Research in Autism Spectrum Disorders 7 (2013) 997–1011

Contents lists available at SciVerse ScienceDirect

Research in Autism Spectrum Disorders

Journal homepage: http://ees.elsevier.com/RASD/default.asp

The role of joint control in teaching listener responding to children with autism and other developmental disabilities☆

Kaitlin G. Causin, Kristin M. Albert, Vincent J. Carbone *, Emily J. Sweeney-Kerwin

Carbone Clinic, Valley Cottage, NY, United States

ARTICLE INFO

Article history:
Received 27 February 2013
Received in revised form 10 April 2013
Accepted 17 April 2013

Keywords:
Autism
Joint control
Listener behavior
Verbal behavior

ABSTRACT

This study evaluated the effectiveness of a teaching procedure derived from the analysis of joint control in increasing listener responses for three children with autism using a multiple probe design across participants. One nonvocal and two vocal children with autism were taught to select multiple pictures of items from a large array in the order in which they were requested (e.g., "Give me the ball, cup, and spoon") using the joint control teaching procedure. The effect of these procedures on the emission of accurate selection responses to both trained and novel stimulus sets was measured. The results indicated that listener responses to trained stimuli increased following the implementation of the independent variable and untrained responses across novel stimulus sets also emerged. Implications for designing language training programs for children with autism based on an analysis of joint control are discussed.

Skinner's (1957) analysis of verbal behavior offered an alternative to the prevailing structural conceptions of language in which words and sentences (i.e., formal properties of language) were considered the important units of analysis. Whereas structural accounts emphasized the topography of language (e.g., syntax, grammar, morphemes, mean length of utterance), Skinner's behavior analytic account identified the functional relation between a response and its controlling variables, or the verbal operant, as the important unit of analysis. This behavior analytic account of language suggests important implications for the treatment of children with autism and other developmental disabilities (Sundberg & Michael, 2001) and a growing body of clinical work and research has documented the value of including this taxonomy in language training programs (see Sautter & LeBlanc, 2006 for a review). Much of this literature, however, has focused on the application of Skinner's analysis to teaching speaker behavior, with less work dedicated to a thorough analysis of the contingencies operating on the behavior of the listener (Schlinger, 2008).

1. Listener behavior

Listener behavior did not receive extensive coverage in Skinner's (1957) work mainly because much of listener behavior is not distinguished from other behavior under discriminative control. Possibly due to this lack of attention, cognitive explanations that describe the listener as a "passive receptacle" (Schlinger, 2008, p. 149), "recipient" (Lowenkron, 1998, p. 339), or "processor" (Sidener, 2006, p. 119) of information have persisted. Although Skinner's (1957) analysis emphasized

☆ The authors would like to thank Francesca degli Espinosa for her comments on an early version of this paper. We would like to thank the parents of the participants who took part in this study, the instructors who conducted the sessions, and our colleagues, Christy Winder and Laura Kohlberg Peterhoff, who assisted with training and measuring treatment fidelity.

* Corresponding author at: Carbone Clinic, 614 Corporate Way, Suite 1, Valley Cottage, NY 10989, United States. Tel.: +1 8452670920.
E-mail address: drvjc@aol.com (V.J. Carbone).

1750-9467/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.rasd.2013.04.011

speaker behavior, he did not ignore the listener. Skinner suggested that the control exerted by verbal stimuli was at least partially dependent upon the listener having an existing verbal repertoire of speaker behavior. He stated, "...some of the behavior of listening resembles the behavior of speaking, particularly when the listener understands what is said" (Skinner, 1957, p. 10). Schlinger (2008) extended Skinner's analysis of listener behavior and refined the difference between listener behavior as a repertoire of discriminated operants (i.e., mediation of reinforcement for a speaker) and *listening*. Schlinger asserted that listening is behaving verbally. He stated, "...the behavior of listeners and speakers may be inseparable, especially when we say the listener *listens, pays attention to*, or *understands* the speaker" (p. 148). Schlinger argued that, in fact, listening and speaking may not be functionally different and said, "In other words, the listener also behaves verbally when he or she is said to be listening" (Schlinger, 2008, p. 150). All of this suggests that listening may be predicated upon a complex verbal repertoire that mediates listener responses.

2. Joint control

Consistent with Schlinger's (2008) analysis of verbally mediated listener responding, Lowenkron (1991, 1998, 2004, 2006a) has offered joint control as a conceptually systematic explanation of various complex human behaviors, including listener behavior. Lowenkron (1998) defined joint control as "the effect of two [discriminative stimuli] S^Ds acting jointly to exert stimulus control over a common response topography" (pp. 328–329). Lowenkron (1998) stated:

> Joint control occurs when the currently rehearsed topography of a verbal operant, as evoked by one stimulus, is simultaneously evoked by another stimulus. This event, the onset of joint stimulus control by two stimuli over a common response topography, then sets the occasion for a response appropriate to this special relation between the stimuli. (p. 327)

In other words, one verbal response is simultaneously emitted under two distinct sources of stimulus control. For example, two possible sources of control are: (1) a verbal stimulus that evokes an echoic (vocal imitation) or self-echoic (imitation of the speaker's own verbal behavior) and (2) a nonverbal antecedent S^D that evokes a tact (label). The emission of a single verbal response under two joint sources of stimulus control is a unique event that then exerts control over a third response, typically a selection response or listener response (Lowenkron, 1998). According to this analysis the selection response is mediated by the verbal responses.

An illustrative example of joint control clarifies the nature of this type of event (see Fig. 1). A speaker presents a listener with the task of locating a particular arrangement of stimuli from an array containing various configurations of the same stimuli. The listener is told to find the arrangement that shows an "oval over arrow over rectangle." In order to preserve the stimulus while searching for the correct configuration, the listener, becoming a speaker, emits the response, "oval over arrow over rectangle," first as an echoic (vocal imitation) of the vocal verbal stimulus produced by the speaker and then as a self-echoic (vocal imitation) of the verbal stimulus produced by the listener's own verbal behavior. When shown the comparison stimuli, the listener then directly tests the properties of each comparison encountered relative to the sample provided. Upon the occasion in which the listener emits, "oval over arrow over rectangle," not only as a self-echoic (vocal imitation), but also as a tact (label) of the specified stimulus configuration, a unique form of stimulus control, joint control, occurs. The effect of this single response topography occurring under two sources of stimulus control then sets the occasion for a selection or listener response.

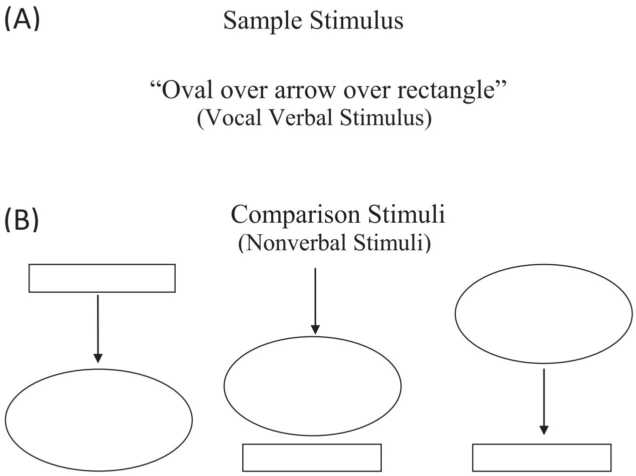

Fig. 1. Example of sample stimulus and comparison stimuli for a joint control listener task. (A) Sample stimulus presented as a vocal verbal stimulus. (B) Comparison nonverbal stimuli.

Sidener (2006) provided a practical example of joint control, highlighting the value of teaching children with autism to repeat instructional commands in order to improve delayed selection responding. Sidener suggested that teaching a child to emit echoic and self-echoic responses following instructions such as, ''Go to the window and get the book,'' when tacts for these items have been previously taught, may facilitate the emission of accurate responses of both going to the correct location and retrieving the item. For example, following the delay provided by walking to the window, if the response, ''book,'' is emitted simultaneously as an echoic (vocal imitation) or self-echoic (imitation of the speaker's own verbal behavior) and as a tact (label), the onset of joint control occasions a selection response.

3. Mediated stimulus selection

The analysis of joint control is consistent with Skinner's (1957) account of verbal behavior and provides a cogent explanation of selection responses within the boundaries of a behavioral science. Furthermore, by identifying the role of verbal mediation, the analysis of joint control provides a plausible interpretation of the occurrence of generalized responding which unmediated accounts are insufficient to explain (Lowenkron, 1984, 1988, 1989, 1991, 1996, 1997, 1998, 2006a, 2006b; Lowenkron & Colvin, 1992, 1995). In unmediated stimulus selection, reinforcement is contingent upon the selection of a specific comparison stimulus, given the sample stimulus presented. Based on a history of reinforcement, the probability of selecting a particular comparison stimulus in the presence of a particular sample stimulus increases. This analysis adequately accounts for stimulus arrangements in which the relation between the sample and comparison has been explicitly arranged but cannot sufficiently explain selection responses to novel stimuli. Conversely, according to the mediated stimulus selection account suggested by joint control, responses are not dependent upon a history of reinforcement related to a particular stimulus or set of stimuli (Lowenkron, 1998; Palmer, 2006; Schlinger, 2008). Instead, the selection response is emitted under the control of the various stimuli within the selection task. In other words, the selection response is determined by the occasion in which one response topography is emitted under two sources of control and is hence a generic event serving as the basis for generalized responding (Lowenkron, 1991, 1998, 2006a).

Practically speaking, the distinction between unmediated and mediated stimulus selection accounts present important implications for the arrangement of language training programs for children with autism. For example, if a child with autism was to be taught to select two items from a larger field (e.g., ''Give me the crayon and the ball'' when presented with a field of 10 items), an unmediated stimulus selection account would require that each set of two items be specifically trained and reinforced. Conversely, according to the mediated stimulus selection account, if the procedures designed to teach this skill included an analysis of mediating responses, the child's responses may be brought under the control of the stimuli produced by the task itself and not the specific sample and comparison stimuli used, thus facilitating generalized responding. For example, if the child is taught to select items from a larger field under the control of a single response topography being emitted both as a self-echoic and as a tact, rather than being taught to emit specific response to specific stimuli, responding under the control of the generic joint control event would increase the likelihood of generalized responding.

In a series of studies investigating the use of sample-coding to teach matching-to-sample, Lowenkron (1984, 1988, 1989) provided empirical support for the role of mediating responses in the occurrence of generalized relational matching. In the first study of this type, Lowenkron (1984) demonstrated that generalized responding was dependent upon the occurrence of a coding response. Participants were taught to code a sample stimulus using an arrow and select a correct comparison stimulus by matching the coded arrow with the comparisons. Results indicated that when the coding stimulus was available, high levels of generalized responding occurred across novel stimuli. However, when the coding stimulus was not available, generalized responding was low. Additional support for the role of mediating responses in generalized relational matching was provided when the methods of Lowenkron (1984) were extended to teaching matching based on size relations (Lowenkron, 1989) and to children with disabilities using hand signs as the coding response (Lowenkron, 1988). In a later study, Lowenkron and Colvin (1995) taught generalized relational matching to preschool children with vocal verbal behavior in the form of tacts and intraverbals (answering questions) as the mediating responses. More recently, Sidener and Michael (2006) replicated Lowenkron's (1984) study by using joint control to mediate stimulus selection based on the relative spatial orientation of stimuli. Lowenkron and Colvin (1992) also conducted a study that expanded the analysis of joint control to include an account of responding that occurs in the absence of a stimulus, or responding when something *is not*.

Two studies have also examined the nature of the rehearsal response during generalized sequencing tasks. Gutierrez (2006) taught English-speaking women generalized sequencing behavior by teaching echoic and tact behavior in Mandarin Chinese as mediating responses. During sequencing tasks, when self-echoic behavior or rehearsal was blocked by requiring the emission of a different response, accurate sequence production decreased. When self-echoic behavior was permitted again, responding improved. DeGraaf and Schlinger (2012) replicated and extended the findings of Gutierrez (2006) and also found when the self-echoic component of the joint control training procedure was prevented, accurate responding deteriorated.

4. Joint control and listener behavior

Much of the empirical work investigating joint control has focused on establishing generalized relational matching or generalized sequencing responses (DeGraaf & Schlinger, 2012; Gutierrez, 2006; Lowenkron, 1984, 1988, 1989; Lowenkron &

Colvin, 1992, 1995; Sidener & Michael, 2006). Less work has examined the role of joint control in establishing listener responding. Lowenkron (2006b) used joint control to mediate the selection of complex objects based on their conformance to a spoken description of the features (i.e., recognition of an object from its description). The first experiment examined the relationship between the acquisition of tact responses and the emergence of generalized selection responding. The results revealed that correct responses to novel stimuli during generalized selection tasks only occurred after tacts of these stimuli were taught. The second experiment examined the maintenance of the sample stimulus through self-echoics. Participants were prevented from emitting self-echoic responses to preserve the sample stimulus and the results showed that rehearsal prevention decreased the accuracy of selection responses. The results of this study suggested that generalized selection responding depended upon both components of joint control.

To our knowledge, the only published applied study which utilized joint control procedures to teach children with autism was conducted by Tu (2006). This study examined the importance of joint control when teaching responses to experimenter vocal requests to both vocal and nonvocal children. Across two experiments the effect of joint self-echoic and tact responding on the emergence of untrained selection responses was examined. Participants were first taught to emit echoic or mimetic (imitation of hand signs) and tact responses to a set of four stimuli. Selection responses were then probed. When participants failed to emit correct responses, a joint control training procedure was employed. Training continued until correct selection responses were occurring within a field of four. A novel set of stimuli were then probed to determine if responding generalized across stimuli. The results of both experiments indicated that following tact and echoic training alone participants failed to emit correct selection responses to the novel stimuli. It was only after participant responses were brought under joint echoic and tact control that accurate selection responses occurred. Tu asserted that the results indicated that language training programs that make use of the analysis of joint control and verbal behavior will likely lead to effective and efficient treatment procedures.

In an unpublished dissertation, degli Espinosa (2011) examined the effects of a teaching procedure derived from a joint control analysis on the selection of picture sets composed of color and item combinations for three children with autism. The joint control teaching procedure required that participants simultaneously emitted self-echoic and tact responses prior to selecting pictures of compound stimuli (e.g., a blue car, cat eating). The results of this study indicated that the joint control training procedures utilized lead to correct selection of trained stimuli and generalized responding across novel stimulus sets.

The purpose of this study was to extend the current body of applied research related to using joint control to teach children with autism. Specifically, the purpose of this study was to teach children with autism who emitted limited listener response repertoires to select multiple pictures of items from a large array in the order in which they were requested (e.g., "Give me the ball, cup, and spoon"). By teaching the selection of multiple items from a large set, this study extends the previous literature which has focused on teaching selection of one stimulus or one compound stimulus out of a field. Furthermore, only one published study and one unpublished study related to teaching listener behavior to children with autism using the analysis of joint control could be located within the literature. Given the limited evidence in support of teaching procedures derived from an analysis of joint control for children with autism, additional investigations of this conceptually systematic account of complex language are warranted.

5. Method

5.1. Participants and setting

Three boys, Billy, Cole, and Abe, served as the participants in this study. All participants were enrolled at a private clinic that provided educational services to individuals with autism and other developmental disabilities. Instruction was provided in a one-on-one intensive teaching format in the form of discrete trial teaching interspersed with teaching in the natural environment and was guided by the principles of applied behavior analysis and incorporated Skinner's (1957) analysis of verbal behavior.

Billy was a 6-year-old nonvocal male diagnosed with autism whose primary means of communication was manual sign language. The inclusion of a nonvocal child with autism who used manual sign language as his primary form of communication may suggest the benefit and utility of these procedures with children who use alternative communication systems. He attended the clinic five days per week for a total of 15 h of instruction per week. During the Verbal Behavior Milestones Assessment and Placement Program (VB-MAPP, Sundberg, 2008) that was administered just prior to beginning the study, Billy's mand repertoire fell within the 18- to 30-month developmental range (i.e., Level 2). Program data indicated that Billy had mastered 198 mands. All mands required formal training and many teaching trials; however, mands were typically maintained once acquired. Billy's mand repertoire consisted of mainly one word manual signs, typically nouns (i.e., items) and a few verbs (i.e., actions). Billy did not mand solely under the control of motivating operations (i.e., mands were typically prompted by the presence of the item or instructor provided physical/gestural prompts), with unfamiliar persons, or mand for information. Billy's tact repertoire also fell within the 18- to 30-month developmental range. Program data indicated that Billy tacted over 266 items including reinforcers, common objects, and pictures of common objects. Once taught a single exemplar of a tact (e.g., tacting a picture of a spoon), responding typically generalized across additional exemplars (i.e., other spoons that were not explicitly taught). Billy did not tact parts and features of objects, actions, or tact by feature, function, or class. Billy did not acquire tacts unless they were formally trained and all required many teaching trials.

Billy's listener repertoire also fell within the 18- to 30-month developmental range. Program data indicated that Billy had acquired 11 listener skills. Billy performed some motor actions on command, responded to his own name, and attended to a speaker's voice. He was beginning to select items/pictures from a field, however, he only did this with singular items/pictures (i.e., Billy did not select multiple items from a field). Billy's motor imitation repertoire fell within the 18- to 30-month developmental range. Program data indicated that Billy had acquired over 249 motor imitation skills. He imitated gross and fine motor movements, motor movements with objects, and engaged in some spontaneous motor imitation. Billy sometimes had difficulty imitating fine motor movements and imitating complex movements (e.g., moving hands in a circular motion while maintaining a certain hand/finger positioning). Billy did not demonstrate any intraverbal skills as measured by the VB-MAPP (Sundberg, 2008), however, when asked by an instructor to sign a word (e.g., ''Sign cat''), Billy would respond intraverbally by signing the word (e.g., signing cat). Billy's echoic repertoire fell within the birth- to 18-month range (i.e., Level 1). Program data indicated that he acquired 30 echoic skills. Echoics were limited to some speech sounds, some combinations of speech sounds/blends, and a few word approximations. Billy's social behavior also fell within the 18- to 30-month developmental range. He visually tracked and showed interest in other people's movements (only adults) and spontaneously engaged in parallel play near other children. He did not show interest in his peers, mand from peers, or respond to mands from peers. He did not spontaneously imitate peers. Billy engaged in low levels of problem behavior (typically less than five episodes per day) and this problem behavior typically did not interfere with instruction.

Cole was a 14-year-old vocal male diagnosed with pervasive developmental disorder and attention deficit hyperactive disorder. He attended the clinic four days per week for a total of 8 hr of instruction per week. During the VB-MAPP (Sundberg, 2008) that was administered just prior to the beginning of this study, Cole's mand repertoire fell within the 18- to 30-month developmental range. Since Cole typically acquired new mands without formal training, program data were not recorded on the number of mands acquired. Cole manded for missing items, emitted multiple component mands (e.g., orange tic tac vs. white tic tac), and manded spontaneously (i.e., manded under the control of the motivating operation alone without the item being present). Cole manded using multiple words and noun/verb combinations. Cole did not mand for information, mand using adjectives, prepositions, or adverbs, give instructions on how to participate in an activity, or mand for others to attend to his own intraverbal behavior. Cole's tact repertoire fell within the 30- to 48-month developmental range (i.e., Level 3). Program data indicated that Cole tacted over 1465 items including reinforcers, common objects, pictures of common objects, familiar people, ongoing actions, pictures of ongoing actions, parts and features of objects, classes of objects, and noun/verb combinations. Cole did not consistently tact prepositions, pronouns, or adjectives and he did not tact using complete sentences. In addition, Cole typically required multiple exemplar training in order to demonstrate generalized responding across stimuli. Cole's listener repertoire also fell within the 30- to 48-month developmental range. Program data indicated that Cole had acquired over 368 listener skills. Cole demonstrated a generalized listener discrimination repertoire across different exemplars, performed specific motor actions on command, followed two-component noun–verb and verb–noun instructions, selected the correct item in a book, picture scene, or natural environment when named, and selected items by color and shape. Cole did not follow instructions involving prepositions or pronouns, select items by adjectives or adverbs, or follow multiple step instructions. Cole demonstrated a generalized motor imitation repertoire. Cole's intraverbal repertoire fell within the 30- to 48-month developmental range. Program data indicated that he acquired over 544 intraverbal skills. Cole completed various fill-in-the-blank phrases related to routines, personal information, and songs. He answered some what, who, or where questions but did so inconsistently. He did not spontaneously emit intraverbal comments, answer questions about a story or video segment watched moments before, or describe events, videos, or stories. He also did not answer ''wh'' questions in a rotating format. Cole's echoic repertoire fell within the 18- to 30-month developmental range. He attempted to echo most sounds, words, and phrases named by an adult when instructed to do so, however, Cole demonstrated difficult with utterances past a certain length and with certain blends. Cole's social skills fell within the 18- to 30-month developmental range. Cole demonstrated some interest in his peers; however, he more typically attended to and showed interest in the behavior of adults. He did not mand to peers or respond to mands from peers nor did he engage in social play with peers. Cole engaged in low levels of problem behavior (typically less than one episode per day) and his problem behavior did not interfere with instruction.

Abe was a 17-year-old vocal male diagnosed with autism. He attended the clinic five days per week for a total of 15 h of instruction per week. During the VB-MAPP (Sundberg, 2008) that was administered just prior to the beginning of the study, Abe's mand repertoire fell within the 18- to 30-month developmental range. Abe's mands were typically one-word utterances, although he did occasionally emit mands that included two words (e.g., orange chip). Abe emitted mands for missing items and manded for others to complete actions. He did not typically emit mands solely under the control of his motivation nor did he mand for information. He did not acquire mands without formal training and training took many teaching trials, however, program data were not recorded on the number of mands acquired at the time the study took place. Abe's tact repertoire fell within the 18- to 30-month developmental range. Program data indicated that Abe tacted over 767 reinforcers, common items, pictures of common items, familiar people, and some adjectives. Abe did not emit two component verb–noun or noun–verb tacts, tact parts and features of items, tact pronouns, prepositions, or adjectives, or tact with complete sentences. Abe did not demonstrate generalized responding across similar stimuli without extensive multiple exemplar training. Abe's listener repertoire also fell within the 18- to 30-month developmental range. Program data indicated that Abe had acquired over 168 listener skills. Abe selected items/pictures of items from a messy array, performed some motor actions on command, and selected the correct item in a book, picture scene, or natural environment when named, however, he did so inconsistently. Abe did not demonstrate a generalized listener discrimination repertoire nor did he follow multiple step instructions. Abe demonstrated a generalized motor imitation repertoire. Abe's intraverbal repertoire fell within the 18- to 30-month developmental range. Program data indicated that Abe had acquired over 495 intraverbal skills. He completed a variety of fill-in-the blank phrases related to animal/object sounds, familiar routines, social games, and songs, and provided some personal information when asked. Abe did not answer "wh" questions, describe events, or spontaneously emit intraverbal comments. Abe's echoic repertoire fell within the 18- to 30-month developmental range. Abe would echo most sounds, blends, and words modeled by an adult, however, he demonstrated difficulty with longer utterances. Abe's social skills fell within the birth- to 18-month developmental range. Although Abe showed interest in others, he typically did not show interest in his peers. He did not spontaneously look at, imitate, or mand to his peers nor did he respond to mands from peers. Abe engaged in low levels of problem behavior (typically less than 5 episodes per session), and typically this problem behavior did not interfere with instruction.

Sessions were conducted every day that the participants attended the clinic. All sessions took place within shared or individual classrooms where an instructor was seated at a table, approximately 1 m long and 0.75 m wide, directly across from the participant. Stimuli and materials were placed on the table in front of the participant. Seven instructors conducted these sessions and recorded data.

5.2. Materials and reinforcer selection

Participants were taught to select pictures from the *Language builder: Picture noun cards* (Stages Learning Materials, 1997). Each card was approximately 13 cm by 9 cm and displayed a color photo of a common object. For each participant, 12 picture cards were selected from a pool of previously mastered echoic/mimetic, tact, and single-step listener selection responses. These 12 cards were used to develop 50 randomly selected stimulus sets for each participant. It is possible that inclusion of the same stimuli across sets may have influenced the outcome of the results due to exposure or practice effects, however, sets containing identical stimuli were relatively low and the order of stimuli always varied across sets. Based on a prebaseline assessment of each participant's skill level, stimulus sets for Billy and Abe consisted of three pictured items and stimulus sets for Cole consisted of four pictured items.

An informal preference assessment was conducted at the beginning of each session in order to determine putative reinforcers for each participant. This assessment was based on the frequency of mands emitted by participants for various items and activities. Those items and activities frequently manded immediately prior to the experimental sessions were selected and delivered following target responses. Additionally, participants were provided opportunities to mand for additional items and activities during the reinforcement period in experimental sessions.

5.3. Response measurement and interobserver agreement

The dependent variables measured in this study were the cumulative number of untrained and trained stimulus sets acquired during probes. Correct responses during probes were defined as selecting all of the stimuli named by the instructor in the same order in which they were stated, within 20 s of the presentation of the vocal stimulus. Responses were followed by a 1 s pause following the delivery of the final item in the set so as not to preclude the possibility that the participant may continue to select additional, irrelevant stimuli. An example of a correct response was when the instructor presented the vocal stimulus, "Give me the car, grapes, and ball," and the participant selected pictures of the car, grapes, and ball in that order, within 20 s, and did not select any additional items during the 1 s pause following the delivery of the last picture card. An incorrect response was defined as: (a) selecting stimuli that did not correspond to the vocal stimulus; (b) selecting the incorrect number of stimuli; (c) selecting stimuli that corresponded with the vocal stimulus but in an order different from that in which they were named; (d) emitting a response beyond the established time criteria of 20 s, (e) initiating a response before the completion of the vocal stimulus; or (f) failing to emit a response. For example, following the vocal stimulus, "Give me the car, grapes, and ball," any of the following participant responses were considered errors: (a) selecting pictures of a shoe, a block, and a dog; (b) selecting pictures of a car, grapes, a ball, and shoes; (c) selecting pictures of grapes, a ball, and then a car; (d) selecting pictures of a car and grapes within 20 s but failing to select the picture of the ball within the interval, (e) selecting the picture of a car immediately after the instructor said "car" and before the vocal stimuli grapes and ball were presented; or (f) not selecting any stimuli from the field.

Probe data on the accuracy of participant responding were collected across conditions. Data sheets included columns for each stimulus set and a column to record "Y" for a correct response or "N" for an incorrect response. Each participant's instructor served as the primary data recorder. A second observer, trained in the data collection procedures, simultaneously but independently recorded measures of the dependent variables for the purposes of determining interobserver agreement (IOA). An agreement between the instructor and second observer was scored when both observers recorded a response as either correct or incorrect. A disagreement occurred when the instructor recorded a correct response and the second observer recorded an incorrect response, or vice versa. Interobserver agreement was calculated by dividing the number of agreements by agreements plus disagreements and converting the ratio to a percentage. For each participant IOA was scored for 29% to 50% of the sessions within each condition. Interobserver agreement was 100% during all sessions except during one untrained probe session with Billy when it was 98%.

5.4. Prebaseline instructor training and treatment fidelity

Task analyses of the instructor's expected behavior during each phase of the study were developed and used as checklists to assess treatment fidelity. Prior to conducting the study, all instructors were trained by the primary researcher to 90% competency (i.e., treatment fidelity) on all procedures. Treatment fidelity was calculated by dividing the number of target instructor responses emitted by the number of target responses emitted plus the number of target responses incorrectly or not emitted and converting the ratio to a percentage. Treatment fidelity was scored for 29–50% of the sessions within each phase for each participant, during all sessions when IOA was scored. Average treatment fidelity for Billy was 98% within each phase. Average treatment fidelity for Cole was 98% during the baseline phase, 97% during the joint control training phase, and 99% during untrained probes. Average treatment fidelity for Abe was 99% within each phase. If at any point an instructor's treatment fidelity score fell below 90%, a booster training session for the instructor was conducted until a 90% competency score was achieved. This occurred on three occasions for Billy's instructor in the training phase. This also occurred on two occasions for Cole's instructor in the training phase.

Two researchers also recorded measures of treatment fidelity simultaneously but independently for the purpose of calculating IOA on treatment fidelity. An agreement between the researchers was scored when both observers scored a target instructor response as either correctly emitted or incorrectly/not emitted. A disagreement occurred when one researcher recorded an expected response as correctly emitted and the other researcher recorded an expected response as incorrectly/not emitted, or vice versa. Treatment fidelity IOA was calculated by dividing agreements by agreements plus disagreements and converting the ratio to a percentage. Treatment fidelity IOA was scored during 25–50% of the sessions during which treatment fidelity was scored across conditions for each participant. Average treatment fidelity IOA for Billy was 92% for baseline, 98% for joint control training, and 99% for untrained probes. Average treatment fidelity IOA for Cole was 100% for baseline, 97% for joint control training, and 100% for untrained probes. Average treatment fidelity IOA for Abe was 99% for baseline, 99% for joint control training, and 100% for untrained probes.

5.5. Experimental design and conditions

A multiple probe across participants design (Horner & Baer, 1978) was used to evaluate the effectiveness of joint control training on teaching listener responding. This experimental design was selected to reduce the possibility that repeated exposure to untrained stimulus sets during the baseline condition would confound the effects of joint control training.

Baseline sessions were scheduled in accordance with the logic of multiple probe design. The baseline condition for the first participant continued until stable responding occurred. The independent variable was then implemented with this participant while baseline measures continued for the other two participants. When simultaneous stability was observed in the joint control training data for the first participant and the baseline data for a second participant, the second participant was subjected to the joint control training condition. The same criteria were used to move the third participant from baseline to training.

The joint control training condition consisted of trained probes, five joint control training trials, and untrained probes. The joint control training trials for the trained stimulus set were typically conducted within a 15–20 min instructional period (depending on the participant). However, if a single training trial exceeded 15 min before the participant emitted a correct response, or if the participant emitted high rates of problem behavior, adjustments to this schedule were made. Only one stimulus set was taught per session. Training of stimulus sets and subsequent probes of all remaining untrained stimulus sets was repeated until the participant met the acquisition criteria for all 50 stimulus sets.

5.6. Procedures

5.6.1. Probes

An outline of the probe procedures is provided in Fig. 2. There were two types of probes: untrained and trained probes. Untrained probes were conducted as a measure of initial levels of responding in the baseline condition and generalization in the joint control training condition. Untrained probes consisted of one probe trial for each of the untrained stimulus sets. These probes were conducted during every session in baseline and sessions immediately following the acquisition of a trained stimulus set. Untrained stimulus sets were considered acquired when the participant emitted a correct selection response during a probe based on the response definition described above in one untrained probe session. Trained probes consisted of one probe trial of the current trained stimulus set and were conducted at the beginning of every training session. If a stimulus set was selected for training and a correct response occurred on the first probe prior to any training being conducted, the set was consider acquired as an untrained stimulus set and a new stimulus set was selected for training. Trained stimulus sets were considered acquired when the participant emitted a correct selection response during a probe based on the response definition described above over two consecutive sessions. In the joint control training condition, acquisition of a trained stimulus set occasioned probes of remaining untrained stimulus sets.

Probes were conducted by the experimenter presenting a messy field of the nonverbal stimuli (i.e., the 12 picture cards) on the table directly in front of the participant. The instructor then presented the vocal stimulus (e.g., "Give me the A, B, and C") and held out one flat hand, with palm facing up, directly in front of his or her body. The instructor's hand remained in this position for 20 s. If the participant did not complete the response within this interval, the instructor withdrew his or her hand

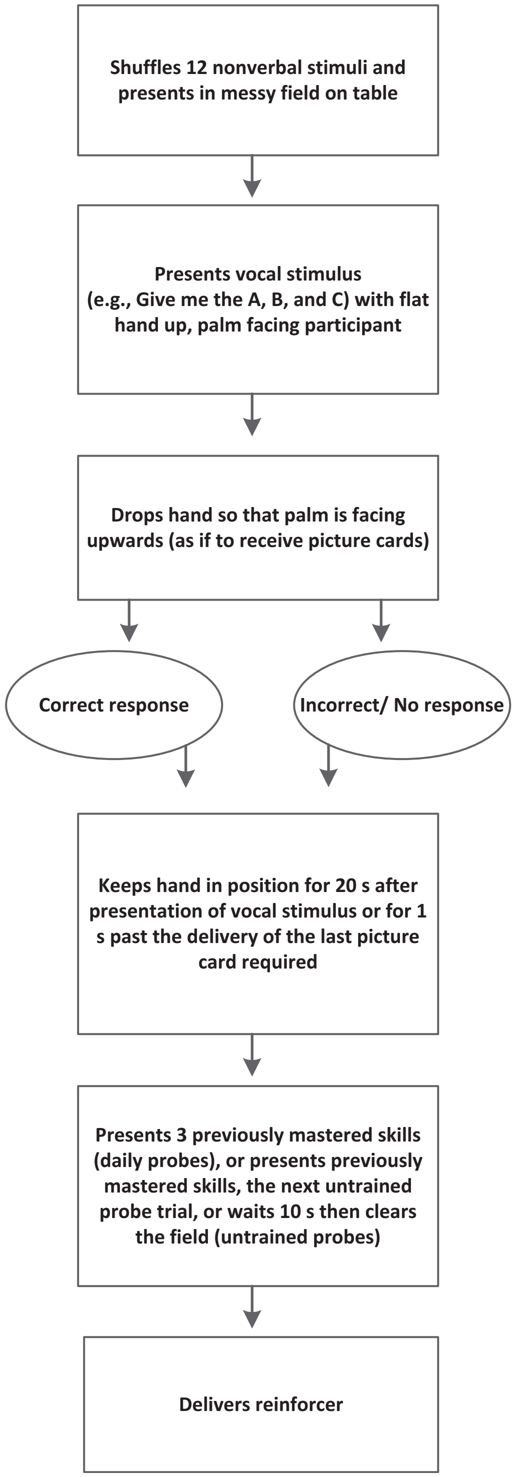

Fig. 2. Overview of probe procedures used across baseline and joint control training for both trained and untrained stimulus sets. Squares indicate a description of teacher behavior, ovals indicate a description of participant behavior. Arrows indicate the sequence of instruction.

at the end of the 20 s interval. If the participant initiated a selection response, the instructor kept his or her hand out until either the 20 s interval was complete or for 1 s beyond the selection of the last required number of picture cards. The withdrawal of the instructor's hand signaled the completion of a trial. For trained probes, following the withdrawal of the instructor's hand, regardless of the accuracy of participant response, three previously mastered responses were required before a reinforcer was delivered. For untrained probes, following the withdrawal of the instructors hand, the instructor moved on to either another trial, a previously mastered skill, or waited for 10 s then cleared the field and delivered non-contingent reinforcement depending upon the schedule of reinforcement previously determined for that participant.

5.6.2. Baseline

The baseline condition consisted of only untrained probes. Fifty untrained stimulus sets were probed in a random order each session. If a participant emitted a correct response to an untrained stimulus set during baseline, that stimulus set was considered to be acquired and replaced with a new, randomly selected stimulus set. For example, in baseline Cole responded correctly across five untrained stimulus sets. Following correct responses, these sets were removed from the pool of potential training sets and replaced with new, randomly selected sets. Consequently, for Cole, a total of 55 stimulus sets were used and for Billy a total of 51 sets were used. Any untrained stimulus set that the participant did not acquire during baseline was eligible to be entered into the training condition. All participants began training with 50 untrained stimulus sets.

5.6.3. Joint control training

In the joint control training condition, one stimulus set was randomly selected to be entered into training. To assess the effect of joint control training on the dependent variables, a daily probe for this trained stimulus set was conducted, prior to any training trials, according to the procedures described above.

A summary of the joint control training procedures is provided in Fig. 3. Following the probe, five trials using joint control training procedures were conducted to teach the participants to emit rehearsal and selection responses. A messy field of nonverbal stimuli was presented as in baseline. The instructor then presented a vocal stimulus (e.g., ''Give me the A, B, and C'') with a flat hand up, palm facing the participant while physically blocking the participant from responding until the entire instruction was completed as needed. Gesture (point), echoic, and/or mimetic prompts were provided in order to evoke echoic and self-echoic or intraverbal and self-mimetic rehearsal responses (i.e., repeating ''A,B,C. . .A,B,C. . .A,B,C. . .etc.'' vocally or with sign hand movements). The nonvocal participant in this study, Billy, emitted manual sign responses to vocal stimuli presented by the instructor. These responses were essentially translation responses, lacking point-to-point correspondence with the vocal stimuli presented and, therefore, technically meeting the definition of intraverbal. Early in training participants were prompted using gestures or other prompts to emit three consecutive self-echoic responses. As training progressed, the intrusiveness of prompts was gradually faded until self-echoic responses were emitted independently. After the participant completed the rehearsal requirement, the instructor paused for 1–2 s before repeating the vocal stimulus (i.e., ''Give me the A, B, and C''), with a flat hand up, palm facing the participant. This re-presentation of the vocal stimulus was faded as quickly as possible across stimulus sets. The instructor kept his or her hand for 20 s after the presentation of the vocal stimulus or for 1 s past the selection of the last required number of picture cards for that participant. If the participant emitted a correct response, a reinforcer was immediately delivered (e.g., preferred edibles/tangibles).

If at any point the participant emitted an incorrect response (echoic, self-echoic/mimetic, or selection), an error correction procedure was implemented. The participant's hands were reset to a neutral position (e.g., folded on the table), the instructor turned his or her head away for a 5 s time out, the field of nonverbal stimuli was shuffled and reset, and the joint control training procedure was restarted. During the joint control training procedure, prompts were adjusted depending upon the nature of the error. If the participant failed to emit self-echoic/mimetic responses, a partial phonemic prompt of the vocal stimulus was provided and the gestural prompts used initially during training to evoke self-echoic responses may have been implemented. If an error occurred while the participant was emitting a self-echoic/mimetic response, the number of rehearsals required prior to a selection response may have been increased. If the participant emitted an incorrect selection response, the hands were reset to a neutral position, the field was shuffled and reset, and the trial was restarted. When necessary, adjustments to the training procedures such as those described above may have been implemented. However, participants were never directly prompted to select the correct items (i.e., gestural or physical prompts of the response scheduled for reinforcement were not provided).

5.6.4. Procedural modifications

Beginning in Session 48 for Billy, 37 for Cole, and 34 for Abe, sessions that consisted solely of echoic and self-echoic or intraverbal and self-mimetic trials were provided prior to requiring selection responses. These sessions were conducted in the absence of the trained stimuli (i.e., without the 12 picture cards on the table) and without requiring the participant to emit a selection response. Reinforcement was contingent only upon the emission of echoic and self-echoic or intraverbal and self-mimetic responses. The purpose of these trials was to further strengthen the echoic and self-echoic or intraverbal and self-mimetic responses emitted by participants. In these sessions, the experimenter first conducted a trained probe of the current training stimuli. Following the probe, the stimuli were removed from the table and five echoic/mimetic training trials, consisting of the echoic and self-echoic or intraverbal and self-mimetic responses relevant to the current training stimuli, were conducted. In these trials gesture (point), echoic, and/or mimetic prompts were provided in order to evoke echoic/self-echoic or intraverbal/self-mimetic responses (i.e., repeating ''A,B,C. . .A,B,C. . .A,B,C. . .etc.'' vocally or with manual sign language). Participants were required to emit three consecutive independent self-echoic or self-mimetic responses and once this response requirement was met, a reinforcer was delivered. The intrusiveness of prompts was gradually faded

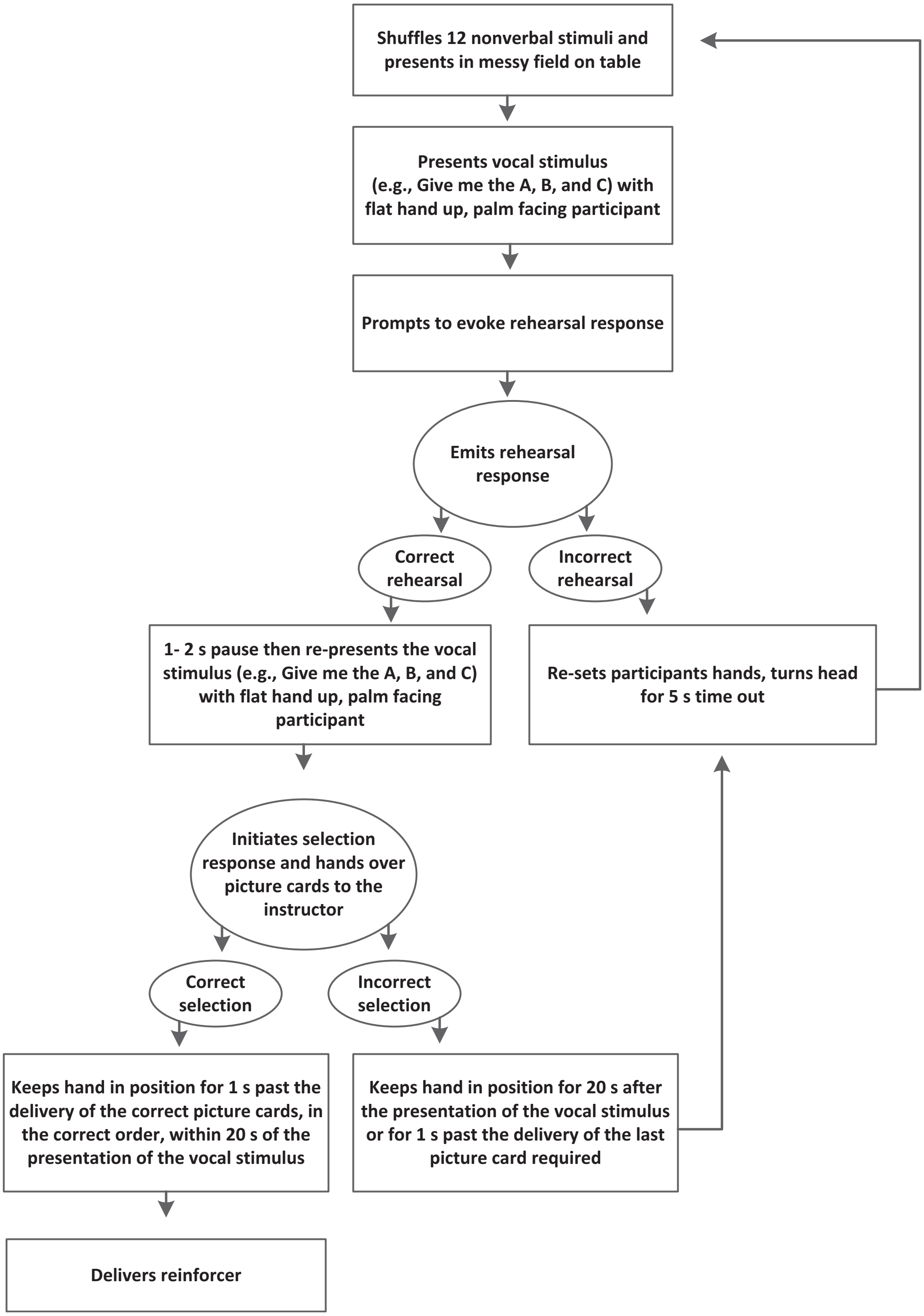

Fig. 3. Overview of joint control training and error correction procedures. Squares indicate a description of teacher behavior, ovals indicate a description of participant behavior. Arrows indicate the sequence of instruction.

within and across sessions until self-echoic or self-mimetic responses were emitted independently. These practice trials were conducted until participant responding achieved the mastery criterion of four out of five trials with accurate and independent echoic and self-echoic or intraverbal and self-mimetic responses within one session. Once this mastery criterion was achieved, the joint control training procedures described above resumed.

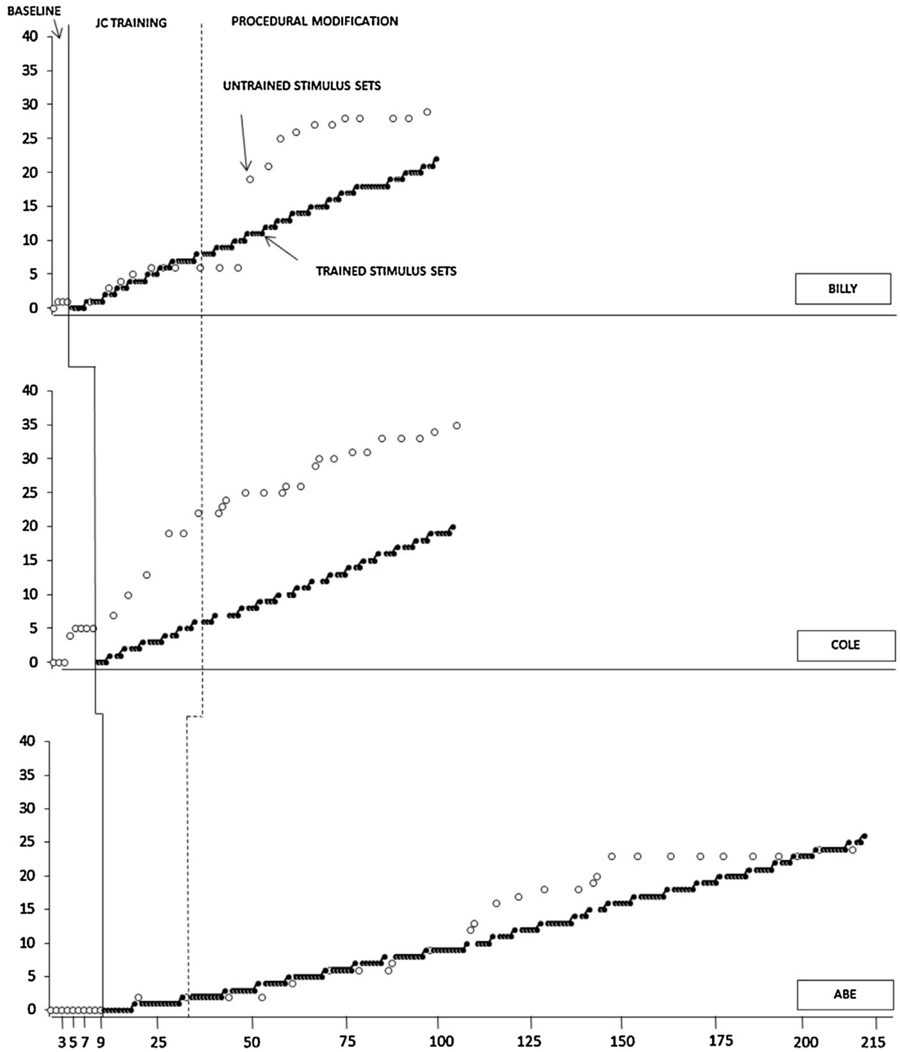

Fig. 4. The cumulative number of trained and untrained stimulus sets acquired during baseline and joint control training for Billy, Cole, and Abe.

6. Results

The cumulative number of trained and untrained stimulus sets acquired per session across baseline, joint control training, and procedural modification conditions are presented in Fig. 4. On the figure, the *y*-axis depicts the cumulative number of trained and untrained stimulus sets acquired and the *x*-axis depicts the number of instructional sessions. The closed circles indicate trained stimulus sets and the open circles indicate untrained stimulus sets. The results for Billy appear in the top tier of the figure, the results for Cole appear in the middle tier, and the results for Abe appear in the bottom tier.

6.1. Billy

Billy's baseline responding stabilized at one untrained stimulus set out of 51 after four baseline sessions. Joint control training began in Session 5 and continued through Session 124. The first trained stimulus set met the acquisition criteria after six teaching sessions. No untrained stimulus sets met the acquisition criterion during the first untrained probe. However, two untrained stimulus sets were acquired after the second trained stimulus set was mastered. Acquisition of untrained stimulus sets continued to increase gradually over the next three untrained probe sessions with a total of five untrained stimulus sets meeting the acquisition criterion during the joint control training condition by Session 46. From Session 5 to 46, an average of four teaching sessions were conducted before trained stimulus sets met the acquisition criteria. Untrained stimulus sets were acquired at a rate of about one set every 11 sessions. From Session 32 to 47 three additional trained stimulus sets met the acquisition criterion but Billy failed to emit any correct responses across three untrained probes. Consequently, the procedural modification consisting of additional echoic and self-echoic or intraverbal and self-mimetic training was introduced in Session 48. An immediate effect upon the emission of correct responses during untrained probes was not observed following this change. However, after training three additional sets to the mastery criterion, there was a substantial increase in the number of untrained stimulus sets acquired with 13 sets meeting the mastery criterion during a single untrained probe sessions for a total of 19 untrained stimulus sets. Acquisition of untrained stimulus sets continued to increase throughout the remainder of the experiment with only two brief plateaus in responding. After the introduction of the procedural modification, an average of about four teaching sessions were required before a trained stimulus set met the acquisition criterion and 23 untrained stimulus sets met the mastery criterion at a rate of about one set acquired every four sessions. For target 17, Billy's responding during trained probes met the acquisition criterion following only trials consisting only of intraverbal and self-mimetic training. In total, Billy acquired 22 trained stimulus sets and 28 untrained stimulus sets across 120 joint control training condition sessions.

6.2. Cole

Cole's baseline responding stabilized at five untrained stimulus sets out of 55 total sets across eight baseline sessions. Joint control training began in Session 9 and continued through Session 105. The first trained stimulus set met the acquisition criteria after four teaching sessions and two untrained stimulus sets were acquired in the first untrained probe sessions. Acquisition of untrained stimulus sets continued to occur across the next five untrained probe sessions, except one brief plateau at Session 32. By Session 36, a total of 17 untrained stimulus sets had met the mastery criterion in the joint control training condition. From Session 9 to 36, an average of four teaching sessions were required before a trained stimulus set met the acquisition criteria and untrained stimulus sets were acquired at a rate of about one set every two sessions. The procedural modification was introduced in Session 37 for Cole. Following this change, Cole's responding met the acquisition for the seventh trained stimulus set within four teaching sessions but correct responding across untrained probes did not occur. However, correct responding did occur across two untrained stimulus sets randomly selected as the next sets to undergo joint control training. No additional untrained stimulus sets were acquired across two untrained probes but acquisition increased again at Session 59 when Cole emitted correct responses for a set randomly selected to undergo joint control training. Acquisition of untrained stimulus sets continued to gradually increase over the remainder of experimental sessions with a few brief plateaus in acquisition. After the introduction of the procedural modification, an average of about four teaching sessions were required before a trained stimulus set met the acquisition criterion and 13 untrained stimulus sets met the mastery criterion at a rate of about one set acquired every five sessions. For stimulus sets 12 and 13, Cole's responding during trained probes met the acquisition criterion following trials consisting only of echoic and self-echoic training. In total, Cole acquired 20 trained stimulus sets and 30 untrained stimulus sets, across 96 joint control training sessions.

6.3. Abe

Baseline responding for Abe remained stable at zero untrained stimulus sets out of 50 total sets for nine sessions. Joint control training began in Session 10 and continued through Session 215. The first trained stimulus set met the acquisition criterion after 10 teaching sessions and, two untrained stimulus sets were acquired immediately following the acquisition of the first trained stimulus set. A second trained stimulus set was acquired after an additional 12 joint control teaching sessions but no additional untrained stimulus sets met the acquisition criterion. From Session 10 to 33 an average of 12 teaching sessions were required before a trained stimulus set met the acquisition criterion and untrained stimulus sets were acquired at a rate of about one set every 12 sessions. The procedural modification was introduced in Session 34. An immediate effect upon the emission of correct responses during untrained probes was not observed following this change. However, after training two additional sets to the mastery criterion, two additional untrained stimulus sets were acquired. Acquisition of untrained stimulus sets continued to increase gradually until 23 sets were acquired. At this point acquisition plateaued across seven untrained probe sessions. One additional untrained stimulus set was acquired in Session 212. In Sessions 88, 110, and 143, Abe's responding achieved the mastery criterion during probes for targets randomly selected to undergo joint control training. After the introduction of the procedural modification, an average of about seven teaching sessions were required before a trained stimulus set met the acquisition criterion and 22 untrained stimulus sets met the mastery criterion at a rate of about one set acquired every nine sessions. Across 206 joint control training sessions, Abe acquired 26 trained stimulus sets and 24 untrained stimulus sets.

7. Discussion

The current study examined the effects of teaching procedures derived from Lowenkron's (1991, 1998, 2004, 2006a) analysis of joint control on the listener behavior of children with autism and other developmental disabilities. In this experiment the emission of a single response topography occurring under two different sources of control (i.e., self-echoic and tact or self-mimetic and tact) occasioned selection responses. Given this analysis, the onset of joint control was a generic event consisting of the simultaneous control of two discriminative stimuli over a single response. The generic nature of joint control events allowed for novel responses or generalized responding to occur. The results of this study indicated that joint control training was effective in increasing trained and untrained listener responses for the three participants involved. The most notable finding of this study was that all participants emitted correct responses to untrained stimulus sets. For Billy and Cole, acquisition of untrained stimulus sets exceeded trained stimulus sets and for Abe, an almost equal number of sets were acquired. These results suggest that teaching mediating responses, consistent with an analysis of joint control, may produce important language benefits for children with language deficits. The procedures described within this study offer one possible method by which to establish responding under joint stimulus control.

The rate of acquisition for trained and untrained sets varied across participants. A review of each participant's rate of acquisition for listener skills prior to the onset of this study revealed acquisition rates similar to those obtained in this experiment. The suggestion that maturation could account for the observed changes in this lengthy experiment is mitigated by two factors. First children with autism typically fail to acquire skills without explicit teaching and do not learn merely from their interactions with the physical and social environments (Smith, 2001). Second, the rate of acquisition of listener responses during this experiment is consistent with the participant's previous instructional history over a similar time period suggesting the minimal role of the passage of time as an independent variable.

For children with autism responding under joint control may be prevented by: (a) insufficient echoic and self-echoic and tact repertoires (Michael, Palmer, & Sundberg, 2011) or (b) a failure of natural contingencies of reinforcement supplied by the environment to select such responses. Consequently, a repertoire of responding to the onset of a joint control event may require systematic and precise programmed instruction for children with autism of the type described in this study. In the current study, participants had previously acquired tacts of the experimental stimuli. However, repertoires related to the emission of echoic and self-echoic or intraverbal and self-mimetic responses following instructional commands had not been previously established. The joint control procedures described within this study were designed to explicitly target these repertoires by teaching participants to emit echoic or intraverbal responses following the presentation of vocal verbal stimuli and to preserve the stimuli through self-echoic or self-mimetic responses. Once these repertoires had been established, the occurrence of selection responses to the onset of a joint control event was possible. Consequently, according to an analysis of joint control, during the selection task when the participant scanned the array of comparison stimuli, tested the properties of the stimuli encountered, and emitted the same response topography as a tact currently being emitted as a self-echoic, this event became discriminative for a selection response. Once participants' responding came under the control of joint control events, regardless of the specific stimuli presented, accurate selection of untrained stimuli emerged. Based on these results it appears that the joint control training facilitated generalized responding. Lowenkron (2006b) recognized that the analysis of joint control is partly interpretative. The mediating responses that are simultaneously emitted under one source of control are often covert and therefore unobservable by individuals other than the person behaving. The procedures of the current study sought to provide some empirical evidence in support of this interpretative analysis by requiring participants to emit overt responses at times when it is suspected that covert mediating responses are occurring for individuals with more sophisticated verbal repertoires.

A number of authors have expounded upon the advantages of a joint control analysis, not only as an explanation for complex human behavior (Lowenkron, 1998), but as a means by which to design language training programs for individuals with language deficits and delays (degli Espinosa, 2011; Michael et al., 2011; Sidener, 2006; Tu, 2006). A major advantage of a mediated stimulus selection account is related to the issue of efficiency (Lowenkron, 1998; Sidener, 2006). Unmediated stimulus selection requires repeated exposure to a stimulus set in order to establish an appropriate history of reinforcement for selection responses (Michael et al., 2011). Repeated instructional opportunities would also need to be arranged for all relevant stimulus pairs across hundreds of stimuli and response classes. The individual instruction of each relevant selection or listener response is not only inefficient, but likely impossible. Conversely, the generic nature of mediated stimulus selection proposed by a joint control analysis facilitates responding to novel stimuli and addition of this account into language training programs presents a highly efficient instructional practice. The findings of the current study support the contention that joint control training offers an efficient means by which to teach language. The results indicated that for all three participants untrained stimulus sets were acquired at a rate of about one set for every trained stimulus set acquired.

Michael et al. (2011) identified joint control as a special case of convergent multiple control, defined as "the convergent control of a response of a particular topography by two concurrent variables" (p. 21). Michael et al. (2011) and Palmer (2006), Palmer (2010) asserted that the convergent control of two or more stimuli extends beyond joint control events and suggested that changes in saltation of response strength may provide a ubiquitous and plausible explanation for the more general phenomena of multiply controlled responses. Palmer (2006) suggested that at any given moment interaction with environmental stimuli strengthens a host of possible responses but weak stimulus control or strong competing responses may prevent emission. The onset of some additional stimulus, however, may strengthen previously potentiated response forms and cause a discriminable "jump" in response strength, leading to response emission. Consistent with Lowenkron's (1998) analysis, Michael et al. (2011) suggested that the occurrence of joint control is a "discriminable event that would control a selection response" (p. 21). Given this description, in the current study the vocal stimulus presented by the instructor strengthened a selection response but the control by the vocal stimulus alone was not sufficient to exert control over the selection response. However, as the participant scanned the array of stimuli, a "jump" in response strength would have occurred when same response topography was simultaneously emitted as both a self-echoic and tact or self-mimetic and tact, occasioning the selection response.

There are three limitations of this study. First, the same 12 stimuli were used in the 50 or more stimulus sets for each participant, although the arrangement of stimuli within each set always varied. For example, for Billy, a trained stimulus set was ball, balloons, and chicken, and an untrained stimulus set consisted of the same stimuli but arranged as ball, chicken, and balloons. Consequently, it is possible that inclusion of the same stimuli across sets may have influenced the outcome as a result of exposure or practice effects and mitigated the results of the independent variable. However, the number of sets containing identical stimuli was relatively low, thereby limiting the influence this variable may have had on the results obtained. In addition, given that the ordering of stimuli varied for each of these sets, the vocal stimuli presented by the instructor and the target responses relevant to each set were not in fact identical. Reinforcement was contingent upon participants emitting selection responses in the order in which the instructor presented the vocal stimuli and any responses that deviated from that order were considered errors. Nevertheless, future research should examine the issue of overlapping stimuli and include conditions which test completely novel stimulus sets as a further indicator of the establishment of generalized responding.

A second limitation was that participants were not required to emit tact responses of the experimental stimuli during the selection tasks. Anecdotal reports of participant performance provide tentative evidence that participants were emitting tact responses during the selection task and therefore implicate a joint control analysis of the results obtained. Billy, a nonvocal participant, was observed to respond intraverbally to the instructor's vocal stimulus and preserve the stimulus through an overt self-mimetic response (i.e., signing the names of the pictured items named by the instructor) during probes even when he was not required to by the experimental contingencies. In addition, he emitted an overt tact of each item before selecting the stimulus. Billy was also observed to continue to rehearse the manual sign movements while scanning the field and move past any stimuli which did not match his sign hand movement. It was only when a stimulus evoked the same manual sign response as a tact that was simultaneously being emitted as a self-mimetic that Billy emitted a selection response.

Cole also consistently emitted self-echoic responses, even when he was not required to do so by the experimental contingencies. Additionally, Cole was observed to consistently overtly tact the item as he selected it. On occasion, he was observed pointing to a target nonverbal stimulus, emitting a vocal tact, and then emitting a selection response. Cole's rehearsal was further strengthened and improved following the implementation of additional echoic training trials. On one occasion Cole was observed to emit echoic and self-echoic responses, scan the field, tact all 12 stimuli encountered, and when he emitted a tact response of the same topography as the self-echoic response being emitted, select the nonverbal stimulus. Nonverbal stimuli that did not evoke the same response topography as the self-echoic were discarded. Furthermore, during probes, if Cole did not emit an echoic response consistent with the teacher's instructional command but instead emitted some other response not scheduled for reinforcement, he consistently emitted an incorrect selection response consistent with whatever stimulus was preserved through the self-echoic. For example, if the teacher's instruction was, "Give me the milkshake, motorcycle, pumpkin, and trampoline," but Cole emitted the response, "Milkshake, macaroni, pumpkin, trampoline" as an echoic and self-echoic, the subsequent selection response emitted would be consistent with the stimuli preserved through a self-echoic.

Abe also emitted self-echoic responses during both trained and untrained probes. When Abe emitted correct self-echoic rehearsal responses, accurate stimulus selection responses consistently occurred. Abe was observed to emit self-echoic rehearsal responses while scanning the stimulus array and occasionally he reached toward a stimulus not included in the target set. However, he would then point to the stimulus, emit a tact response, withdraw his hand, and move on to another stimulus, resuming the emission of self-echoic responses. Only when he encountered a stimulus which evoked a tact response with the same topography as the self-echoic, would he emit a selection response.

In general, when mediating responses (self-echoic, self-mimetic, and tact responses) were overt, correct selection of the stimulus set was more likely. Similar to the studies performed by Gutierrez (2006), Lowenkron (2006b), and DeGraaf and Schlinger (2012), future research should experimentally investigate the individual roles of echoic, self-echoic, and tact responses to provide additional empirical evidence in favor of a joint control analysis.

Finally, only results from three participants were obtained and there was considerable variability in the response repertoires of the participants involved. Therefore replications of these findings with additional learners with autism who have varying characteristics are needed to determine the benefits of these procedures for a broad participant group. For one participant the response form was manual sign language, suggesting the applicability of these procedures to nonvocal children for whom alternative communication systems are necessary. An important point, however, is that responding to joint control events depends upon topography-based verbal behavior (e.g., manual sign language, writing) (Lowenkron, 1991). Consequently, selection-based methods of communication such as the Picture Exchange Communication System (Bondy & Frost, 2002) or icon selection using a touch screen device preclude the occurrence of responding under the control of joint control events. As Lowenkron (1991) stated, "Generalized selection-based verbal behavior is thus dependent on, rather than alternative to, topography based verbal behavior" (p. 125).

Within autism treatment programs skills that are often acquired under the title of "auditory and visual memory" or cognitive skills may actually be acquired through the unwitting effects of the type of verbal mediation that was explicitly taught in this study. For example, tasks that include delayed match to sample of an array of items, completing a complex pattern of items, finding a previously displayed item within a large array, identifying what is missing from a previously displayed array of items, following multiple step instructions, answering yes or no, and counting out a specific number of items from a larger set when requested may best be taught by teaching the verbal mediation over listener behavior described in this study. Introduction of joint control procedures to teach important listener skills within treatment programs for children with autism could add substantially to the outcomes for many individuals. These procedures provide a behavior analytic and conceptually systematic approach to teaching important skills that previously were assumed to rely on the development of cognitive abilities. Consideration of overt and covert verbal mediation as the underlying mechanism by which complex listener responding is established may lead to more effective action by researchers and practitioners charged with developing teaching procedures related to complex "cognitive skills." Palmer (2006) discussed the important role joint control plays in the control of human behavior and notes the general lack of recognition it has previously been given by behavior analysts. He stated, "Joint control is a tool in the workshop of the behavior analyst who would understand complex behavior. It is not a new phenomenon, nor does an analysis of joint control invoke new principles. It has been lying in the toolbox all along, but we are only beginning to appreciate its role in the control of human behavior. I believe that it will have a distinguished future" (p. 214).

References

Bondy, A., & Frost, L. (2002). *The picture exchange communication system training manual* (2nd ed.). Newark, DE: Pyramid Educational Products Inc.

degli Espinosa, F. (2011). *Verbal behaviour development for children with autism*. Unpublished doctoral dissertation or master's thesis. University of Southampton, Southampton, United Kingdom.

DeGraaf, A., & Schlinger, H. D. (2012). The effects of joint control on the acquisition and durability of a sequencing task. *The Analysis of Verbal Behavior, 28*, 59–71.

Gutierrez, R. D. (2006). The role of rehearsal in joint control. *The Analysis of Verbal Behavior, 22*, 183–190.

Horner, R. H., & Baer, D. M. (1978). Multiple-probe technique: A variation on the multiple baseline design. *Journal of Applied Behavior Analysis, 11*, 189–196.

Lowenkron, B. (1984). Coding responses and the generalization of matching to sample in children. *Journal of the Experimental Analysis of Behavior, 42*, 1–18.

Lowenkron, B. (1988). Generalization of delayed identity matching in retarded children. *Journal of the Experimental Analysis of Behavior, 50*, 163–172.

Lowenkron, B. (1989). Instructional control of generalized relational matching to sample in children. *Journal of the Experimental Analysis of Behavior, 52*, 293–309.

Lowenkron, B. (1991). Joint control and the generalization of selection-based verbal behavior. *The Analysis of Verbal Behavior, 9*, 121–126.

Lowenkron, B. (1996). Joint control and word-object bidirectionality. *Journal of the Experimental Analysis of Behavior, 65*, 252–255.

Lowenkron, B. (1997). The role of naming in the development of joint control. *Journal of the Experimental Analysis of Behavior, 68*, 244–247.

Lowenkron, B. (1998). Some logical functions of joint control. *Journal of the Experimental Analysis of Behavior, 69*(3), 327–354.

Lowenkron, B. (2004). Meaning: A verbal behavior account. *The Analysis of Verbal Behavior, 20*, 77–97.

Lowenkron, B. (2006a). An introduction to joint control. *The Analysis of Verbal Behavior, 22*, 123–127.

Lowenkron, B. (2006b). Joint control and the selection of stimuli from their description. *The Analysis of Verbal Behavior, 22*, 129–151.

Lowenkron, B., & Colvin, V. (1992). Joint control and generalized nonidentity matching: Saying when something is not. *The Analysis of Verbal Behavior, 10*, 1–10.

Lowenkron, B., & Colvin, V. (1995). Generalized instructional control and the production of broadly applicable relational responding. *The Analysis of Verbal Behavior, 12*, 13–29.

Michael, J., Palmer, D. C., & Sundberg, M. L. (2011). The multiple control of verbal behavior. *The Analysis of Verbal Behavior, 27*, 3–22.

Palmer, D. C. (2006). Joint control: A discussion of recent research. *The Analysis of Verbal Behavior, 22*, 209–215.

Palmer, D. C. (2010). Behavior under a microscope. *The Behavior Analyst, 33*, 37–45.

Sautter, R. A., & LeBlanc, L. A. (2006). Empirical applications of Skinner's analysis of verbal behavior with humans. *The Analysis of Verbal Behavior, 22*, 35–48.

Schlinger, H. D. (2008). Listening is behaving verbally. *The Behavior Analyst, 31*(2), 145–161.

Sidener, D. (2006). Joint control for dummies: An elaboration of Lowenkron's model of joint (stimulus) contorl. *The Analysis of Verbal Behavior, 22*, 119–123.

Sidener, D. W., & Michael, J. (2006). Generalization of relational matching to sample in children: A direct replication. *The Analysis of Verbal Behavior, 22*, 171–182.

Skinner, B. F. (1957). *Verbal behavior*. Englewood Cliffs, NJ: Prentice-Hall.

Smith, T. (2001). Discrete trial training in the treatment of autism. *Focus on Autism and Other Developmental Disabilities, 16*, 86–92.

Stages Learning Materials. (1997). Language builder: Picture noun card, Chicago, CA.

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

Sundberg, M. L., & Michael, J. (2001). The value of Skinner's analysis of verbal behavior for teaching children with autism. *Behavior Modification, 25*, 698–724.

Tu, J. C. (2006). The role of joint control in the manded selection responses of both vocal and non-vocal children with autism. *The Analysis of Verbal Behavior, 22*, 191–207.