imitation, or source-related reinforcement histories (e.g., automatic or socially mediated) to determine any differences in magnitude and durability of the response, or overall speed of acquisition.

SUMMARY

This study was conducted to further inform the efficacy of stimulus-stimulus pairing to establish speech sounds as conditioned reinforcers for vocalization responses that produce those auditory stimuli. The results indicated an enhanced SSP procedure was effective and reliable across all paired topographies for participants, each of whom had severely delayed speech. Clinically, SSP speech studies are useful to the extent that they explicate a viable procedure for improving early speech acquisition in those with weak or non-existent repertoires. Habilitation is impeded when children emit few or inconsistent verbal responses and little vocal play; thus, a treatment such as the improved SSP procedure reported here offers therapeutic direction for practitioners and informs future research to augment the incipient literature in establishing conditioned reinforcers with humans, and the conditioned value of speech in particular.
APPENDIX A

Pilot Study
Pilot Study

A preliminary study consisting of 2 SSP experiments was conducted prior to the primary investigation reported here. As a result of this evaluation, and based upon recommendations from the researcher's dissertation committee, certain procedures in the pilot were modified, resulting in the research protocol used in the main study. The purpose of the pilot was twofold: (1) to provide a pairing history with respect to particular speech syllables and to observe the effect of these pairings on the frequency of speech responses that produced those auditory stimuli; further, in the event that target responses did not increase following pairings, (2) to evaluate more specifically the value of those (paired) auditory stimuli as conditioned reinforcers.

EXPERIMENT 1

Method

Participant and Setting

One participant completed the pilot study. Conner was 3 years, 10 mos old and had a diagnosis of ASD. Contiguous brief (5-10 min) sessions, lasting about 1 hour total, were conducted in Conner's home, 3-5 days per week. In addition to the experimenter's video equipment, the room contained a desk and small table/chairs, sofa, TV, and assorted low preference toys.

Preliminary Assessments

The same battery of speech and language assessments was administered to Conner as to participants in the primary study. On the KPST, Conner emitted no
imitative (i.e., echoic and oral motor) responses. He received a score of 2 on the vocal imitation section of the BLA, indicating informants reported Conner could repeat a few specific sounds or words. No information was available regarding the specific conditions under which these verbal responses were emitted (e.g., prompted, instructional vs. "natural" setting). The PPVT-III yielded no basal score so this test was not given. An alternative language measure (REEL-3) showed receptive and expressive language ages of 6 months and 5 months, respectively. In the 30-min free operant speech inventory, 7 consonants and 4 vowels were emitted; however, there were no recognizable words in this speech sample.

**Stimulus Preference Assessment**

Items to be used as paired stimuli (i.e., reinforcers) in the study were identified in the manner previously described.

**Target Response and Definition**

Two targets, *ma* and *push*, were chosen from the 11 available phoneme topographies identified in the preliminary speech inventory. These responses were counted according to the same criteria used in the primary investigation.

**Interobserver Agreement**

Two independent observers manually recorded session data during a minimum of 75% of all sessions. IOA was calculated for target responses as described earlier. Mean IOA was 83% (range, 64% to 100%) for Target 1 and 98% (range, 71% to 100%) for Target 2.
Procedure

Experimental design. A concurrent multiple-baseline design across 2 phoneme topographies was combined with an ABC design to evaluate the effects of SSP on the frequency of post-pairing vocalizations. Non-target responses were plotted as a separate baseline that served as a constant-series control (Hayes, Barlow, & Nelson-Gray, 1999).

Baseline. Each baseline session consisted of 3 (pre-pairing, pairing, post-pairing), 3-min periods of observation only. During this 9-min period, Conner played alone in the session setting. Experimenters interacted with him only to the extent necessary to maintain his safety.

Control. Control sessions consisted of the same 3-min periods (pre/post and pairing) as baseline. However, during pairing segments, stimuli selected for SSP (speech syllable plus reinforcer) were presented non-contiguously, separated by 20 s. Speech stimuli were presented in an identical manner to those in SSP sessions (see below). This condition controlled for differential effects of the paired stimuli by isolating the individual effect of either stimulus on subsequent target vocalizations.

Stimulus-stimulus pairings. During each 20-trial session of the pairing condition, the experimenter repeated a syllable, 1 per second, in triads (e.g., ba ba ba) then immediately delivered a reinforcer; these were available for 10 s or, in the case of edibles, until consumed at which point a 10-s ITI occurred. Data were collected on target and non-target vocalizations during 3-min pre- and post-pairing observation.
periods in which Conner played alone with low-preference toys and activities. As in the other conditions, interaction between Conner and the experimenter was kept to a minimum.

*Independent Variable Integrity*

An independent observer assessed treatment integrity for 77% of baseline, control, and SSP sessions in Experiment 1. Observations were made either during sessions or later from videotaped recordings. Trials were scored as completely correct or incorrect. The number of correct trials, divided by the number of correct plus incorrect trials, and multiplied by 100, yielded an IVI percentage score that was averaged across sessions (mean treatment integrity percentage). Baseline trials were correct if no syllables and no reinforcers were presented. Control trials were correct if (a) 3 syllables were presented within 5 s, (b) syllables and reinforcers were delivered, alternately, every 20 s, and (c) no other tangibles were delivered. A pairing segment trial was correct if (a) 3 syllables were presented within 5 s, (b) reinforcers were delivered within 5 s after the last syllable was presented, and (c) no other tangibles were delivered. Mean treatment integrity was 100% for both targets.

**RESULTS AND DISCUSSION**

For all conditions, Conner emitted no target vocalizations (with the exception of Target 1 data point 17) during pre- or post-pairing observations (Figure 4, upper 2 panels). Non-target vocalizations (lower panel) occurred at similar rates throughout (5.1 pre-pairing RPM across all conditions compared to 5.2 RPM overall during post-
pairing observations). Thus, although Conner emitted frequent vocalizations, there was no differential effect of the previously paired syllables on responding.

If pairing had effectively established speech stimuli as conditioned reinforcers for speech responses, increased vocalizations as a result of those pairings would be
expected. Since this did not occur, it is uncertain whether these particular stimuli (i.e., \textit{ma, push}) were indeed conditioned as automatic reinforcers. If not, perhaps the procedure itself did not effectively highlight relevant stimuli to condition the paired stimulus (i.e., speech sounds). Alternatively, the pairing procedure may have been effective but the evaluation method may have lacked sensitivity to measure the behavior change. Experiment 2 was thus conducted to evaluate the conditioned reinforcing value of paired vocal stimuli (i.e., speech syllables) by circumventing the vocal response system and substituting a response requirement whose topography was already strong in the repertoire (i.e., button pressing). This had the added advantage of introducing a more observable response.

**EXPERIMENT 2**

**Method**

**Participants and Setting**

Participant, settings, reinforcers, and the recording system were identical to those in Experiment 1.

**Materials**

Three colored 5-inch plastic buttons were used as the response operand. Each button was programmed to produce a specific auditory stimulus via a built-in speaker. Depressing the yellow (target) button resulted in the auditory stimulus \textit{too}; the blue (non-target) button produced \textit{we}, and no consequences were programmed for the red button. Depressing red resulted in no auditory stimulus other than the momentary...
activation click common to all the buttons. Volume level on the speakers was set at
the highest output level (approximately 65 dB) to ensure salience of the auditory
stimuli. Daily equipment checks were conducted to ensure proper functioning.

At any time during Experiment 1, Conner could have emitted vocal responses
that resulted in auditory stimuli. Therefore, during intervention phases of Experiment
2, where responding analogous to laryngeal (and other articulatory) movement was
under investigation, the operanda were similarly available in the environment and
programmed such that depressing a particular button produced either the target
(paired) syllable, a non-target (non-paired) syllable, or no programmed speech
syllable. In other words, as with his own laryngeal musculature, Conner could access
the buttons at any time to emit a given response. Buttons were placed on a table in
front of Conner approximately 28 cm from his torso and button-pressing responses
were recorded. As in Experiment 1, toys and free-play were available in the
environment.

Response Definition

The target response was defined as any motor response that effectively
depressed a button to the depth required to produce the programmed auditory
stimulus or, in the case where no auditory stimuli were programmed on a particular
button, to the depth necessary to produce the characteristic “click” sound inherent in
activating the device. Each motor response that produced an auditory stimulus as
described was counted as one response and responses that did not produce such
stimuli were not counted. Motor responses in which there was no direct contact between Conner’s body and the button surface were not counted (e.g., hitting button with book). Button presses made during any transport of buttons were not counted.

Interobserver Agreement

Two independent observers manually recorded data during 75% of all sessions. IOA on frequency of responding on target and non-target buttons was calculated using the block-by-block method by dividing the smaller frequency of button presses recorded in each 30-s interval by the larger frequency averaged across sessions and multiplied by 100 to yield a percentage of agreement. Mean agreement across sessions was 98% (range, 71% to 100%).

Procedure

Experimental design. An AB design was combined with a 2-tiered constant series control to evaluate SSP effects on the frequency of free-operant button pressing. Responses on target (T), non-target (NT), and control (C) buttons were plotted across separate conditions of baseline and automatic reinforcement of target responding before and after each SSP session.

Pre-intervention assessment. Prior to baseline, an assessment period was conducted to evaluate the extent to which, in the absence of programmed consequences for responding, Conner pressed the various buttons. This assessment controlled for biased responding as a result of previous experience or preference for or aversion to pressing buttons in general (e.g., on toys). The button arrangement was
placed in front of Conner (relative position was randomly varied) and he was instructed to “press all the buttons.” The buttons were in the “off” position; thus, there were no programmed consequences for pressing any of the buttons. That is, when any button was pressed, the only auditory output was the “click” produced by depressing the button itself. If no responding occurred, Conner was prompted verbally and visually to press each button, in random order, 3 times, or fewer if he began to respond independently. If he did not press buttons after 2 verbal/visual prompts, he was physically prompted to make a response. This procedure ensured that Conner could and would press each button. Results showed immediate unprompted button pressing and no preferential responding on any specific button.

Pre-baseline training. Buttons were in the “on” position, having been pre-programmed with target and non-target syllables. Prior to each baseline session, forced-choice training trials were conducted to ensure Conner contacted the contingency in effect for each button. During these training trials, Conner was prompted, in random order, to press each button three times. Prompting procedures were identical to those in pre-intervention assessment. To increase contact with the contingency for target responding, 3 additional training trials were conducted on the target button. For identification purposes, buttons were coded on data sheets as target (T), non-target (NT), and control (C). Pressing the T button resulted in the target (i.e., too) auditory vocal stimulus. Pressing the NT button produced the syllable we. Pressing the C button produced no auditory stimulus except ambient noise.
characteristic of the device itself. Availability of the NT button controlled for any auditory vocal stimulus as a reinforcer for button pressing, whereas availability of the C button controlled for button pressing as a reinforcing activity regardless of its consequences (i.e., auditory stimulus). Thus, in the subsequent intervention condition (pairing), differential responding on the T button could be more confidently interpreted as a result of the conditioned reinforcing aspects of the particular auditory stimulus programmed on that particular button.

**Baseline.** Each baseline session consisted of 3 segments: 3 min pre-pairing, 3 min pairing (in baseline, this was observation only), and 6 min post-pairing. Responses were observed and recorded during the pre- and post-pairing segments, during which Conner had access to toys and free-play. Within-session (pairing segment) data were not recorded. Interactions with the experimenter occurred only if necessary to maintain Conner’s safety (e.g., prevent from climbing).

**Stimulus-stimulus pairing.** The purpose of SSP was to determine the extent to which a previously paired vocal stimulus functioned as a conditioned (automatic) reinforcer for an arbitrary motor response as an analog to laryngeal responding. Timings of pairings and reinforcer delivery were identical to those in Experiment 1. Button programming remained constant across sessions (e.g., yellow button produced the target sound) to maximize discrimination but button position was varied randomly from session to session. The experimenter pressed the target button 3 times (*too, too, too*) and immediately delivered a preferred item. If the child pressed any button
between pairings and reinforcer delivery, a 20-s delay was imposed during which the experimenter maintained a neutral position and no pairings or reinforcers were delivered. This procedure controlled for adventitious (indirect) reinforcement of responding on a particular button (e.g., mand training); thus, any intervention effects could be more confidently interpreted as due to the auditory stimulus itself acquiring reinforcing properties. Data were collected on target, non-target, and control (blank) button responding during observations conducted for 3-min pre-pairing and 6-min post-pairing during which Conner had access to toys and free-play. As in baseline, interactions with the experimenter were kept to a minimum.

Independent Variable Integrity

An independent observer assessed treatment integrity for 100% of SSP sessions. Observations were made either during sessions or later from videotaped recordings. Trials were scored as completely correct or incorrect. The number of correct trials, divided by the number of correct plus incorrect trials, and multiplied by 100, yielded an IVI percentage score that was averaged across sessions (mean treatment integrity percentage). A pairing trial was correct if (a) 3 syllables were presented within 5 s, (b) reinforcers were delivered within 5 s after the last syllable was presented, and (c) no other tangibles were delivered. Mean treatment integrity was 100%.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
RESULTS AND DISCUSSION

Before Experiment 2 could be completed, Conner withdrew from the study. Figure 5 shows results available to that point. Conner made no target (upper panel) button presses during baseline. In SSP, overall rate of pre-pairing responding averaged 0.18 RPM; post-pairing target responding only occurred once (session 12) for an overall average of 0.01 RPM. Thus, data indicate that SSP failed to establish the target auditory stimulus (too) as a conditioned reinforcer for the motor response (button press) required to produce the reinforcer. Non-target (middle panel) and control (lower panel) responding also occurred at undifferentiated low rates throughout. Overall non-target rates averaged 0.25 RPM and 0.01 RPM during pre- and post-pairing, respectively. Similarly, rates for responding on the control button averaged 0.08 RPM and 0.01 during pre- and post-pairing, respectively.

GENERAL DISCUSSION

Lack of target responding in this study underscores the difficulty of evaluating the efficacy of the SSP procedure (see Esch et al., 2005). If pairing established speech stimuli as conditioned reinforcers for the vocal (Experiment 1) or non-vocal (Experiment 2) responses that produce those stimuli, increased responding as a result of those pairings would be expected. Since this did not occur, it indicates that these particular stimuli (i.e., ma, push, too) were not conditioned as automatic reinforcers and, furthermore, suggests that speech stimuli in general may not acquire reinforcing value as a result of SSP (although perhaps through some other contingent relation; for
example, mand training). The results of the pilot study led researchers to propose several modifications to the current SSP procedure; these would be designed to maximize the strength of the pairing relation between the established reinforcer and the arbitrary (neutral) stimulus. Thus, several sources of variability could be accounted for and, if weak effects persisted, researchers could then look to other explanations.
APPENDIX B

Research Protocol Clearance
Date: May 5, 2005

To: James Carr, Principal Investigator
    Barbara Esch, Student Investigator for dissertation

From: Mary Lagerwey, Ph.D., Chair

Re: HSIRB Project Number: 05-04-09

This letter will serve as confirmation that, pending receipt of the approval letter from Croyden Avenue School, your research project entitled “The Role of Automatic Reinforcement in Speech Acquisition” has been approved under the full category of review by the Human Subjects Institutional Review Board. The conditions and duration of this approval are specified in the Policies of Western Michigan University. You may now begin to implement the research as described in the application.

Please note that you may only conduct this research exactly in the form it was approved. You must seek specific board approval for any changes in this project. You must also seek reapproval if the project extends beyond the termination date noted below. In addition if there are any unanticipated adverse reactions or unanticipated events associated with the conduct of this research, you should immediately suspend the project and contact the Chair of the HSIRB for consultation.

The Board wishes you success in the pursuit of your research goals.

Approval Termination: April 20, 2006
REFERENCES


