JOURNAL OF APPLIED BEHAVIOR ANALYSIS 2008, **41,** 53–68 NUMBER 1 (SPRING 2008)

AN EVALUATION OF INTRAVERBAL TRAINING AND LISTENER TRAINING FOR TEACHING CATEGORIZATION SKILLS

ANNA INGEBORG PETURSDOTTIR, JAMES E. CARR, SARAH A. LECHAGO, AND SEASON M. ALMASON

WESTERN MICHIGAN UNIVERSITY

The effects of vocal intraverbal training and listener training on the emergence of untrained categorization skills were evaluated. Five typically developing preschool children initially learned to name a number of previously unfamiliar visual stimuli. Each child then received one of two types of training. Intraverbal training involved reinforcing vocally emitted category names by the child in the presence of a spoken exemplar name. Listener training involved reinforcing the selection of visual stimuli by the child in the presence of a spoken category name. A multiple baseline design was used to evaluate the effects of training on each child's intraverbal category naming and stimulus selection. Pre- and posttests were used to assess additional emergent relations. Four of 5 participants did not demonstrate the emergence of any untrained relations. The current findings suggest that additional research is needed to determine the optimal sequencing of conditions in programs designed to teach categorization skills.

DESCRIPTORS: categorization, functional independence, intraverbals, listener behavior, language training

It has been observed that comprehension often appears to precede production in typical language development (e.g., Fraser, Bellugi, & Brown, 1963). This observation was among factors that led Lovaas (1977, 1981) to recommend teaching receptive language skills before expressive language skills in early and intensive behavioral intervention (EIBI) programs for children with autism. Mastery of receptive programs is still listed as a prerequisite for the introduction of corresponding expressive programs in some contemporary EIBI manuals (Leaf & McEachin, 1999; Maurice, Green, & Luce, 1996). However, as Lovaas (1977, 2003) has pointed out, the sequence in which typically developing children acquire language skills may not necessarily be the ideal sequence in which to teach those same skills to children who do not acquire them in their natural environment. Rather, the sequencing of programs should be based on empirical data.

From a behavioral perspective, receptive and expressive language skills can be understood as instances of different types of stimulus control over topographically distinct responses. Skinner's (1957) analysis of verbal behavior provides one behavior-analytic framework for describing receptive and expressive skills in terms of stimulus control and reinforcement history. In terms of Skinner's analysis, expressive language may take the form of a number of different verbal operants emitted by a speaker. One such operant, the mand, consists of a verbal response evoked by a motivating operation (MO; Laraway, Snycerski, Michael, & Poling, 2003) due to a history of response-specific reinforcement. Other verbal operants, such as the tact and the intraverbal, consist of verbal responses under

This article is based on a dissertation submitted by the first author under the supervision of the second author in partial fulfillment of the requirements for the PhD degree in psychology. We thank committee members Jack Michael, Scott Gaynor, Cynthia Pietras, and Norm Peterson for helpful comments on the dissertation. We also thank Kelsey Behnke, Heather DeForest, Allison DeGraaf, Juliet Flynn, Adam Freeman, Jeana Koerber, Tracy Lepper, Catrina Litzenburg, Sadie Lovett, Jada Rae-Traub, Stacy Ruse, and Kelly Stone for their assistance with data collection. Anna Ingeborg Petursdottir is now at Texas Christian University.

Address correspondence to James E. Carr, Department of Psychology, Western Michigan University, 1903 W. Michigan Ave., Kalamazoo, Michigan 49008 (e-mail: jim.carr@wmich.edu).

doi: 10.1901/jaba.2008.41-53

discriminative control. The tact is evoked by a nonverbal discriminative stimulus (S^D); for example, the sight of a cat may evoke the vocal response ''cat.'' The intraverbal, on the other hand, is evoked by a verbal S^D that does not have point-by-point correspondence with the response product. For example, the vocal response ''cat'' may be evoked by a vocally presented stimulus such as ''animal'' or ''whiskers.'' These basic verbal operants comprise only portions of the complex verbal interchanges that may occur between speaker and listener (Skinner). However, the early stages of EIBI programs target them in relatively pure form, in that brief vocal utterances are reinforced in the presence of specific MOs or S^Ds while other components of verbal interaction are held constant. EIBI curricula (e.g., Leaf & McEachin, 1999; Maurice et al., 1996) contain a large number of expressive programs that involve the reinforcement of either tacts or intraverbals. Tacts are established in programs that teach children to name, for example, objects and object attributes, whereas intraverbals are established in programs designed to teach such skills as responding to social questions and comments, answering general knowledge questions, and describing objects in terms of their functions, features, or category membership.

In contrast to training protocols designed to produce expressive language, receptive protocols are designed to teach a child to respond to verbal stimuli by emitting motor responses or stimulus-selection responses (e.g., pointing to stimuli). In EIBI curricula, the implementation of programs that target tacts and intraverbals is frequently preceded by receptive programs in which the child's selection of visual stimuli is reinforced following a trainer's vocal presentation of, for example, the names of objects, categories, functions, or features. For example, in the presence of the vocally presented stimulus ''cat,'' the child's selection of a picture of a cat from an array of pictures of various animals may be reinforced. In line with Skinner's (1957) analysis, this type of training has sometimes been said to establish listener behavior (Horne & Lowe, 1996; Michael, 1985).

Evidence exists that among individuals with language deficits, speaker and listener repertoires can be independent of one another, such that the establishment of each may require a separate reinforcement history (Guess & Baer, 1973; Lee, 1981). Nevertheless, most studies have found some degree of interdependence between the speaker and listener repertoires of children with autism and other developmental disabilities as well as those of young, typically developing children (such interdependence has been addressed conceptually by recent behavior-analytic accounts of language, e.g., Barnes-Holmes, Barnes-Holmes, & Cullinan, 2000; Horne & Lowe, 1996). Of relevance to the sequencing of receptive and expressive programs, several studies have produced evidence that the training of vocal or signed tacts, in Skinner's (1957) taxonomy, generated untrained listener relations more readily than listener training generated tacts (Connell & McReynolds, 1981; Keller & Bucher, 1979; Lee; Miller, Cuvo, & Borakove, 1977; Smeets & Striefel, 1976; Williams & McReynolds, 1975; Wynn & Smith, 2003). Tact training also has been found to require fewer training trials than listener training (Connell & McReynolds; Cuvo & Riva, 1980; Smeets, 1978), and it has been demonstrated that prior listener training does not necessarily facilitate subsequent tact training (Miller et al.; Watters, Wheeler, & Watters, 1981). Further, the results of a series of recent experiments (Horne, Hughes, & Lowe, 2006; Horne, Lowe, & Randle, 2004; Lowe, Horne, Harris, & Randle, 2002; Lowe, Horne, & Hughes, 2005; Miguel, Petursdottir, Carr, & Michael, in press) indicate that when either tact or listener training successfully establishes both tact and listener repertoires, additional untrained relations may emerge.

The above findings appear to indicate a greater efficiency of tact than listener training in

terms of both training effort and collateral gains. Consequently, they do not appear to support the practice of training receptive before expressive skills, when the expressive skill in question is the tact. If these findings generalize to teaching intraverbal relations, then intraverbal training may be expected to generate listener relations more readily than listener training generates intraverbals as well as perhaps to produce additional emergent relations. However, such generalization cannot be taken for granted. A vocal intraverbal relation differs from the tact in that it does not contain a visual (or other nonverbal) stimulus that a listener can select in response to a verbal stimulus. Due to its absence, it intuitively appears unlikely that appropriate listener behavior should emerge as a result of intraverbal training alone, or vice versa. For example, a child who has learned to emit the vocal response ''cat'' in the presence of the vocal auditory stimulus ''animal'' may be unlikely to select a picture of a cat in the presence of the same stimulus, in the absence of any additional training involving cats as visual stimuli. However, by the time such programs are implemented, students typically have acquired the relevant tacts and listener relations (e.g., a child has typically learned to tact various individual animals before being introduced to a program that introduces ''animal'' as a verbal stimulus), and it is conceivable that such relations may serve to link intraverbals to additional listener relations (Horne & Lowe, 1996).

To date, very few studies have investigated the effects of intraverbal training on the emergence of listener relations, or vice versa. Data on a secondary dependent measure in a study by Luciano (1986) indicated that the establishment of intraverbal relations (vocally emitting the names of category exemplars when presented with category names) was accompanied by the emergence of appropriate listener relations (selecting visual stimuli when presented with category names) among children with developmental disabilities. Miguel, Petursdottir, and Carr (2005) evaluated the effects of listener training on the emergence of intraverbal repertoires and found them to be minimal among young, typically developing children. These preliminary findings provide tentative evidence that intraverbal training may generate both intraverbal and listener relations more readily than does listener training. However, no studies have employed similar procedures to investigate the effects of both types of training with similar participants.

The purpose of the present study was to provide an analogue evaluation of the effects of listener training and intraverbal training on the emergence of listener and intraverbal relations to assess the emergence of intraverbal and listener relations. We also assessed two additional vocal relations (termed category tacts and reverse intraverbals) as well as performance on a category match-to-sample task. In previous research on tacts and listener relations (Horne et al., 2004, 2006; Lowe et al., 2002, 2005; Miguel et al., in press),[1] the emergence of listener relations as a result of tact training, and vice versa, predicted the emergence of additional categorization skills. The present study addressed the extent to which a similar covariation might be observed in the context of intraverbal and listener training.

METHOD

Participants and Setting

Five children without any known developmental delays participated with their parents' permission. The children's ages ranged from 3 years 5 months to 3 years 8 months at the

[1]For readers familiar with the work of Horne, Lowe, and colleagues, it may be helpful to note that those researchers trained either common tact (Lowe et al., 2002, 2005) or listener (Horne et al., 2004, 2006) relations, in which one name corresponded to several category exemplars. Listener training in those studies parallels category-name listener training in the present study, whereas tact training in the present study was a preexperimental condition and did not involve training common names. Instead, common intraverbal responses were trained.

beginning of the study and from 3 years 5 months to 4 years 2 months at its completion. Informal observation indicated that the children spoke in sentences, readily followed simple instructions, and could follow instructions to imitate each of the vocal response topographies that we trained and tested during the study.

Sessions were conducted in a partitioned area (1.2 m by 2.1 m) in a room at the children's preschool. The participant and the experimenter were seated side by side at a child-sized table, facing a wall. A second observer, when present, was seated at one end of the table. Each child participated in sessions once or twice per day, 5 days a week, and each session lasted approximately 10 to 20 min. At the end of each session, the child selected a small toy or snack for completing the entire session.

Stimuli

The 30 visual stimuli used in the experiment consisted of framed line drawings of 12 outline maps of foreign countries, 12 characters from foreign writing systems (two of each from Greek, Cyrillic, Hiragana, and Katakana, and four from Hebrew), and six line drawings of familiar items (dog, cat, apple, banana, shoe, and flower). The stimuli measured 6.3 cm by 6.3 cm and were printed in black ink on white sheets of paper (21.6 cm by 27.9 cm) (trial sheets) as well as on white cards (6.3 cm by 6.3 cm) (stimulus cards). Trial sheets were inserted into plastic sheet protectors contained in a three-ring binder. Underneath each trial sheet was a blank sheet, followed by the next trial sheet. Depending on the specific training or testing condition, a trial sheet contained either one stimulus centered in the middle of the sheet, three horizontally aligned stimuli, or one stimulus centered in the upper half of the sheet and three horizontally aligned stimuli in the bottom half of the sheet, covered by a paper flap. Stimulus cards were inserted into transparent hard plastic covers (7.5 cm by 10 cm).

Sixteen of the stimuli (eight outline maps and eight foreign characters) were organized into four stimulus sets, in which each stimulus was assigned a category name and an exemplar name. As shown in Figure 1, each set contained two categories, and within each category two exemplars (e.g., in the north–south set, the north category contained Rocco and Mali and the south category contained Mozam and Bia). Eight additional stimuli (four outline maps and four Hebrew characters) were assigned exemplar names (Chile, Tina, Fiji, Nepal, Al, Mem, Kaf, and Hey) but were not assigned category membership. Each child received training on a maximum of two stimulus sets and a maximum of four additional maps and characters. The six familiar stimuli were used only in the pretraining phase of the experiment. They included two categories (fruits and animals) with two exemplars in each and two stimuli that were not assigned category membership.

Data Collection

Dependent variables. The primary dependent variables were (a) category-name listener responding, defined as selecting a visual category exemplar given its spoken category name (e.g., selecting Rocco given the auditory stimulus ''north''); and (b) category-name intraverbals, defined as vocalizing the name of a category given a spoken exemplar name (e.g., saying ''north'' given the auditory stimulus ''Rocco''). Secondary dependent measures were performance on pre- and posttests of (a) category tacts, saying a category name when presented with a visual category exemplar (e.g., saying ''north'' given the visual stimulus of Rocco); (b) reverse intraverbals, saying an exemplar name given a spoken category name (e.g., saying either ''Rocco'' or ''Mali'' given the auditory stimulus ''north''); and (c) category match to sample, defined as selecting a visual category exemplar when presented with another visual exemplar of the same category (e.g., selecting Mali given the visual stimulus of Rocco).

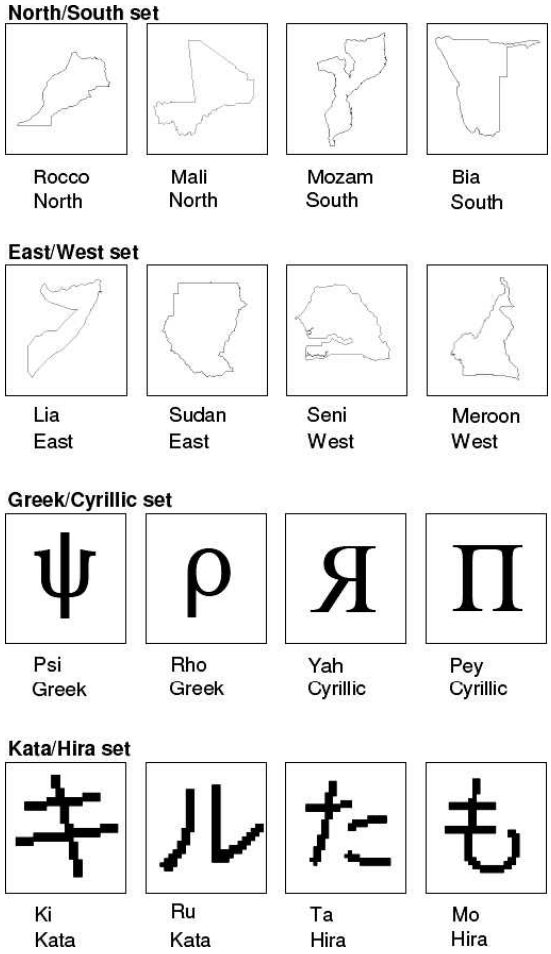

Figure 1. Stimulus sets used in the experiment. Exemplar names are printed below the visual stimuli to which they were assigned, and category names are printed below the exemplar names.

Scoring. The experimenter recorded the participant's responses on a data sheet. In trials conducted in a match-to-sample format (listener and category match-to-sample trials), the experimenter scored a response as correct if the child touched the positive comparison within 10 s of stimulus presentation. The experimenter scored an incorrect response if the child touched another comparison or did not touch a stimulus within 10 s. If the child touched more than one stimulus, the experimenter recorded only the first response. On trials that required a vocal response, the experimenter recorded a correct response if the child vocalized the target category or exemplar name. The experimenter scored an incorrect response if the child did not emit the target response within 10 s of the experimenter's instruction, or if the child vocalized a different category or exemplar name. If the child vocalized the name of more than one category or exemplar, the experimenter scored only the first response. Reverse intraverbal trials were an exception to this rule, because up to two correct responses were possible in each trial, and the experimenter scored the first two responses.

Interobserver agreement. A second observer independently collected data on at least 57% of all trial blocks during category-name listener and intraverbal testing and at least 50% of all pre- and posttests for each child. We scored an agreement if the experimenter and the observer both recorded a correct or an incorrect response on the same trial; otherwise we scored a disagreement. Point-by-point agreement was calculated for each trial block or test by dividing the number of agreements by the sum of agreements and disagreements and multiplying by 100%. Mean agreement for listener and intraverbal testing was 100% for Marc, Chanele, and Sanjay; 99% for Greg (range, 87% to 100%); and 99% for Erika (range, 87% to 100%). Mean agreement for pre- and posttests for category match to sample, category tacts, and reverse intraverbals was 100% for Marc, Erika, and Chanele; 99% (range, 90% to 100%) for Greg; and 99% (range, 95% to 100%) for Sanjay.

During category-name listener and intraverbal training, we assessed interobserver agreement in the same manner on at least 49% of all trial blocks. Mean agreement across the four-trial training blocks was at least 99% for each participant. During exemplar-name training and maintenance, we assessed agreement on at least 33% of all trial blocks. Mean agreement was at least 99% during training and 99% during maintenance for each participant.

Procedural fidelity. Procedural fidelity was assessed for each child on at least 20% of all trial blocks during category-name listener and intraverbal testing. An independent observer recorded antecedents and consequences delivered by the experimenter on each trial. The observer recorded a test trial as correctly implemented if the experimenter delivered the appropriate instruction and did not deliver any consequences. A procedural fidelity score was computed for each trial block by dividing the number of correctly implemented trials by the total number of trials and multiplying by 100%. Mean procedural fidelity during testing was 100% for all participants.

Procedure

Experimental design. The order of training and testing conditions is shown in Figure 2. Following exemplar-name tact and listener training, each child received either category-name intraverbal training (Marc and Greg) or category-name listener training (Erika, Chanele, and Sanjay) on one (Chanele and Sanjay) or two (Marc, Greg, and Erika) category sets. A concurrent multiple baseline design across stimulus sets (Marc, Greg, and Erika) or a nonconcurrent multiple baseline design across participants (Chanele and Sanjay) was used to evaluate the effects of training on category-name intraverbal and listener relations. Pre- and posttests were used to assess the effects of training on category match to sample, category tacts, and reverse intraverbals.

General procedure. The order of trial presentations was preprogrammed, as was trial type (the location and combination of positive and negative comparison stimuli) in match-to-sample tasks. The pre- and posttests of category tacts and category match to sample each consisted of 20 randomly ordered trials in which the experimenter presented each visual stimulus five times. The reverse intraverbals pre- and posttests consisted of 10 randomly ordered trials in which the experimenter presented each spoken category name five times. In all other

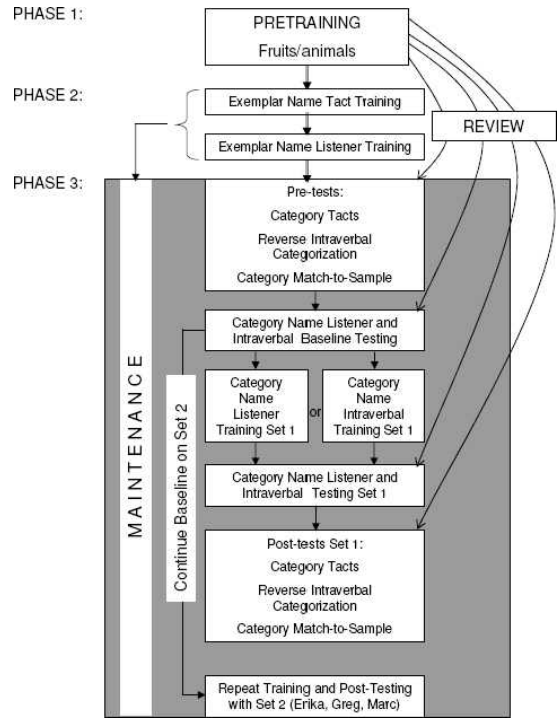

Figure 2. The sequence of training and testing conditions.

conditions, the experimenter had up to five preprogrammed trial blocks available. Presentation order and trial type were programmed to vary unsystematically across trial blocks, with the restriction that positive comparisons appeared an equal number of times in each position across trial blocks in listener training and testing conditions. No intertrial intervals were programmed, but unprogrammed intertrial intervals could occur on occasion.

Phase 1: Pretraining and review with familiar categories. Pretraining was conducted with two exemplars from each of two familiar categories to familiarize the children with the experimental procedures. Each child received exemplar-name tact and listener training, category-name intraverbal and listener training, and training on pre- and posttesting procedures (category match to sample, category tacts, and reverse intraverbals). Stimulus presentation and instructions were identical to those in the

corresponding experimental conditions in Phases 2 and 3. Each pretraining condition continued until the child responded with 100% accuracy in a single four-trial block, except the pretraining of reverse intraverbals in which a trial block consisted of only two trials, but two correct responses were required on each trial to meet criterion.

The experimenter reviewed the testing procedures with pretraining stimuli briefly prior to each testing condition. The experimenter initially presented one block of trials under extinction. If performance was less than 100% accurate, the experimenter conducted pretraining until 100% accuracy was achieved in one trial block followed by 100% accuracy in a second trial block under extinction.

Phase 2: Exemplar-name training and maintenance. In exemplar-name training, the experimenter trained the child first to tact each of 12 stimuli (six maps and six foreign characters) and then to respond as a listener by selecting those same stimuli given their spoken exemplar names. Eight of the stimuli belonged to two different category sets; either Kata–Hira and north–south or Greek–Cyrillic and east–west. The remaining four stimuli were two additional maps and two Hebrew characters. For tact training, the 12 stimuli were divided arbitrarily into three groups of four stimuli. The experimenter conducted training trials in 12-trial blocks and proceeded through the following stages. In Stage 1, the experimenter alternated presentations of the four stimuli in one stimulus group until responding was 100% accurate on one trial block that contained three presentations of each stimulus. The experimenter repeated this procedure in Stages 2 and 3 for the remaining two stimulus groups. Finally, in Stage 4, the experimenter alternated presentations of all 12 stimuli until responding was 100% accurate in three consecutive trial blocks. Training procedures were identical during all stages. On each trial, the experimenter asked "What is this?" and then presented a visual stimulus. The experimenter praised correct responses. When the child made an incorrect response, the experimenter prompted a correct response (e.g., "This is Mali, can you say it?").

Following the completion of tact training, the experimenter conducted one 12-trial block of listener training. The experimenter presented an instruction containing an exemplar name (e.g., "Which one is Mali?"), and then presented a trial sheet containing 3 of the 12 stimuli—one positive comparison and two negative comparisons (one map and one character). The experimenter praised correct responses. Following an incorrect response, the experimenter prompted the child to touch the correct comparison by pointing and saying "It is this one." If the child performed with 100% accuracy, further listener training was omitted. Otherwise, training continued until performance was 100% accurate on three consecutive trial blocks.

Exemplar-name tacts and listener relations were maintained during all subsequent experimental conditions. Maintenance trials were conducted immediately before approximately every other testing or training session; tacts were targeted on approximately half of these occasions and listener relations on the other half. Maintenance procedures were identical to listener training and Stage 4 tact training. If in the first trial block, the child made an error in one or more of the eight trials that targeted category exemplars, training continued until the child responded correctly in all eight trials. Errors made with the remaining four stimuli were corrected but did not result in additional trial blocks as long as the child responded correctly on all trials that targeted category exemplars.

Phase 3: Pre- and posttests. Pretests for category match to sample, category tacts, and reverse intraverbals were conducted prior to the category-name listener and intraverbal baseline. Posttests were conducted following the completion of posttraining category-name intraverbal

and listener testing on each set. All tests were conducted under extinction. The participant received up to 10 s to respond following stimulus presentation, and no consequences were delivered for correct or incorrect responses.

The experimenter initiated a category match-to-sample trial by presenting a trial page with only the sample stimulus visible to the child and by delivering the instruction "Touch this." Once the child had touched the sample stimulus, the experimenter instructed the child to "Find the one that goes with it" and lifted the flap that covered the three comparison stimuli. The positive comparison was a stimulus that belonged to the same category as the sample stimulus (e.g., the other north stimulus if the sample stimulus was a north stimulus). The negative comparisons were (a) a stimulus from the other category in the set (e.g., the south category), and (b) a stimulus previously used in exemplar-name training that did not belong to an experimentally defined category (a map when the sample was a map and a Hebrew character when the sample was a character).

On a category tact trial, the experimenter presented a trial page containing one stimulus and asked "What is this?" If the child responded with the previously trained exemplar name rather than a category name, the experimenter asked "Is it called anything else?" and waited an additional 10 s for a response.

No visual stimuli were used on the reverse intraverbal test. The experimenter initiated a trial by saying, "Name something that is [category name]." If the child responded with two or more exemplar names or did not respond with an exemplar name, the trial was over and the next trial was initiated. If the child initially responded with one correct or incorrect exemplar name, the experimenter instructed the child to "Name something else that is [category name]" and waited an additional 10 s for another response.

Phase 3: Category-name listener and intraverbal testing. Category-name listener and intraverbal relations were tested in baseline and immediately following each category-name training phase. One category set was tested at a time. Each block of test trials consisted of one listener and one intraverbal trial per category exemplar, for a total of eight trials.

At the beginning of each listener trial, the experimenter presented the instruction "Which one is [category name]?" and then presented a trial sheet containing the positive comparison and two negative comparisons (one stimulus from the other category in the set and one stimulus that did not belong to an experimentally defined category but was included in exemplar-name training). In intraverbal trials, no visual stimuli were used. The experimenter delivered the instruction "[Exemplar name] is—" and waited up to 10 s for a response, at which time the experimenter initiated the next trial. The experimenter did not deliver a consequence for correct or incorrect responses.

In baseline, testing continued until stable performance was observed in both listener and intraverbal trials. In the posttraining testing condition, the untrained repertoire was considered successfully acquired if performance was 100% correct (four correct responses) in three consecutive blocks within six blocks of testing. If the criterion was not met, testing continued until responding was stable in the untrained repertoire.

Phase 3: Category-name intraverbal training. During category-name intraverbal training, the child was taught to respond to four spoken exemplar names by vocalizing the appropriate category names (two exemplar names corresponded to each category name). Training consisted of a continuous reinforcement (CRF) condition and an extinction (EXT) condition. In both conditions, trials were presented in four-trial blocks. At the beginning of each trial, the experimenter delivered the instruction "[Exemplar name] is—." All correct responses were praised during CRF. In addition, if the child responded correctly on both of a pair of consecutive trials, the experimenter placed a token on a token board that could hold

4, 8, 12, or 16 tokens. Trial pairs were arranged such that some pairs contained two exemplar names from the same category (e.g., two north trials), whereas others contained one exemplar name from each category (e.g., one north and one south trial). If the child made an incorrect response on one or both trials in a pair, no token was delivered for that pair of trials. The experimenter vocally prompted a correct response (e.g., "north"), and following the presentation of both trials in the pair, the experimenter repeated the trial pair until the child responded correctly on both trials. Early in training, four tokens were required to complete a session. As training progressed, the token board was gradually enlarged up to 16 tokens. The token board and trial-pair procedures were introduced for Greg after the first 100 training trials on each set, whereas these procedures were used from the beginning of training for Marc.

The CRF condition was divided into three stages for Marc. In Stage 1, presentations of two exemplar names (one from each category) were alternated. When Marc obtained 100% accuracy on three consecutive blocks, Stage 2 commenced, in which presentations of the two remaining stimuli were alternated until the same criterion was met. In Stage 3, each four-trial block contained one presentation of each of the four stimuli, and this condition continued until Marc achieved 100% accuracy in five consecutive blocks. Training began at Stage 3 for Greg. The EXT condition began immediately following completion of CRF. As in Stage 3 of CRF, the experimenter presented all four stimuli in four-trial blocks; however, the experimenter presented trials one by one rather than in pairs. No consequences were delivered following correct responses. If the child made an incorrect response, the experimenter vocally prompted a correct response and then presented the next trial. EXT continued until the child achieved 100% accuracy on three consecutive trial blocks. If performance declined, Stage 3 CRF was reintroduced.

Phase 3: Category-name listener training. During category-name listener training, the child was trained to respond to two spoken category names by selecting visual stimuli representing two exemplars from each category. The experimenter presented trials in four-trial blocks. At the beginning of each trial, the experimenter delivered the instruction "Which one is [category name]?" and presented the positive comparison along with two negative comparisons (one from the alternative category and one that did not belong to either category). Listener training was divided into CRF and EXT conditions. The CRF condition was divided into three stages for Erika and Sanjay, corresponding to the three stages of intraverbal training for Marc. Stages 1 and 2 were omitted and training began at Stage 3 for Chanele. The criteria for completing each condition were the same as those during intraverbal training.

During CRF, stimuli were presented on stimulus cards (for Chanele, trial sheets were used initially but were replaced with stimulus cards after 153 trial blocks). The experimenter could either line up the cards on the table or hold them directly in front of the child. As in intraverbal training, the experimenter presented listener trials in pairs. The experimenter praised correct responses and delivered a token when the child responded correctly on both trials in the pair. If the child made an incorrect response, the experimenter pointed to the correct comparison, said "It is this one," and prompted the child to touch the stimulus. The experimenter then repeated the trial pair until the child responded correctly on both trials, The token board and trial-pair procedures were introduced for Chanele after the first 100 trials of training, whereas they were used from the beginning of training for Erika and Sanjay.

During EXT, the stimulus cards were replaced with trial sheets identical to those used during testing, and trials were presented one by one. No consequences were delivered following correct responses, but the experimenter prompt-

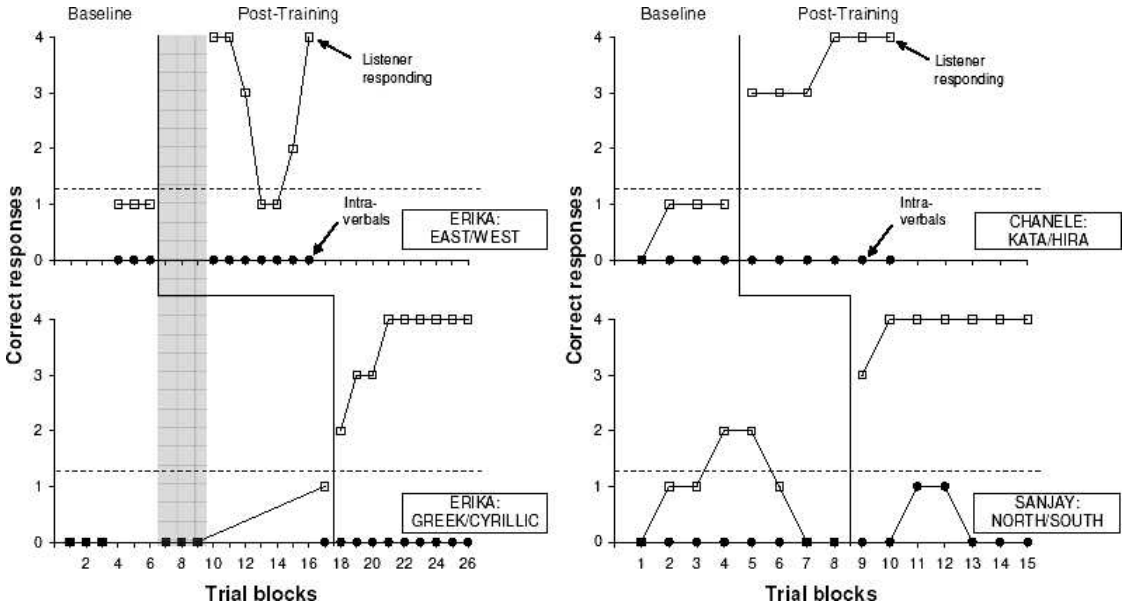

Figure 3. Correct responding on category-name intraverbal (filled circles) and listener (open squares) test trials in baseline and following listener training for Erika, Chanele, and Sanjay. The horizontal dotted line represents chance-level performance on listener trials. The shaded part of the left panels contains baseline data collected on the second category set during training on the first set.

ed a correct response if an incorrect response was made.

RESULTS

Phases 1 and 2. None of the participants required extensive pretraining on experimental procedures; however, all required more than one four-trial block of training to reach criterion in at least one pretraining condition. Erika, Chanele, and Greg required additional training with familiar stimuli on one to two occasions when testing procedures were reviewed immediately prior to testing. Chanele and Marc completed exemplar-name tact training in 43 and 149 12-trial blocks, respectively. Chanele, Sanjay, Marc, and Greg responded with 100% accuracy on the first 12-trial block of exemplar-name listener training and did not require additional training. Erika completed listener training in four blocks. All children maintained high accuracy on exemplar-name tact and listener relations throughout the remainder of the experiment (acquisition and maintenance data are available from the authors).

Phase 3. Sanjay and Chanele required 53 and 233 four-trial blocks, respectively, to complete category-name listener training, whereas Marc and Greg completed category-name intraverbal training in 93 (Greek–Cyrillic set) and 317 (north–south set) blocks, respectively.

Figure 3 shows category-name testing data for children who received category-name listener training. In baseline, all 3 children responded around chance level on listener trials, and no child emitted any intraverbal responses. Following listener training, performance on listener trials in most cases remained highly accurate, except that consistent accurate responding was not obtained for Erika on the east–west set. In no case, however, was an increase in intraverbal responding observed consistently.

Figure 4 depicts category-name testing data for Greg and Marc, who received category-name intraverbal training. Greg responded around chance level on listener trials in baseline

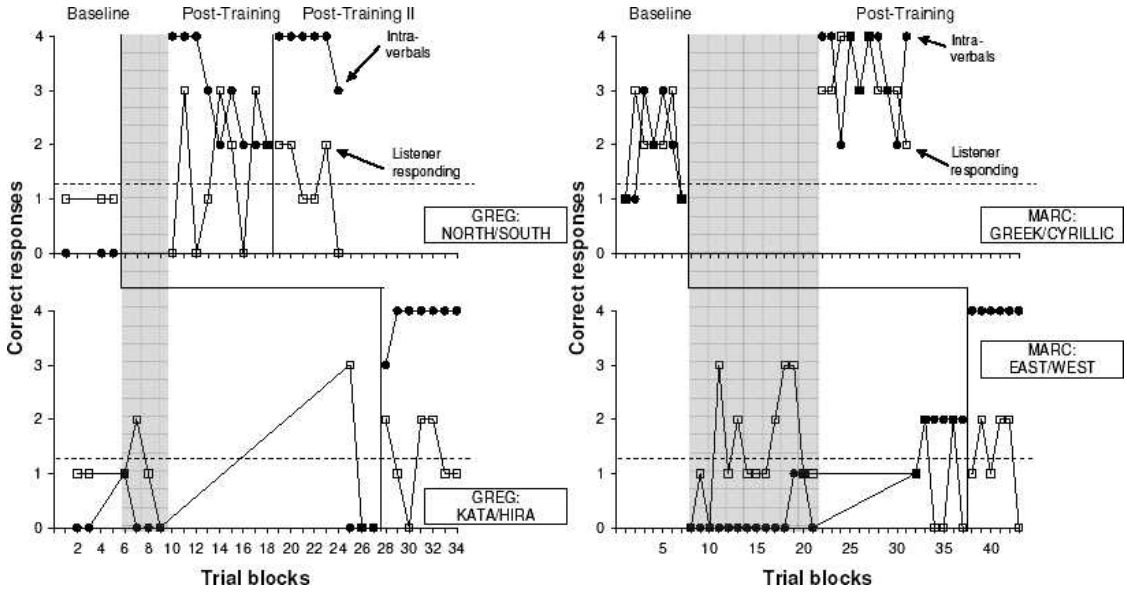

Figure 4. Correct responding on category-name intraverbal (filled circles) and listener (open squares) test trials in baseline and following intraverbal training for Greg and Marc. The horizontal dotted line represents chance-level performance on listener trials. The shaded parts of the graphs contain baseline data collected on the second category set during training on the first set.

and emitted only one correct intraverbal response in the Kata–Hira baseline. Following intraverbal training on the north–south set, Greg's performance on intraverbal trials declined during testing, which resulted in retraining of the intraverbal repertoire. Following retraining, intraverbal performance remained accurate; however, performance in listener trials did not exceed chance level. Following training on the Kata–Hira set, Greg's performance in listener trials also remained at chance level, but intraverbal responding remained accurate.

Marc was the only child who consistently responded with category names on intraverbal baseline trials. In the Greek–Cyrillic baseline, he responded with "Greek" or "Cyrillic" on most trials, resulting in one to three correct responses per block. In the east–west baseline, no intraverbal responding was observed initially, but towards the end of baseline he responded with "east" on most trials, which was the correct response in 50% of all trials. Following intraverbal training on the Greek–Cyrillic set, Marc's intraverbal responding was 100% accurate on 6 of the 10 testing blocks. Accuracy on listener trials also increased from baseline, as during the first nine blocks of testing, Marc responded correctly on three to four trials in each block. This was, however, not sufficient to meet the acquisition criterion. The effect on listener responding was not replicated with the east–west set. Following training on that set, Marc continued to respond at chance levels on listener trials, whereas intraverbal responding remained highly accurate.

Performance on category match-to-sample, category tact, and reverse intraverbal pre- and posttests is shown in Table 1. No substantial increases were observed on any of these measures, except that Marc responded with 80% accuracy on the category tact posttest and with 70% accuracy on the reverse intraverbal posttest on the east–west set.

DISCUSSION

No convincing evidence emerged that listener training resulted in the emergence of an intraverbal repertoire in the context of teaching

Table 1
Percentage of Correct Responses on Pre- and Posttests of Category Match to Sample, Category Tacts, and Reverse Intraverbals

	Erika		Chanele	Sanjay	Marc		Greg	
	Set 1	Set 2			Set 1	Set 2	Set 1	Set 2
Category tacts								
Pretest	0	0	0	0	0	0	0	0
Posttest	0	0	5	20	80	0	40	0
Posttest after retraining on Set 1							0	
Reverse intraverbals								
Pretest	0	0	0	0	0	0	0	5
Posttest	0	0	0	20	70	0	0	5
Posttest after retraining on Set 1							0	
Category match to sample								
Pretest	20	55	30	30	50	50	25	40
Posttest	10	30	30	60	55	40	40	50
Posttest after retraining on Set 1							50	

categorization skills to young children or vice versa. Further, there was little evidence that either type of training produced any additional categorization skills. The only exception occurred for Marc, whose performance on tests of category-name listener relations, category tacts, and reverse intraverbals improved following intraverbal training on one stimulus set. However, experimental control was not achieved over this effect, in that it was not observed following training on a second set. These findings were obtained in spite of the fact that (a) tacts and listener relations with respect to exemplar names were maintained at high levels of accuracy throughout all experimental conditions, and (b) the participants received pretraining on all experimental instructions.

The absence of an effect of listener training on intraverbal responding is consistent with the findings of Miguel et al. (2005). By contrast, previously observed effects of intraverbal training on the emergence of listener relations (Luciano, 1986) were not replicated in the present study, in that only one of four instances of training appeared to affect the listener repertoire. A number of possible reasons exist for this discrepancy. First, the presence of visual stimuli as prompts during intraverbal training may have facilitated the concurrent emergence of listener relations in Luciano's study. Second, the participants in Luciano's study were trained to vocalize exemplar names when presented with a category name, whereas the reverse intraverbal relation was trained in the present study. In Luciano's study, therefore, the trained intraverbal response had the same topography as an existing tact, which was not the case in the present study. Third, intraverbal training in the present study involved establishing the same intraverbal response in the presence of two different S^Ds (many-to-one training), whereas Luciano trained multiple intraverbal responses to the same S^D (one-to-many training). Most empirical research on intraverbal training to date has been conducted in the context of teaching categorization skills (Braam & Poling, 1983; Luciano; Miguel et al.; Partington & Bailey, 1993), and thus the trained intraverbal relations have necessarily been one-to-many or many-to-one relations. Other teaching programs that include both listener and intraverbal components may, however, involve one-to-one relations (e.g., animal sounds). Future research might explore further the training of one-to-one, many-to-one, and one-to-many intraverbal relations and investigate how such structures might interact with the direction of intraverbal relation trained.

In addition to category-name intraverbals, two additional vocal relations were tested: reverse intraverbals and category tacts. With

the exception of Marc's performance on the first stimulus set, these relations did not emerge as a result of either type of training. These data are consistent with prior research indicating that listener training does not necessarily generate intraverbals (Miguel et al., 2005) or tacts (Connell & McReynolds, 1981; Horne et al., 2004), and intraverbal training does not necessarily generate tacts (Watkins, Pack-Teixteira, & Howard, 1989) or reverse intraverbal relations (Perez-Gonzalez, Garcia-Asenjo, Williams, & Carnerero, 2007). The only instance in which an increase in performance on those two tests was observed was also the only instance in which training appeared to affect the primary dependent variable. The data are therefore consistent with prior research (Horne et al., 2004, 2006; Lowe et al., 2002, 2005; Miguel et al., in press) in which covariation was observed between the emergence of different relations. It may be speculated that in Marc's case, the presence of reverse intraverbals and category tacts mediated or facilitated the emergence of listener relations, or alternatively, that the emergence of all three relations exemplified some other process analogous to that by which responding in accordance with symmetry and transitivity arises following match-to-sample training in research on stimulus equivalence (see Sidman & Tailby, 1982).

The final category skill tested was category match to sample. An increase in accuracy above chance levels on this test was never observed following either intraverbal or listener training. It may be surprising that category match to sample did not emerge as a result of listener training, given that a similar type of auditory-visual match-to-sample training has tended to reliably generate such performances in the stimulus equivalence literature, even among children as young as those in the present study (e.g., Devany, Hayes, & Nelson, 1986). Horne et al. (2004, 2006), however, observed that such performances were correlated with the extent to which listener training produced common (category) tacts of the stimuli, which it never did in the present study. Further, because the study was designed to simulate the type of training that children with autism might receive in an EIBI program, training and testing arrangements differed somewhat from common practices in research on stimulus equivalence (see Green & Saunders, 1998). It may be noted that at least one EIBI textbook (Maurice et al., 1996) recommends teaching category match to sample prior to teaching relations involving category names. Future research might assess the extent to which that sequence produces an overall saving of trials or other benefits.

Although the training of intraverbal and listener relations in the present study tended not to generate any emergent relations, exemplar-name tact training immediately generated 100% accurate performance on exemplar-name listener trials for 4 of the 5 children. A practical implication may be that for children who readily acquire listener relations as a result of tact training (e.g., Wynn & Smith, 2003), expressive programs that target intraverbals may not necessarily have the same effect. This finding may reflect a developmental trend or increased complexity of the task.

Some limitations of the present study should be highlighted. First, the study was designed only to evaluate the effects of each type of training and not to directly compare the two. Thus, it is necessary to interpret the data on Marc's performance on the Greek–Cyrillic set with caution and not take them as indicative of a superiority of intraverbal over listener training. Second, the purpose of the study was merely to evaluate the initial effects of training procedures similar to those described in EIBI curricula. Therefore, no attempt was made to manipulate additional variables to establish the untrained repertoires, either with or without direct reinforcement, when they failed to emerge as a result of the training of other repertoires. Third, the duration of Chanele's and Sanjay's participation was too brief to allow

a within-subject demonstration of experimental control. Instead, a nonconcurrent multiple baseline design was used, which may have yielded a lesser degree of experimental control than did the concurrent multiple baselines across stimulus sets employed with Marc, Greg, and Erika. Fourth, a large number of training trials were required in most cases to establish criterion performance during both listener and intraverbal training. During those conditions, each child also was required to maintain the large number of arbitrary relations previously established during exemplar-name training, which may have impeded acquisition of additional relations. Steps taken to remedy slow acquisition included introducing a token system and training the discriminations sequentially. The children who were exposed to those procedures from the beginning of training (Marc, Erika, and Sanjay) in all cases achieved the criterion faster than did Greg and Chanele, for whom the procedures were implemented when they failed to make progress without them. However, it is possible that even faster acquisition could have been achieved with the use of errorless learning procedures or by using preference assessments to identify more powerful back-up reinforcers. A final limitation, from an experimental standpoint, is that all training and testing was accomplished using a tabletop procedure. Such procedures may introduce inconsistencies in the timing of stimulus presentation and consequence delivery, along with raising the possibility of experimenter cueing. At present, no technology is available that would allow fully automated training of vocal responses when such training requires immediate scoring and consequence delivery of individual responses. To minimize the potential impact of human error and bias, future research on vocal relations might attempt to combine automated procedures with the use of human observers who communicate remotely with experiment-running software.

These limitations notwithstanding, the present data provide evidence that histories of intraverbal or listener training, analogous to those established in programs designed to teach categorization skills in EIBI curricula, may not easily yield collateral benefits in the form of the emergence of untrained relations. Given that the language-able children in the present study generally acquired only directly reinforced relations, it appears unlikely that children in need of language interventions would acquire untrained relations under similar training arrangements. Consequently, the results do not at present suggest any specific benefits of teaching listener and intraverbal relations in a particular sequence. It remains possible, however, that the training of one type of relation may serve to facilitate the subsequent training of another, resulting in an overall savings of trials. This possibility remains to be evaluated, as do the effects of interspersing intraverbal and listener trials during training, as recommended by some EIBI practitioners (Sundberg & Partington, 1998). In addition, it may be possible to arrange training in ways that are more likely to bring about the emergence of additional relations. A clinically important area of future research may be to evaluate ways to overcome the independence of speaker and listener repertoires via specific reinforcement histories. One candidate for such a history may be multiple-exemplar training, which has been found to result in listener training generating tacts (Greer, Stolfi, Chavez-Brown, & Rivera-Valdes, 2005), tact training generating mands (Nuzzolo-Gomez & Greer, 2004), and intraverbal training generating reverse intraverbal relations (Perez-Gonzales et al., 2007). Future research might evaluate such procedures in the context of intraverbals and listener relations as well.

REFERENCES

Barnes-Holmes, D., Barnes-Holmes, Y., & Cullinan, V. (2000). Relational frame theory and Skinner's *Verbal Behavior*: A possible synthesis. *The Behavior Analyst*, *23*, 69–84.

Braam, S. J., & Poling, A. (1983). Development of intraverbal behavior in mentally retarded individuals through transfer of stimulus control procedures: Classification of verbal responses. *Applied Research in Mental Retardation*, *4*, 279–302.

Connell, P. J., & McReynolds, L. (1981). An experimental analysis of children's generalization during lexical learning: Comprehension or production. *Applied Psycholinguistics*, *2*, 309–332.

Cuvo, A. J., & Riva, M. T. (1980). Generalization and transfer between comprehension and production: A comparison of retarded and nonretarded persons. *Journal of Applied Behavior Analysis*, *13*, 315–331.

Devany, J. M., Hayes, S. C., & Nelson, R. O. (1986). Equivalence-class formation in language-able and language-disabled children. *Journal of the Experimental Analysis of Behavior*, *46*, 243–257.

Fraser, C., Bellugi, U., & Brown, R. (1963). Control of grammar in imitation, comprehension, and production. *Journal of Verbal Learning and Verbal Behavior*, *2*, 121–135.

Green, G., & Saunders, R. R. (1998). Stimulus equivalence. In K. A. Lattal & M. Perone (Eds.), *Handbook of research methods in human operant behavior* (pp. 229–262). New York: Plenum.

Greer, R. D., Stolfi, L., Chavez-Brown, M., & Rivera-Valdes, C. (2005). The emergence of the listener to speaker component of naming in children as a function of multiple exemplar instruction. *The Analysis of Verbal Behavior*, *21*, 123–134.

Guess, D., & Baer, D. M. (1973). An analysis of individual differences in generalization between receptive and productive language in retarded children. *Journal of Applied Behavior Analysis*, *6*, 311–329.

Horne, P. J., Hughes, J. C., & Lowe, C. F. (2006). Naming and categorization in young children: IV. Listener behavior training and transfer of function. *Journal of the Experimental Analysis of Behavior*, *85*, 247–273.

Horne, P. J., & Lowe, C. F. (1996). On the origins of naming and other symbolic behavior. *Journal of the Experimental Analysis of Behavior*, *65*, 185–241.

Horne, P. J., Lowe, C. F., & Randle, V. R. L. (2004). Naming and categorization in young children: II. Listener behavior training. *Journal of the Experimental Analysis of Behavior*, *81*, 267–288.

Keller, M. F., & Bucher, B. D. (1979). Transfer between receptive and productive language in developmentally disabled children [Abstract]. *Journal of Applied Behavior Analysis*, *12*, 311.

Laraway, S., Snycerski, S., Michael, J., & Poling, A. (2003). Motivating operations and the terms to describe them: Some further refinements. *Journal of Applied Behavior Analysis*, *36*, 407–414.

Leaf, R., & McEachin, J. (1999). *A work in progress. Behavior management strategies and a curriculum for intensive behavioral treatment of autism*. New York: DRL.

Lee, V. L. (1981). Prepositional phrases spoken and heard. *Journal of the Experimental Analysis of Behavior*, *35*, 227–242.

Lovaas, O. I. (1977). *The autistic child*. New York: Irvine.

Lovaas, O. I. (1981). *Teaching developmentally disabled children: The ME book*. Austin, TX: Pro-Ed.

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

Lowe, C. F., Horne, P. J., Harris, F. D. A., & Randle, V. R. L. (2002). Naming and categorization in young children: Vocal tact training. *Journal of the Experimental Analysis of Behavior*, *78*, 527–549.

Lowe, C. F., Horne, P. J., & Hughes, J. C. (2005). Naming and categorization in young children: III. Vocal tact training and transfer of function. *Journal of the Experimental Analysis of Behavior*, *83*, 47–65.

Luciano, M. C. (1986). Acquisition, maintenance, and generalization of productive intraverbal behavior through a transfer of stimulus control procedures. *Applied Research in Mental Retardation*, *7*, 1–20.

Maurice, C., Green, G., & Luce, S. (Eds.). (1996). *Behavioral intervention for young children with autism: A manual for parents and professionals*. Austin, TX: Pro-Ed.

Michael, J. (1985). Two kinds of verbal behavior plus a possible third. *The Analysis of Verbal Behavior*, *3*, 1–4.

Miguel, C. M., Petursdottir, A. I., & Carr, J. E. (2005). The effects of multiple-tact and receptive-discrimination training on the acquisition of intraverbal behavior. *The Analysis of Verbal Behavior*, *21*, 27–41.

Miguel, C. F., Petursdottir, A. I., Carr, J. E., & Michael, J. (in press). The role of naming in stimulus categorization by preschool children. *Journal of the Experimental Analysis of Behavior*.

Miller, M. A., Cuvo, A. J., & Borakove, L. (1977). Teaching naming of coin values—Comprehension before production versus production alone [Abstract]. *Journal of Applied Behavior Analysis*, *10*, 735–736.

Nuzzolo-Gomez, R., & Greer, R. D. (2004). Emergence of untaught mands or tacts of novel adjective-object pairs as a function of instructional history. *The Analysis of Verbal Behavior*, *20*, 63–76.

Partington, J. W., & Bailey, J. S. (1993). Teaching intraverbal behavior to preschool children. *The Analysis of Verbal Behavior*, *11*, 9–18.

Perez-Gonzalez, L. A., Garcia-Asenjo, L., Williams, G., & Carnerero, J. (2007). Emergence of intraverbal antonyms in children with pervasive developmental disorder. *Journal of Applied Behavior Analysis*, *40*, 697–701.

Sidman, M., & Tailby, W. (1982). Conditional discrimination vs. matching to sample: An expansion of the testing paradigm. *Journal of the Experimental Analysis of Behavior*, *37*, 5–22.

Skinner, B. F. (1957). *Verbal behavior*. Acton, MA: Copley.

Smeets, P. M. (1978). Establishing generative performance and cross modal generalization of the manual plural sign in a severely retarded deaf girl. *British Journal of Communication Disorders*, *13*, 49–57.

Smeets, P. M., & Striefel, S. (1976). Acquisition and cross modal generalization of receptive and expressive signing skills in a retarded deaf girl. *Journal of Mental Deficiency Research*, *20*, 251–259.

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

Watkins, C. L., Pack-Teixteira, L., & Howard, J. S. (1989). Teaching intraverbal behavior to severely retarded children. *The Analysis of Verbal Behavior*, *7*, 69–81.

Watters, R. G., Wheeler, L. J., & Watters, W. E. (1981). The relative efficiency of two orders for training autistic children in the expressive and receptive use of manual signs. *Journal of Communication Disorders*, *14*, 273–285.

Williams, G. C., & McReynolds, L. V. (1975). The relationship between discrimination and articulation training in children with misarticulations. *Journal of Speech and Hearing Research*, *18*, 401–412.

Wynn, J. W., & Smith, T. (2003). Generalization between receptive and expressive language in young children with autism. *Behavioral Interventions*, *18*, 245–266.

Received August 1, 2006
Final acceptance February 22, 2007
Action Editor, Mark Dixon