RESEARCH ARTICLE

An Abbreviated Evaluation of the Efficiency of Listener and Tact Instruction for Children with Autism

Sarah E. Frampton¹ · Hannah C. Robinson¹ · Daniel E. Conine¹ · Caitlin H. Delfs^{1,2}

Published online: 13 February 2017
© Association for Behavior Analysis International 2017

Abstract We assessed the efficiency of tact and listener training for eight participants with autism spectrum disorder. Tact and listener probes were conducted in baseline for all target sets, and then tact training was initiated with one and listener training with another. Following mastery of one set, tact and listener probes were conducted with only the sets assigned to the same modality of training (i.e., sets 1, 3, and 5 for tact; sets 2, 4, and 6 for listener). Training and probes were repeated for all sets. The measures of efficiency included the number of skills mastered through direct training, the number of skills that emerged without training, the number of trials-to-criterion, and maintenance of skills. Clinical programming based on each participant's results is discussed. For six participants, tact training was more efficient than listener training across

multiple measures. For the remaining two participants, tact training and listener training were considered equivalent.

Keywords Efficiency · Emergent responding · Listener responding · Tact · Verbal behavior

Addressing language skills for children with autism spectrum disorder (ASD) is of critical importance as success in this domain appears to be a strong predictor of better overall outcomes (Venter, Lord, Schopler, & 1992; Szatmari, Bryson, Boyle, Streiner, Duku, & 2003). Skinner (1957) provided a conceptual framework to analyze deficits in verbal behavior, and from that foundation, behavioral approaches to language training have grown (Sundberg & Michael, 2001). Two major areas of behavioral language intervention are listener responding and tact training. Listener training often begins with teaching learners to select common items after hearing the names of those items. In contrast, tact training often begins with teaching the learner to state the names of common objects. These programs are commonly included in early intensive behavioral intervention (EIBI) curricula; however, the sequence in which they should be introduced has been on area of ongoing empirical investigation (see Petursdottir & Carr, 2011).

The relationship between listener and tact repertoires is complex. Though these repertoires are usually initially independent (Skinner, 1957) they can rapidly become interdependent. Recent theoretical accounts propose that these repertoires become interdependent throughout the course of typical development and use the term *naming* to describe that functional interdependence (Horne & Lowe, 1996). As children's verbal repertoires become more sophisticated, they may behave as both speaker and listener in each interaction with their environment. For example, upon being told to "point to the ball," the child may emit overt or covert echoic responses

Danny Conine is now at the University of Florida.

Implications for Practice

- Tact and listener comparisons can be conducted in the context of daily clinical services, by assigned clinical staff, with individualized procedures for each participant.
- Probe and training procedures can be adapted to require less intrusive data collection while still allowing clinicians to make data-based decisions regarding the outcomes.
- Tact training was shown to be more efficient or as efficient for all participants; however, differences across measures were detected for most participants.
- The proposed decision-making model may assist clinicians in using their data to guide clinical programming.

✉ Sarah E. Frampton
sarah.frampton@choa.org

¹ Marcus Autism Center, 1920 Briarcliff Road NE,
Atlanta, GA 30329, USA

² Emory University School of Medicine, Atlanta, GA, USA

(e.g., “ball”) as they orient to the ball. Any reinforcement subsequently provided will serve to simultaneously strengthen the speaker behavior of overtly or covertly tacting the ball in its presence as well as the listener behavior of selecting the ball. Similarly, when asked to tact the ball, the child will orient toward the ball while hearing their own tact response (e.g., “ball”). Subsequent reinforcement will simultaneously strengthen the tact as well as the listener behavior of orienting to the ball upon hearing the word “ball.” Horne and Lowe (1996) proposed that this interdependence allows children to acquire both tact and listener responses through instruction in either form. This suggestion has been supported in a variety of experimental studies with typically developing children (e.g., Lowe, Horne, Harris, & Randle, 2002; Horne, Lowe, & Randle, 2004; Lowe, Horne, & Hughes, 2005; Miguel, Petursdottir, Carr, & Michael, 2008) and children with ASD (e.g., Miguel & Kobari-Wright, 2013; Kobari-Wright & Miguel, 2014; Lee, Miguel, Darcey, & Jennings, 2015). Naming also allows children to acquire tact and listener skills through language pairings (e.g., the instructor labels an object in the presence of the learner), as demonstrated in experimental studies with typically developing children (e.g., Gilic & Greer, 2011; Longano & Greer, 2014) and children with ASD (e.g., Fiorile & Greer, 2007; Longano & Greer, 2014).

Existing research suggests that tact training is more likely to produce the emergence of listener selection responses than vice-versa; however, idiosyncrasies reported in the literature give clinicians cause to make this determination for their individual clients (for a review of comparison studies, see Petursdottir & Carr, 2011). Several recent experimental studies have specifically compared tact and listener training with participants with ASD (Wynn & Smith, 2003; Sprinkle & Miguel, 2012; Delfs, Conine, Frampton, Shillingsburg, & Robinson, 2014). Across studies, sets of stimuli were randomly assigned to either tact or listener training and probes were conducted to evaluate the emergence of the untrained responses. In all three studies, tact training led to emergent listener responses more often than listener training led to emergent tact responses.

However, the results of all three studies were also variable across and within participants, suggesting that the overall relative superiority of tact training may not constitute a sufficient basis for EIBI clinicians to make decisions about sequencing tact and listener training programs. First, for one participant in the Wynn and Smith (2003) study, listener training appeared to produce more emergent relations than tact training. For three other participants, listener training did produce emergent relations, though less consistently than tact training. These results suggest that individualized assessment may still be necessary to detect individuals for whom the above assumptions are not true. Second, emergent relations were equivalent across training conditions for some subjects in Delfs et al. (2014) and Sprinkle and Miguel (2012). In such cases,

determining the most efficient training condition would require the evaluation of additional criteria such as trials-to-criterion, maintenance, and learner preference (Wolery, Ault, Doyle, 1992). Third, in the Wynn and Smith and Delfs et al. studies, participants showed variation in their responsiveness to tact and listener intervention over time. Thus, it may be necessary for clinicians to conduct repeated assessments with individual learners to determine whether patterns of emergence may change over time. Finally, the trials-to-criterion data from the Delfs et al. and Sprinkle and Miguel (2012) study highlight that tact and listener training for individuals with ASD may be a lengthy process. Therefore, clinicians must carefully weigh not only the likelihood of producing emergence of untrained skills but also the effectiveness of tact or listener training in terms of mastery of the explicitly targeted relations. Therefore, while the consensus of the published literature suggests that tact training is generally more efficient in terms of emergent responding, clinicians still may need to assess the relative efficiency of listener and tact instruction for individual learners on an ongoing basis using broader criteria.

Examining the most recently published comparison study by Delfs et al. (2014), some procedural limitations likely limit widespread clinical implementation. In the Delfs et al. study, probe sessions were used to determine mastery. During these probe sessions, each target was presented multiple times, with a range of 32–48 total trials per session. Similarly, during daily probe sessions, which were used to determine completion of the training conditions, each target was presented 3–4 times and in both modalities. When considering the constraints on time and resources faced by EIBI clinicians, decreasing the time spent probing responses relative to the time spent training would likely enhance the feasibility of an assessment procedure. Furthermore, Delfs et al. used a parallel treatment design that occasionally required a mastered set to remain in training until the paired set met mastery or termination criteria. For clinicians, continuing to conduct intensive training sessions for targets that have reached mastery criteria may be considered an inefficient use of resources. Lastly, trial-by-trial data was collected during all phases of the study. This may be overly burdensome in some clinical settings. While the Delfs et al. paper incorporated some individualized procedures for each participant to increase the acceptability of the procedures (e.g., array sizes, number of training trials), modifications to the procedures as a whole would likely be necessary before they could be integrated into ongoing clinical service provision.

The purpose of the current study was to replicate and extend the procedures of Delfs et al. (2014) while incorporating modifications aimed at increasing clinical utility. Specifically, the procedures implemented by Delfs et al. were used with the following modifications: (1) reduced number of probe trials, (2) use of cold probe data collection rather than trial-by-trial data collection, and (3) use of a multiple probe

design across behaviors to limit unnecessary training sessions. In addition, sessions were conducted within the context of clinical services, with the assigned clinical teams conducting the sessions and the case managers overseeing the services making decisions related to individualized components of the training procedures. This was specifically done to evaluate how easily the procedures used by Delfs et al. could be incorporated into clinical programming. The primary measures of efficiency were emergence of skills, trials-to-criterion, and maintenance over time.

Method

Participants and Setting

The current study was conducted with eight participants, ranging in age from 2 years 8 months to 15 years 3 months (see Table 1). All participants were male and reportedly diagnosed with ASD. Mark, Paul, and Mike had diagnoses confirmed through independent psychological testing. All participants received services through a clinic-based program focused on teaching functional verbal behavior. Participants attended the clinic for 2 to 3 h each day, 5 days a week. The length of previous language intervention in the clinic ranged from 15 days to 5 months with a mean length of 2 months.

All sessions took place in the clinic, in a classroom with other children and therapists present. Sessions were conducted by the therapists that typically conducted clinical treatment sessions with the participants. The participants' individual teaching areas contained a table, chairs, preferred toys and foods, and other classroom materials. Sessions and therapists were supervised by a Board Certified Behavior Analyst (BCBA) and a licensed psychologist.

Upon admission to the clinic, the Verbal Behavior–Milestones Assessment and Placement Program (VB-MAPP; Sundberg, 2008) was administered by the BCBA and trained therapists. Clients were referred by the BCBA for inclusion in the current study. Inclusion criteria, based loosely on recommendations from

Miguel and Petursdottir (2009), included a generalized echoic repertoire and previously mastered targets in both the tact and listener domains (i.e., effective teaching methods had been established by the clinical teams). There were no specific exclusion criteria.

Stimuli Selection

Eighteen unknown pictures were identified for five participants whereas 12 items were identified for the other three participants. Clinical teams were directed to select targets that would be appropriate for the participant based on current skills and programming. The clinical teams were directed to select targets of similar (1) familiarity to the participant and (2) complexity (e.g., number of letters and syllables). Half of the targets were assigned to listener training, and the other half were assigned to tact training. Of the 6–9 targets assigned to each condition, they were divided into 2–3 training sets. Sets 1, 3, and 5 were always taught as tact targets and sets 2, 4, and 6 were always taught as listener targets. If only two sets were assigned per condition, then sets 1 through 4 were assigned in the same manner. Clinical teams were directed to counterbalance the complexity of targets within and across sets based on familiarity and complexity. See Table 2 for an example of target assignments for one participant.

Response Measurement

Correct tact responses and *correct listener responses* served as the primary dependent variables. A *correct tact response* was defined as the participant vocally stating the name of an object or a pre-identified approximation (e.g., “gwill” for “grill”) within 3 or 5 s of the discriminative stimulus (S^D) presentation (i.e., the picture of the object and the vocal stimulus “What is it?”). The amount of time allowed for responding (e.g., 3 or 5 s) depended on the participant's typical programming and prompting strategies. This interval was kept constant within participants and across conditions. A *correct listener response* was defined as the participant touching the correct picture within 3 or 5 s of

Table 1 Participant characteristics

Participant	Age	Months in treatment prior to start	VB-MAPP milestones score	Tact domain score	Listener domain score
Kevin	4	1.5	71	7	6
Josh	8	1.5	50.5	6	7.5
Mark	4	1	28.5	3	2.5
Jake	15	1.5	107	8	8.5
Yale	8	2.5	63	7	8
Neal	7	1	130.5	13	11
Mike	4	2	66.5	7	5
Paul	2	5	54.5	6.5	4.5

Table 2 Example of targets assigned to groups (Mike)

Tact training		Listener training	
Set 1	Envelope Cutting board Iron	Set 2	Outlet Mixer Oven mitt
Set 3	Vase Lightbulb Binder	Set 4	Whisk Flute Lawn mower
Set 5	Can opener Saw Trumpet	Set 6	Pliers Onion Spatula

the S^D presentation (i.e., an array of pictures and the vocal stimulus “point to the...” or “show me the...” followed by the name of the object).

Experimental Design

Multiple probe designs across behaviors (sets) were used to evaluate the effects of two different instructional methods (i.e.,

listener training, tact training) on the acquisition of both tact and listener responses. Probe and training sessions were conducted in accordance with the logic of the multiple probe design within each group of sets (e.g., sets 1, 3, and 5 assigned to tact training and sets 2, 4, and 6 assigned to listener training). For example, if mastery criteria were met for set 1 (i.e., tact training), full probes were conducted only for sets assigned to tact training (see Fig. 1). Probes were not conducted for the sets assigned to listener training until mastery criteria were met for set 2, which was assigned to listener training. Thus, the assessment could progress independently for each training condition.

General Procedures

During all phases, preferred items were identified prior to and periodically throughout each session through the participants’ vocal requests. The therapist either waited for the participant to request an item present in the environment or provided a general vocal prompt to evoke requesting (e.g., “What do you want to work for?”) before initiating a session. These methods were

Fig. 1 Flowchart of training and probe conditions

similar to those used in all participants' clinical programming (Delfs & Frampton, 2014). Requested items were used in the ensuing session.

Mastered targets drawn from the participants' clinical programming were interspersed throughout all phases of the study; targets from the current experiment were not included. The mastered targets could include any response type (e.g., tact, listener selection, motor imitation). These varied each session and were selected semi-randomly by the therapist by drawing a card that indicated which target was to be presented. Consequences for compliance with mastered targets varied according to the phase, but general praise was always provided.

The experimenters consulted with each participant's clinical case manager to individualize several elements of the procedures; the clinical case managers were directed to select procedures consistent with typical clinical programming. The elements that were individualized included (1) array sizes, (2) reinforcement schedules, (3) and the number of training trials per session. See Table 3 for a breakdown of these variations across participants.

Baseline

The baseline phase was conducted to ensure that no targets were mastered as either tact or listener responses prior to training. A target was considered mastered, if correct responses occurred on at least two out of three trials. Additionally, one other target in the set had to occur at or above this level of proficiency to ensure responding was discriminated and prevent false mastery (e.g., emitting the same response on every trial). All targets assigned to each training condition were evaluated (see Fig. 1). All targets in all sets were initially probed on one trial as a tact. Next, each target in each set was probed on one trial as a listener skill. This continued until each target was probed a total of three times as a tact response and three times as a listener response. The order of trial presentation was quasi-randomized on the data sheet throughout the baseline condition.

When presenting tact trials, the therapist held up a picture and asked, "What is it?" When presenting listener trials, pictures were presented in an array of 3 or 6. When an array of 6 was used, target cards from both the tact and listener sets were included in this array. When an array of 3 was used, the only cards that were presented were the listener target cards in that set. Thus, in both arrangements, all target stimuli served as both S^D s for correct responding on some trials, and as S^Δ on all other trials. Once the cards were placed in an array, the therapist said, "point to the..." or "show me the..." followed by the item name. The therapist then waited 3 or 5 s for a response. No programmed consequences were provided for correct or incorrect responses in baseline. Mastered targets were interspersed throughout the probes to maintain responding, and reinforcement for mastered tasks was provided based on the reinforcement schedule used in the participant's clinical programming.

Tact and Listener Training

For tact and listener training, the therapist presented targets in the same manner as in baseline. Contingent on a correct response, the therapist delivered general praise and a preferred tangible item. Contingent on an incorrect response or no response, the therapist immediately presented two prompted trials to evoke two correct prompted responses. On prompted trials, a prompt was delivered immediately following the S^D for that trial. For example, on a tact trial the therapist repeated the vocal instruction, "What is it?" then immediately used an echoic prompt to occasion a correct vocal response, "wrench." On a listener trial, the therapist repeated the vocal instruction, "Point to whisk," then immediately pointed at the target stimulus to occasion a correct selection response. If the participant did not respond to the gestural prompt within 2 s a more intrusive prompt was used (e.g., a model prompt or physical guidance). After the two consecutive prompted trials, another independent opportunity (hereafter referred to as a *transfer trial*) was provided. On the transfer trial, the therapist allowed 3–5 s for the participant to respond. Contingent on a correct response during

Table 3 Participant-specific treatment components

Participant	Visual array size	Differential reinforcement on transfer trial	Number of training trials	Tact targets Average number of letters per word	Listener targets Average number of letters per word
Kevin	3	No	5	6.33	7.55
Josh	3	Yes	5	6.00	5.33
Mark	6	Yes	10	5.55	6.00
Jake	6	No	5	6.22	7.00
Yale	6	No	5	6.83	6.67
Neal	6	Yes	5	5.78	7.11
Mike	6	Yes	5	6.89	6.22
Paul	6	No	5	5.16	5.16

transfer trials, the therapist provided praise and/or a preferred tangible item (see Table 3).

During training, the therapist checked off a box on a data sheet following the completion of each initial independent opportunity and data sheets were individualized to have the corresponding number of boxes (see Appendix Table 5 for example data sheet). Thus, no data were collected on whether responses during training were correct or incorrect, only that the target number of trials was conducted. Cold probe data rather than trial-by-trial data were used to determine the completion of the training phase, which allowed evaluation of whether cold probe data collection could be used while still maintaining the integrity of the tact and listener comparison. Training trials, along with mastered tasks, were interspersed throughout the participants' 2- to 3-h therapy sessions. Clinical teams were directed to conduct the trials as a routine part of the participant's clinical programming; thus, the intervals between trials varied day-to-day across participants. The number of training trials varied from 5 to 10 (see Table 3) and was kept constant across conditions (i.e., if 5 trials were conducted in the tact condition, 5 trials were also conducted in the listener condition).

Cold Probes To determine if the training phase was complete, cold probe data were collected for the sets of targets currently in training. Cold probe data were collected prior to the initiation of training sessions. The therapist presented the targets as in baseline, providing no programmed consequences for correct or incorrect responses. Each target in the training set was probed one time and only in the modality undergoing training (see Appendix Table 5 for example data sheet). That is, tact responses were probed only as tacts and listener responses were probed only as listener responses. Completion of training was determined based on meeting the mastery criteria or the termination criteria. The mastery criterion consisted of two consecutive days with correct responses during cold probes for all targets in the set. The termination criterion was met if the total number of cold probe sessions reached twice the maximum number of sessions that had been conducted with a previous set of targets (in either operant). The therapist then consulted with the case manager to examine the trend of the last several cold probe data points. If there was an increasing trend, the therapist conducted two additional sessions and then reassessed for trend. If no increasing trend was detected, the training phase ended. For example, if the mastery criterion was met for set 1 after 5 teaching sessions and the mastery criterion were met after 8 sessions for set 3, if mastery was not achieved after 16 sessions for set 5, then the case manager would examine the data to check for an increasing trend. If no increasing trend was detected, training would cease as, the termination criteria were met.

Full Probes

To evaluate emergence of untrained responses and maintenance of trained and emergent targets, full probes were conducted following the completion of training, regardless of whether the

mastery or termination criteria were met. Procedures were identical to baseline with the exception that only the sets assigned to the particular training condition were probed. Thus, if the mastery or termination criteria were met for set 1 in tact training, full probes were conducted with sets 1, 3, and 5 only. Training continued with additional sets following full probes, unless all target sets had already completed training.

Interobserver Agreement

A second trained independent observer collected data using pen and paper during a percentage of baseline sessions, full probes, and cold probes for 6 of the 8 participants. The observer recorded the number of correct or incorrect responses during probe sessions. Trial-by-trial interobserver agreement (IOA) was calculated for each participant by taking the number of trials with agreements and dividing by the total number of trials with agreements and disagreements then multiplying by 100. No reliability data were collected for Josh or Neal for this protocol. Excluding these two participants, IOA data was collected for an average of 47% (range 27–75%) of sessions across participants. The percentage of agreements averaged 98% (range 95–100%) for the participants for whom these data were collected.

Fidelity of Implementation

Procedural fidelity data were also collected by a trained independent observer across all phases. Procedures were outlined in a step-by-step format in a checklist and the independent observer recorded whether the therapist followed the procedures exactly as detailed (scored as a "+") or with any deviation (scored as a "-"). Scores were calculated by dividing the number of correctly implemented steps by the total number of steps then multiplying by 100. Procedural fidelity data were collected for four of the eight participants. The average percentage of sessions with procedural fidelity across those four participants was 30% (range from 23 to 44%). The percentage of correct implementation averaged 99% (range 98–100%) for the participants for whom these data were collected.

Results

Figures 2, 3, and 4 display results of tact and listener training across full probes for all participants, and Table 4 includes a summary of the collected measures. Table 4 provides the number of targets trained to mastery in each modality, the number of untrained targets that emerged at mastery levels for each participant, and the average number of trials to criteria for each modality. For Table 4, mastery of trained and untrained targets was determined based on the full probe that immediately followed training. This was done to ensure targets from sets 1 and 2 could be compared to targets from sets 5 and 6. If delayed emergence was observed, it is reported below and is displayed on the participants' graph. Maintenance

Fig. 2 Results of tact and listener training are shown for Kevin (*top*), Josh (*middle*), and Mark (*bottom*). The graphs on the *left side* display results of tact training, and graphs on the *right side* display results of

listener training. The number of tact targets mastered is represented by *black bars*. The number of listener targets mastered is represented by *white bars*

was evaluated by comparing the number of targets at mastery criteria across full probes. For consistency across sets, maintenance was evaluated across only one interval (e.g., full probe 1 to full probe 2 for sets 1 and 2; full probe 2 to full probe 3 for sets 3 and 4). The final sets (i.e., sets 5 and 6) did not receive maintenance probes.

For Kevin, the mastery criterion was met for all three sets assigned to tact training after a range of 20–40 trials; the average

number of trials-to-criterion was 28 (Table 4). Tact training led to mastery of 8/9 targets as tacts and the emergence of 8/9 untrained listener targets (Fig. 2 and Table 4). Following one maintenance interval, Kevin emitted 100% of the trained tact targets and 80% of the untrained listener targets at mastery level. For sets assigned to listener training, mastery criteria were met for sets 2 and 6 after 30 trials each and termination criteria were met for set 4 following 16

Fig. 3 Results of tact and listener training are shown for Jake (*top*), Yale (*middle*), and Neal (*bottom*). The graphs on the *left side* display results of tact training, and graphs on the *right side* display results of listener

training. The number of tact targets mastered is represented by *black bars*. The number of listener targets mastered is represented by *white bars*

training sessions; the average trials to criteria was 47. Listener training led to mastery of only 3/9 listener targets and the emergence of 0/9 untrained tact targets. Following one maintenance interval, Kevin emitted 0% of the trained listener targets at mastery level. Thus, on the basis of targets mastered from direct training, emergent responses, maintenance, and trials-to-criteria, tact training was determined to be the more efficient instructional format for Kevin.

For Josh, the mastery criterion was met for all three sets assigned to tact training after a range of 30–55 trials; the average

number of trials-to-criteria was 40 (Table 4). Tact training led to mastery of 6/6 tact targets, and the emergence of 3/6 untrained listener targets (Fig. 2 and Table 4). Following one maintenance interval, Josh emitted 100% of the trained tact targets at mastery level. Delayed emergence was observed, as all three listener targets in set 1 were emitted at mastery levels during full probe 2, although they were not emitted at mastery levels during full probe 1. For sets assigned to listener training, the mastery criterion was met for sets 2 and 6 after a range of 30–65 trials and the

Fig. 4 Results of tact and listener training are shown for Mike (*top*) and Paul (*bottom*). The graphs on the *left side* display results of tact training, and graphs on the *right side* display results of listener training. The

number of tact targets mastered is represented by *black bars*. The number of listener targets mastered is represented by *white bars*

termination criterion was met for set 4 after 95 trials; the average number of trials to criteria was 62. Listener training led to the mastery of 3/6 trained listener targets and the emergence of 3/6 untrained tact targets. Following one maintenance interval, Josh emitted 100% of the trained listener targets and 100% of the untrained tact targets at mastery level. Thus, on the basis of targets

mastered from direct training and trials-to-criterion, tact training was determined to be the more efficient instructional format for Josh. As regards to emergence of untrained skills and maintenance, tact and listener training were equivalent.

For Mark, the mastery criterion was met for all three sets assigned to tact training after a range of 40–60 trials; the average

Table 4 Results of tact and listener training across participants

Participant	Tact training			Listener training		
	Targets mastered from direct teaching	Emergent listener targets	Average trials to mastery (per set)	Targets mastered from direct teaching	Emergent tact targets	Average trials to mastery (per set)
Kevin	8/9	8/9	28	3/9	0/9	47
Josh	6/6	3/6 ^a	40	3/6	3/6	62
Mark	8/9	9/9	50	8/9	5/9 ^a	40
Jake	9/9	9/9	50	9/9	0/9 ^a	43
Yale	5/6	5/6	90	5/6	2/6	80
Neal	9/9	9/9	18	8/9	7/9	18
Mike	9/9	9/9	52	8/9	2/9 ^a	28
Paul	6/6	5/6 ^a	23	6/6	5/6	18

^a Indicates more targets emerged upon subsequent Full Probes

number of trials-to-criterion was 50 (Table 4). Tact training led to the mastery of 8/9 tact targets and the emergence of 9/9 untrained listener targets (Fig. 2 and Table 4). Following one maintenance interval, Mark emitted 83% of the trained tact targets and 100% of the untrained listener targets at mastery level. The mastery criterion was met for all three sets assigned to listener training after a range of 30–50 trials; the average number of trials to the criterion was 40. Listener training led to the mastery of 8/9 listener targets and the emergence of 5/9 untrained tact targets. Following one maintenance interval, Mark emitted 83% of trained listener targets at mastery level. Delayed emergence was observed, as Mark emitted an additional untrained tact from set 2 at mastery level during full probe 2; however, only 2/3 untrained tact targets in set 4 maintained at mastery level. Overall, 83% of the untrained tact targets maintained following listener training. The emergent targets, both tact and listener, maintained as well or better than the directly trained targets. Thus, on the basis of the number of emergent skills, tact training was determined to be a more efficient instructional format for Mark. As regards to the number of targets mastered from direct training, maintenance, and average trials-to-criterion, tact and listener training were similar. The difference of 10 trials-to-criterion was judged to be of minimal difference by Mark's clinical case manager.

For Jake, the mastery criterion was met for all three sets assigned to tact training after a range of 40–55 trials; the average number of trials-to-criterion was 50 (Table 4). Tact training led to the mastery of 9/9 trained tact targets and the emergence of 9/9 untrained listener skills (Fig. 3 and Table 4). Following one maintenance interval, Jake emitted 83% trained tact targets and 83% untrained listener targets at mastery level. The mastery criterion was met for all three sets assigned to listener training after a range of 15–95 trials; the average number of trials to criteria was 43. Listener training led to the mastery of 9/9 trained listener skills but no untrained tact targets emerged. Following one maintenance interval, Jake displayed 83% trained listener targets at mastery level. Delayed emergence was observed, as four untrained tacts occurred at mastery level. Thus, on the basis of emergence and maintenance of untrained skills, tact training was determined to be a more efficient instructional format for Jake. As regards to trials-to-criterion and direct training, tact and listener training were similar. The difference of 7 trials-to-criterion was judged to be of minimal difference by the clinical case manager.

For Yale, the mastery criterion was met for set 1 after 60 trials and the termination criteria were met for set 3 after 120 trials; the average number of trials to criteria was 90 (Table 4). Tact training led to the mastery of 5/6 tact targets and 5/6 untrained listener targets (Fig. 3 and Table 4). While Yale's responding did not meet mastery criteria for set 3 during the training condition, when full probes were conducted, his responding met mastery criteria for some targets. Following one maintenance interval, Yale emitted 67% of trained tacts and 0% of the untrained listener targets at mastery level. The mastery criterion was met for set 2 after 40 trials and the termination criteria were met for set 4 after 120 trials;

the average number of trials to criteria was 80. Listener training led to the mastery of 5/6 listener targets and the emergence of 2/6 untrained tact skills. Following one maintenance interval, Yale emitted 67% of trained listener targets and 0% of the untrained tact targets at mastery level. Thus, on the basis of emergence, tact training was determined to be a more efficient instructional format for Yale. As regards to trials-to-criterion, direct training, and maintenance the training formats were similar. The difference of 10 trials-to-criterion was judged to be of minimal difference by the clinical case manager.

For Neal, the mastery criterion was met for all three sets assigned to tact training after a range of 15–25 trials; the average number of trials-to-criterion was 18 (Table 4). Tact training led to the mastery of 9/9 trained tact targets and the emergence of 9/9 untrained listener targets (Fig. 3 and Table 4). Following one maintenance interval, Neal emitted 100% trained tact targets and 100% of the untrained listener targets at mastery level. The mastery criterion was met for all three sets assigned to listener training after a range of 15–25 trials; the average number of trials-to-criterion was 18. Listener training led to the mastery of 8/9 trained listener targets and the emergence of 7/9 untrained tact targets. Following one maintenance interval, Neal emitted 100% of the trained listener targets and 100% of the untrained tact targets. Delayed emergence was observed, as an additional trained listener target occurred at mastery level during full probe 2. Thus, on the basis of emergence and results of direct training, tact training was determined to be a more efficient instructional format for Neal. As regards to trials-to-criterion and maintenance, the training formats were similar.

For Mike, the mastery criterion was met for all three sets assigned to tact training after a range of 35–65 trials; the average number of trials-to-criterion was 52 (Table 4). Tact training led to the mastery of 9/9 trained tact targets and the emergence of 9/9 untrained listener targets (Fig. 4 and Table 4). Following one maintenance interval, Mike emitted 0% of the trained tact targets and 100% of the untrained listener targets at mastery level. The mastery criterion was met for all three sets assigned to listener training after a range of 15–40 trials; the average number of trials-to-criterion was 28. Listener training led to the mastery of 8/9 listener targets, and the emergence of 2/9 untrained tact targets. Following one maintenance interval, Mike emitted 83% of the trained listener targets at mastery level and 100% of the untrained tact targets. Delayed emergence was observed during full probe 3, as one additional untrained tact skill occurred at mastery level from set 4. Thus, on the basis of emergence and results of direct training, tact training was determined to be a more efficient instructional format for Mike. However, as regards to trials-to-criterion and maintenance, listener training was determined to be the more efficient instructional format.

For Paul, the mastery criterion was met for both sets assigned to tact training after a range of 15–30 trials; the average number of trials-to-criterion was 23 (Table 4). Tact training led to the mastery

of 6/6 trained tact targets and the emergence of 5/6 untrained listener targets (Fig. 4 and Table 4). Following one maintenance interval, Paul emitted 100% of the trained tact targets at mastery level and 100% of the untrained listener targets. Delayed emergence was observed, as Paul correctly emitted an additional untrained listener response from set 1 during full probe 2. The mastery criterion was met for both sets assigned to listener training after a range of 15–20 trials; the average number of trials-to-criterion was 18. Listener training led to the mastery of 6/6 trained listener targets and 5/6 untrained tact targets. Following one maintenance interval, Paul emitted 100% of the trained listener targets and 100% of the untrained tact targets at mastery level. Thus, for Paul, the two training conditions were determined to be equally efficient on the basis of all measures. The difference of 5 trials-to-criterion was judged to be of minimal difference by the clinical case manager.

Discussion

The purpose of the current study was to evaluate an abbreviated, individualized assessment procedure to compare tact and listener training for children with ASD in a clinical setting. The current procedures were abbreviated from Delfs et al., in terms of the following: (1) the number of probe trials, (2) use of cold probe rather than trial-by-trial data collection, and (3) use of a multiple probe design across behaviors. These assessment procedures yielded several useful measures of individual performance (e.g., emergence of untrained responses, trials to criteria, maintenance), which were incorporated into the participants' ongoing clinical programming.

For six of the eight participants in the current study, tact training was more likely to lead to the emergence of listener skills than vice versa, replicating findings from previous research (e.g., Delfs et al., 2014; Sprinkle & Miguel, 2012; Wynn & Smith, 2003). For the remaining two participants, tact training and listener training were equally likely to lead to the emergence of additional responses. Thus, if emergence of untrained skills is the only measure of interest, tact training should be selected. However, if measures other than emergence are of interest, systematic comparison procedures similar to those included in the current study may be useful.

Results from the current study also suggest that there are several other factors that must be considered when selecting tact or listener training, aside from the emergence of untrained targets. Factors such as mastery of directly trained targets, trials-to-criterion, and maintenance should be carefully evaluated as well, and the training condition that is superior with respect to these measures may be different from that which produces the greatest degree of emergent responding. Included below are five possible outcomes that could be reached using the current study's procedures along with suggestions for how to program for these outcomes. For simplicity, the measures related to emergence of untrained skills and trials-to-criterion were selected as the primary factors for the

decision-making matrix below, though the other measures are incorporated when equivalent outcomes are observed.

Outcome 1 indicates that trials-to-criterion and emergence of untrained skills support the same form of training. Kevin's findings match this outcome. For a clinician, interpretation of these results is straightforward; training conducted in the indicated condition will be most efficient. Outcome 2 occurs if both tact and listener training are equally likely to produce the emergence of untrained responses, but there is a significant difference in trials-to-criterion between the two training conditions. Josh's results fit this outcome. In this case, it is recommended that the clinician use the training condition with the fewest trials-to-criterion. Outcome 3 occurs if trials-to-criterion are roughly equivalent, but one form of training is more likely to produce untrained responses. Mark, Jake, Yale, and Neal's findings match this outcome. In this case, a clinician should select the training condition more likely to produce emergent responses. Outcome 4 suggests that additional investigation and analysis is warranted, as the two measures indicate the use of different training conditions. In this case, as with Mike, clinicians would be best served to consider the relative strengths of each measure. For example, if trials-to-criterion were vastly lower for one training condition and that condition was only marginally less likely to produce untrained responses, that training condition would be indicated. If both findings are starkly divergent, the clinician would need to evaluate whether a longer training procedure that reliably leads to the emergence of the untrained skill is better suited to the client's needs, or whether using the faster training procedure, but needing to retrain the same targets in the opposite modality is preferred. Further, maintenance data could also aid with making this decision. Comparison of maintenance data for Mike's targets indicates that no tact responses that were directly trained maintained throughout the current evaluation. Interestingly, tact targets that emerged as a product of listener training were more likely to be maintained. However, all listener responses maintained, regardless of whether they had been directly trained or had emerged. Thus, overall listener training led to better maintenance of trained and untrained skills for Mike, potentially tipping the scales in favor of listener training. In outcome 5, both forms of training are equivalent (as was the case for Paul). As with outcome 4, other variables should be considered to determine relative strengths.

The clinician may also ask themselves the following questions. Which training procedure is easier to implement with regard to material preparation or desk space for presenting visual arrays? If an attending cue is necessary, is it easier to incorporate into one of these training conditions? Is the targeted content likely to be probed as a listener skill or tact skill when presented in the context of typical social interaction or academic testing? Can I (or my team) teach this task with the same level of fidelity in both tact and listener training? Have controlling prompts (i.e., prompts that ensure correct responding) been identified for both training approaches? Does this client have a relative preference for tact or listener training? Does the client's teacher or parents have a preference for working on skills that require vocal or motor output?

When services are provided in clinical settings, decisions related to optimal training conditions require consideration of a wide variety of factors, such as those listed here.

Of note, the need for clinicians to choose between tact and listener training is only of concern as long as these repertoires remain independent of one another. Once these repertoires become interdependent, as would be suggested by outcome 5, the choice regarding tact or listener training becomes much simpler. Once interdependence between speaker and listener repertoires (i.e., naming) is established, either training condition will simultaneously establish both tact and listener behavior (Horne & Lowe, 1996). While preliminary findings suggest this interdependence begins to occur around 30 months in typical development (Lee et al., 2015), several studies involving children with disabilities have suggested that this interdependence can be established through specific training procedures, such as multiple exemplar instruction (Greer, Stolfi, Chavez-Brown, & Rivera-Valdes, 2005; Fiorile & Greer, 2007; Greer, Stolfi, & Pistoljevic, 2007).

The aim of the current study was to take another step in aligning experimentally valid procedures with clinical practice. In order to accomplish this aim, individual differences with respect to array sizes, prompting hierarchies, and reinforcement schedules were allowed based on the judgment of each participant's clinical case manager (a BCBA). The methods selected were those typically used in tact and listener programming for each participant and balanced across tact and listener conditions. While these adjustments limit the generality of findings across participants, they enhance the likelihood that results from this comparison will effectively steer treatment decisions for each individual. How these individualized differences may have influenced overall outcomes is unclear from these data. For example, Kevin and Josh's listener targets were presented in arrays of 3 instead of 6, which is presumably an easier array to scan and seems likely to facilitate correct responding. Yet the numbers of listener target mastered from direct training were lower and average trials to mastery were higher for these participants than for those whose targets were presented in arrays of 6. Additionally, tangible reinforcement was not provided on the transfer trials for Paul during training, and his average trials to criteria were among the lowest of all participants. Future studies could utilize more standardized criteria or assessments to determine how to individualize these training variations across participants. For example, the Barriers Assessment of the VB-MAPP (Sundberg, 2008) could be used as a guide to select array sizes, number of training trials, and reinforcement schedules.

One limitation of the current study is that IOA and treatment fidelity data were not collected for all participants. However, this limitation may be mitigated by the high percentages of IOA and treatment fidelity obtained with the participants for whom these data were collected. Secondly, the sets selected for each participant were only informally counterbalanced in terms of complexity. Future research should employ more systematic procedures

already established in the literature (Gast, 2010) to overcome this limitation. Lastly, the mastery criteria required two targets within a set be emitted at mastery level, which may have underrepresented the number of targets that emerged. This decision was made in an effort to require discriminated responding for trained/untrained targets and exclude targets that only appeared to be mastered due to chance-level responding. Future studies may want to include mastery criteria measures normally used for each individual participant's specific programming.

From a clinical perspective, one limitation of the current study is the lack of programmed maintenance procedures (i.e., ongoing training and reinforcement for trained targets that met mastery criteria). As durability of responding over time was a measure used in our comparison of relative efficiency, it was determined that no additional training should be provided to targets once the training phase was completed. Another possible limitation from a clinical perspective was the use of termination criteria in the training phase. In clinical practice, practitioners would likely make modifications to training procedures as soon as failure to acquire a target is observed. In the current study, training was ceased only if cold probe data suggested that acquisition of the current targets was significantly deviating from prior sets. Similarly, if targets did maintain during full probes, in a clinical setting these targets would likely be put back into training until mastery criteria were met in full probes. These features were put into the current study to maintain the integrity of the experimental design and overall validity of the comparison procedures.

In summary, the current work directly extends the Delfs et al. (2014) study to a broader pool of participants, and simplifies the assessment procedures in accordance with the recommendations made by Delfs and Frampton (2014). Adaptations that make procedures accessible to instructors from a variety of skill backgrounds are critical for the overall dissemination of effective treatment (Corrigan, 2001), and are in line with our ethical obligation to provide effective procedures that are adapted to fit the participant's environment (Baer, Wolf, & Risley, 1968). This study represents a small step in this direction that we hope is pursued further in a variety of applied settings.

Acknowledgements Sarah E. Frampton, Language and Learning Clinic, Marcus Autism Center; Hannah C. Robinson, Language and Learning Clinic, Marcus Autism Center; Daniel E. Conine, Language and Learning Clinic, Marcus Autism Center; Caitlin H. Delfs, Language and Learning Clinic, Marcus Autism Center, and Department of Pediatrics, Emory University School of Medicine

Compliance with Ethical Standards All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Informed consent was obtained for all individuals in the study.

Conflict of Interest The authors declare that they have no conflict of interest.

Appendix

Table 5 Example data sheet used in training for sets 1 and 2 for Mike

Session #:	5	
Date:	11/29/16	
Initials:	SF	
Primary/Reliability:	Primary	
Tact Cold Probe Set 1	Correct?	Correct?
Envelope	Y	
Cutting board	N	
Iron	Y	
% Correct	66.66%	
Tact Training Set 1 (Teaching Trials)		
Envelope	X X X X X	
Cutting board	X X X X X	
Iron	X X X X X	
Listener Cold Probe Set 2	Correct?	Correct?
Outlet	N	
Mixer	N	
Oven mitt	Y	
% Correct	33.33%	
Listener Training Set 2 (Teaching Trials)		
Outlet	X X X X X	
Mixer	X X X X X	
Oven mitt	X X X X X	

Cold Probe data collected on 1 trial for each target in the set here.

Teaching trials checked off throughout session to indicate they were conducted.

References

- Baer, D. M., Wolf, M. M., & Risley, T. R. (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis, 1*, 91–97. doi:10.1901/jaba.1968.1-91.
- Corrigan, P. W. (2001). Strategies for disseminating evidence-based practices to staff who treat people with serious mental illness. *Psychiatric Services, 52*(12), 1598–1606. doi:10.1176/appi.ps.52.12.1598.
- Delfs, C. H., & Frampton, S. E. (2014). Practical implications of evaluating the efficiency of listener and tact instruction for children with autism. *Journal of Applied Behavior Analysis, 47*, 810–813. doi:10.1002/jaba.176.
- Delfs, C. H., Conine, D. E., Frampton, S. E., Shillingsburg, M. A., & Robinson, H. C. (2014). Evaluation of the efficiency of listener and tact instruction for children with autism. *Journal of Applied Behavior Analysis, 47*, 793–809. doi:10.1002/jaba.166.
- Fiorile, C. A., & Greer, D. A. (2007). The induction of naming in children with no prior tact responses as a function of multiple exemplar histories of instruction. *The Analysis of Verbal Behavior, 23*, 71–87.
- Gast, D. L. (2010). *Single subject research methodology in behavioral sciences*. New York: Routledge.
- Gilic, L., & Greer, R. D. (2011). Establishing naming in typically developing children as a function of multiple exemplar speaker and listener experiences. *The Analysis of Verbal Behavior, 27*, 157–178.
- Greer, R. D., Stolfi, L., Chavez-Brown, M., & Rivera-Valdez, 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.
- Greer, R. D., Stolfi, L., & Pistoljevic, N. (2007). Emergence of naming in preschoolers: a comparison of multiple and single exemplar instruction. *European Journal of Behavior Analysis, 8*(2), 119–131.
- 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. doi:10.1901/jeab.1996.65-185.
- Horne, P. J., Lowe, C. F., & Randle, V. R. (2004). Naming and categorization in young children: II. Listener behavior training. *Journal of the Experimental Analysis of Behavior, 81*, 267–288. doi:10.1901/jeab.2004.81-267.
- Kobari-Wright, V. V., & Miguel, C. F. (2014). The effects of listener training on the emergence of categorization and speaker behavior in children with autism. *Journal of Applied Behavior Analysis, 47*, 431–436. doi:10.1002/jaba.115.
- Lee, G. P., Miguel, C. F., Darcey, E. K., & Jennings, A. M. (2015). A further evaluation of the effects of listener training on derived categorization and speaker behavior in children with autism. *Research in Autism Spectrum Disorders, 19*, 72–81. doi:10.1016/j.rasd.2015.04.007.
- Longano, J. M., & Greer, R. D. (2014). Is the source of reinforcement for naming multiple conditioned reinforcers for observing responses?

- The Analysis of Verbal Behavior*, 31, 96–117. doi:[10.1007/s40616-014-0022-y](https://doi.org/10.1007/s40616-014-0022-y).
- Lowe, C. F., Horne, P. J., Harris, F. D., & Randle, V. R. (2002). Naming and categorization in young children: vocal tact training. *Journal of the Experimental Analysis of Behavior*, 78, 527–549. doi:[10.1901/jeab.2002.78-527](https://doi.org/10.1901/jeab.2002.78-527).
- 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 Experimental Analysis of Behavior*, 83, 47–65. doi:[10.1901/jeab.2005.31-04](https://doi.org/10.1901/jeab.2005.31-04).
- Miguel, C. F., & Kobari-Wright, V. V. (2013). The effects of tact training on the emergence of categorization and listener behavior in children with autism. *Journal of Applied Behavior Analysis*, 46, 669–673. doi:[10.1002/jaba.62](https://doi.org/10.1002/jaba.62).
- Miguel, C. F., & Petursdottir, A. I. (2009). Naming and frames of coordination. In R. A. Rehfeldt & Y. Barnes-Holmes (Eds.), *Derived relational responding: applications for learnings with autism and other developmental disabilities* (pp. 129–148). Oakland, CA: New Harbinger Publications.
- Miguel, C. F., Petursdottir, A. I., Carr, J. E., & Michael, J. (2008). The role of naming in stimulus categorization by preschool children. *Journal of Experimental Analysis of Behavior*, 89, 383–405. doi:[10.1901/jeab.2008-89-383](https://doi.org/10.1901/jeab.2008-89-383).
- Petursdottir, A. I., & Carr, J. E. (2011). A review of recommendations for sequencing receptive and expressive language instruction. *Journal of Applied Behavior Analysis*, 44, 859–876. doi:[10.1901/jaba.2011.44-859](https://doi.org/10.1901/jaba.2011.44-859).
- Skinner, B. F. (1957). *Verbal Behavior*. New York: Appleton-Century-Crofts.
- Sprinkle, E. C., & Miguel, C. F. (2012). The effects of listener and speaker training on emergent relations in children with autism. *The Analysis of Verbal Behavior*, 28, 111–117.
- Sundberg, M. L. (2008). *VB-MAPP Verbal Behavior Milestones Assessment and Placement Program: a language and social skills assessment program for children with autism or other developmental disabilities: guide*. Concord, CA: AVB Press.
- Sundberg, M. L., & Michael, J. (2001). The benefits of Skinner's analysis of verbal behavior for children with autism. *Behavior Modification*, 25, 698–724.
- Szatmari, P., Bryson, S., Boyle, M., Streiner, D., & Duku, E. (2003). Predictors of outcome among high functioning children with autism and asperger syndrome. *Journal of Child Psychology and Psychiatry*, 44(4), 520–528. doi:[10.1111/1469-7610.00141](https://doi.org/10.1111/1469-7610.00141).
- Venter, A., Lord, C., & Schopler, E. (1992). A follow-up study of high-functioning autistic children. *Journal of Child Psychology and Psychiatry*, 33, 489–597. doi:[10.1111/j.1469-7610.1992.tb0087.x](https://doi.org/10.1111/j.1469-7610.1992.tb0087.x).
- Wolery, M., Ault, M. J., & Doyle, P. M. (1992). *Teaching students with moderate to severe disabilities: use of response prompting strategies*. New York, NY: Longman.
- Wynn, J. W., & Smith, T. (2003). Generalization between receptive and expressive language in young children with autism. *Behavioral Interventions*, 18, 245–266. doi:[10.1002/bin.142](https://doi.org/10.1002/bin.142).