ORIGINAL PAPER

The Effects of Instructive Feedback and Stimulus Equivalence Procedures on Group Instructional Outcomes

Christopher A. Tullis[1] · **Sarah E. Frampton**[2] · **Caitlin H. Delfs**[3] · **Kayla Greene**[4] · **Sandra Reed**[4]

Abstract

The current study combined equivalence-based instruction and instructive feedback (IF) with two groups of children with autism spectrum disorder. For group 1, three sets of three targets were tested, and for group 2, two sets of three targets were tested. For each target stimulus, the following verbal operants were evaluated: (1) tact name, (2) tact feature, (3) name-feature intraverbal, and (4) feature-name intraverbal. Following individual across set probes (ASP) sessions evaluating all targets in all sets, group treatment was conducted with set 1. During group treatment, tact training was provided to each participant relating to his/her assigned target and IF relating to a feature of the target was provided. Once daily probe data indicated all targets were mastered by the assigned participant, ASP were repeated. This process was repeated across all evaluated sets in the form of a multiple probe design. Results indicated robust responding for three participants, with evidence of responding to observational targets, secondary targets (related to the IF), and emergence of intraverbal responses. A fourth participant showed moderate levels of correct responding, particularly with observational targets. Overall, results suggest that group IF and stimulus equivalence are viable interventions for some children with ASD.

Keywords Emergence · Equivalence · Instructive feedback · Intraverbals · Observational learning · Verbal behavior

✉ Christopher A. Tullis
   ctullis2@gsu.edu

[1] Department of Learning Sciences, Georgia State University, P.O. Box 3965, Atlanta, GA 30302, USA

[2] May Institute, Inc., Randolph, MA, USA

[3] Village Autism Center, Marietta, GA, USA

[4] Marcus Autism Center, Atlanta, GA, USA

Introduction

Autism spectrum disorder (ASD) is a neurodevelopmental disorder that greatly impacts social, communication, and educational outcomes (American Psychiatric Association 2013). One-to-one instruction (e.g., discrete trial instruction) is one of the most widespread and heavily researched interventions for learners with ASD (Smith 2001; Stahmer et al. 2005). Although effective, intensive one-to-one teaching methodologies may be costly and difficult to implement in many non-clinical educational settings (Collins 2012; Smith 2001). Additionally, other instructional delivery methods (e.g., group instruction) presumably require learners to engage in observational learning, which may be difficult for some learners with ASD because of deficits related to attending to less salient environmental stimuli (Plavnick and Hume 2014).

Small group instruction has been defined as one instructor presenting instruction to a whole group or to individuals in a whole group that is composed of two to ten learners (Collins 2012). Within groups, specific formats may include teaching the same skills with similar or different stimuli, different skills, imbedded one-to-one instruction, and the implementation of various prompting procedures (Collins et al. 1991). For example, Cihak et al. (2006) implemented a modified model-lead-test prompting procedure to teach banking skills to six students with moderate to severe intellectual and developmental disabilities (IDD). Across all participants, the same skills were taught using the same stimuli and prompting procedures, illustrating one method of formatting group instruction.

Group instructional arrangements present a number of advantages over one-to-one teaching arrangements. First, group instruction allows for opportunities for observational learning to occur (e.g., Ledford et al. 2008). Second, group instructional arrangements, when used systematically, may result in more efficient instruction with less instructional staff found in inclusive educational settings (Collins et al. 1991). Previous work has supported the efficacy of this type of instruction on long-term educational outcomes for people with more severe disabilities (e.g., Jimenez et al. 2012). Despite the advantages and associated benefits of instructional arrangements present in inclusive or general educational settings (i.e., group instruction), many learners with more severe disabilities are placed in segregated educational environments (Kleinert et al. 2015; Morningstar and Kurth 2017; Morningstar et al. 2017).

Two methodologies that may be appropriate for small group instruction are instructive feedback (IF; Werts et al. 1995) and equivalence-based instruction (EBI; Stanley et al. 2018; Sidman 1994). IF is a teaching methodology where secondary targets are presented during the learning trial for primary targets, and the learner is not required to respond to the secondary target (Werts et al. 1995). For example, if an instructor is teaching tacts to a learner (e.g., tiger), the secondary targets that are presented after the tact occurs may be a feature of that stimulus (e.g., tigers have claws). Although typically used in one-to-one teaching arrangements, recent investigations using IF methodology embedded in group contexts have yielded promising results for learners with ASD. For example, Leaf et al.

(2017) implemented IF procedures within discrete trial teaching (DTT; Smith 2001) with nine children diagnosed with ASD in a group instructional setting composed of three learners. Across all participants, all primary, secondary, and a portion of observational targets (i.e., those delivered to others in the group) were acquired. These data are similar to the findings of previous research focusing on implementing IF procedures in a group instructional context (e.g., Ledford and Wolery 2015).

EBI (e.g., Stanley et al. 2018) is an instructional paradigm that involves directly teaching specific targets, such that others emerge without direct teaching. For example, an instructor may directly teach a learner to select a picture of an apple when presented with printed text, and to select a two-dimensional apple when presented with the same text. During subsequent testing, the instructor may determine whether the learner can match the stimulus to itself (reflexivity), reverse the previously taught relations (symmetry), and then match the relation between the dissimilar stimuli that were involved in the directly taught relations (transitivity). EBI procedures have been demonstrated with people diagnosed with ASD, but few investigations implemented procedures in a group instructional context (McLay et al. 2013). In one exception, MacDonald et al. (1986) investigated the acquisition of equivalence classes in a dyad instructional context with four adults diagnosed with severe intellectual disability. Of the four participants, only one demonstrated the emergence of a full equivalence class, indicating that observational learning did not occur for the other three.

Rehfeldt et al. (2003) extended the work of MacDonald et al. (1986) by implementing similar procedures in dyads with people with intellectual and developmental disabilities, and typically developing models. In Experiment 1, adult participants were directly taught two relations (A–B and B–C) to mastery. After three teaching trials, participants were prompted to observe a typically developing model engaging in a randomly identified relation from a separate set of stimuli. All participants demonstrated the three classes that were directly taught, but none demonstrated the full emergence of stimulus classes via observational learning (i.e., the skill demonstrated by the typical peer). In Experiment 2, the same procedures were implemented with children diagnosed with ASD, and a sibling. Similar to Experiment 1, only one participant demonstrated the emergence of a full equivalence class, and the remaining two demonstrated partial emergence.

EBI and IF procedures have been demonstrated to be effective as individual interventions in less intensive teaching formats, but two main limitations warrant further investigation. First, Leaf et al. (2017) demonstrated one method of teaching learners with ASD using IF in a group format, but the manner in which target presentation was structured limited acquisition to only one relation (e.g., tacting a feature of Magneto). Though the name and feature of each character were trained, tests for the emergence of intraverbals were not included (e.g., "What is Magneto's superpower?" or "Who controls metal?"). This may be particularly relevant for learners with ASD, as intraverbals have been reportedly deficient in comparison with other verbal operants (Sundberg and Sundberg 2011). Second, the studies evaluating the emergence of equivalence classes in group instructional arrangements has been limited to only dyad instruction. Although this arrangement may be advantageous for maintaining

attending to observational stimuli, it is often untenable in many educational settings (e.g., resource classrooms, general education) because of staffing availability or cost (Stahmer et al. 2005). For group teaching methods to be maximally effective, stimuli should be structured such that multiple relations are presented directly, and observationally (LeBlanc and Ruggles 1982). In the current study, EBI and IF procedures were combined in a group instructional context to determine the effects on skill acquisition of directly taught, secondary, and observational targets.

Method

Participants and Setting

A total of six children, two groups of three, participated in the current study. Participants were receiving services in an applied behavior analysis (ABA) based program at a language clinic, 3–5 days per week, for 3 h per day. As all participants were preparing to discharge from the services at the language clinic, the primary case manager (the second author) determined that assessment of each participant's learning in a group instruction format was warranted. All children were reported by their caregivers to have a diagnosis of ASD and were receiving services in a clinic specifically for children with ASD. Review of medical records confirmed these reports for Dee and Charlie. No additional diagnostic information was available for the other participants. Demographic information for all participants is displayed in Table 1. The Verbal Behavior Milestones Assessment and Placement Program (VB-MAPP; Sundberg 2008) was conducted with each participant at the time of their admission into the language clinic and updated as part of their ongoing services. Scores from the VB-MAPP Milestones indicate that the students scored in Level 3 (Group 1: 140–163; Group 2: 124–138).

Sessions were conducted in the participants' classrooms in the language clinic with the staff members on each participant's clinical team. Sessions were conducted 1–3 times per week, when all participants were present. During treatment sessions, all three participants were seated at a large U-shaped table and the instructor sat in the center of the table. During probe sessions, one participant and an instructor were seated at a small table in the same classroom.

Table 1 Participant characteristics and groups

Group	Participant	Age (years)	Ethnicity
1	Ronald	8	African American
	Frank	7	Asian
	Charlie	10	Hispanic
2	Maureen	3	African American
	Dee	4.5	African American
	Dennis	3.5	Asian

Materials and Target Selection

Materials for sessions consisted of pictures of either historical figures (Group 1) or cartoon characters (Group 2). The pictures were approximately 12 cm by 12 cm and printed onto a white background. For Group 1, three sets of three historical figure targets were selected. For Group 2, two sets of three cartoon character targets were selected. Only two sets were selected for Group 2 to ensure that all sets would be evaluated prior to their discharge from clinical services. For each set, one target was assigned to a participant as their primary target (e.g., Benjamin Franklin for Ronald, William Shakespeare for Charlie, etc.). For each target, an additional feature of the person/character was identified that would be considered the secondary target. For Group 1, the secondary target related to the achievement of each historical figure (e.g., for Benjamin Franklin, *discovered electricity* was his designated achievement). For Group 2, the secondary target related to a friend of each cartoon character (e.g., for Mowgli, *Baloo* was his designated friend). These targets were not presented in other contexts (i.e., by other service providers) and were reviewed with the caregivers to promote acceptability of the procedures. Primary and secondary targets are presented in Table 2.

Response Measurement

Responses were defined according to Skinner's (1957) classification of verbal operants and given alphabetical designations as commonly used in EBI research. These alphabetical designations were modeled from the stimulus class arrangement used

Table 2 Primary and secondary targets per participant

Set	Participant assignment	Primary target (Tact; B–D)	Secondary target (Tact feature; B–E)
1	Ronald	Ben Franklin	Discovered electricity
	Charlie	William Shakespeare	Wrote plays
	Frank	Leonardo DaVinci	Painted the Mona Lisa
2	Ronald	Susan B. Anthony	Helped women vote
	Charlie	Florence Nightengale	Pioneered modern nursing
	Frank	Noah Webster	Wrote the dictionary
3	Ronald	Clara Barton	Founded the Red Cross
	Charlie	Harriet Tubman	Rescued slaves
	Frank	Bessie Coleman	Flew planes
1	Maureen	Evil Queen	Magic Mirror
	Dee	Pocahontas	Flit
	Dennis	Mowgli	Baloo
2	Maureen	Mulan	Mushu
	Dee	Pinoccio	Jiminy Cricket
	Dennis	Stitch	Lilo

in Shillingsburg et al. (2018) with differentiated verbal stimuli and verbal responses. As the current study is a synthesis of procedures from the EBI, verbal behavior, and IF literature, we attempted to build a link between the approaches as shown in Figs. 1 and 2.

The primary dependent variable was the percentage of correct tact responses (B–D), which consisted of stating the name of the historical figure/character (D) after being presented with a corresponding picture (B) and the question, "Who is it?" Note, for each participant, only one tact per set was considered their primary target; the other tact targets were primary for their peers (see Table 2 for target assignments). The secondary dependent variable was the percentage of correct tact feature responses (B–E), which consisted of stating a feature of the historical figure/character (E) after being presented with a corresponding picture (B) and the question, "What did he/she do?" (Group 1) or "Who is her/his friend?" (Group 2). Note, for each participant, only one tact feature per set was considered their secondary target; the other tact feature targets were secondary for their peers.

Data were collected on the percentage of correct name-feature intraverbal responses (A–E), which consisted of stating the achievement or friend (E) after being asked a question with the historical figure/character's name (A). For example, the instructor asked, "What did Benjamin Franklin do?" and the participant said, "invent electricity" for Group 1. For Group 2, the instructor asked, "Who is Mulan's friend?" and the participant said "Mushu." Data were also collected on the
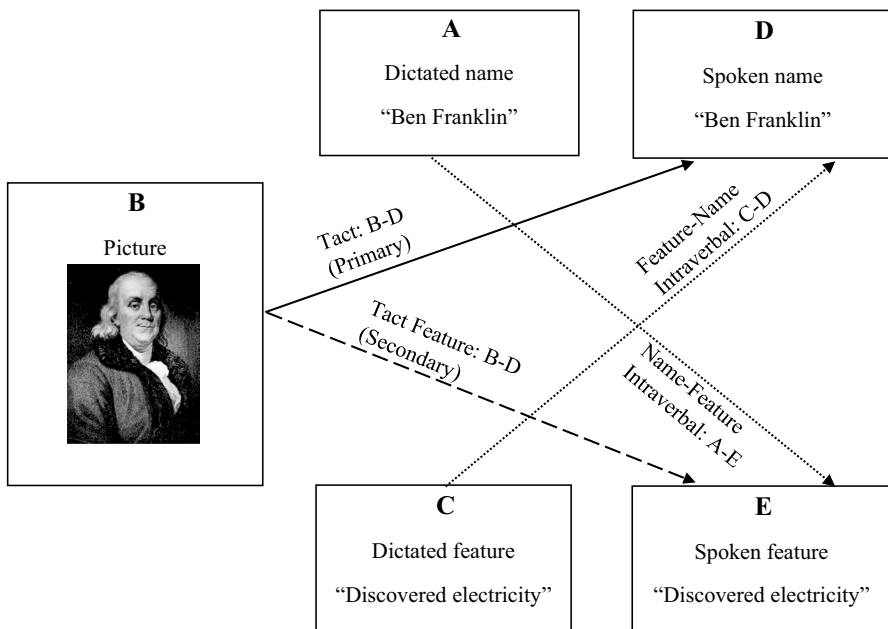

Fig. 1 Example map of relations evaluated for stimulus classes including the participant's primary and secondary targets (Ronald, set 1). The black line indicates a relation that was directly trained in group intervention sessions (B–D). The broken black line indicates a relation presented as instructive feedback in the group sessions (B–E). The dotted line indicates relations that were not trained (A–E, C–D)
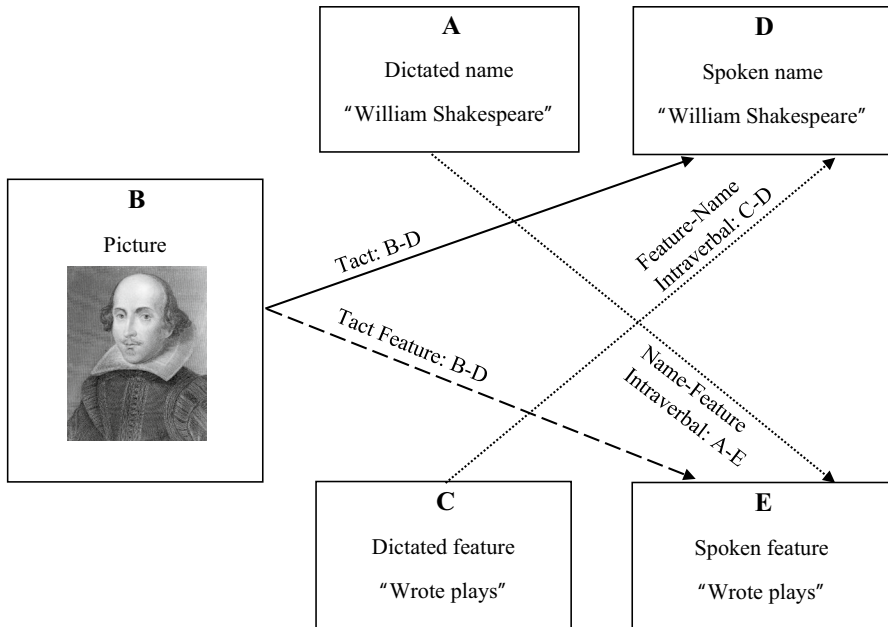
Fig. 2 Example map of relations evaluated for stimulus classes for peer targets (Ronald, set 1). The black line indicates a relation that was directly trained in group intervention sessions to a peer (B–D). The broken black line indicates a relation presented as instructive feedback in the group sessions to a peer (B–E). The dotted line indicates relations that were not trained (A–E, C–D)

percentage of feature-name intraverbal responses (C–D), which consisted of stating the name of the historical figure/character (D) after being asked a question about their designated feature (C). For example, the instructor asked "Who invented electricity?" and the participant said, "Benjamin Franklin") for Group 1. For Group 2, the instructor asked "Who is Mushu's friend?" and the participant said "Mulan." Participants were evaluated on all intraverbals, regardless of whether they corresponded to their primary or secondary targets.

During probe sessions, all responses were considered correct if they occurred within 5 s of the instructor's vocal question. During training sessions, responses were considered correct if they occurred within 3 s of the instructor's vocal question. The latency was 3 s to promote fluent responding during training.

Interobserver Agreement Data, and Treatment Integrity Data

Reliability data were recorded by trained observers and trial-by-trial interobserver agreement (IOA) was calculated separately for each participant. For group 1, reliability data were collected during 34%, 39%, and 48% of sessions for Ronald, Frank, and Charlie, respectively. Mean agreement was 99% (range 83–100%), 100%, and 99% (range 83–100%) for Ronald, Frank, and Charlie, respectively. For group 2,

data were collected during teaching for 78%, 65%, and 74% of sessions for Maureen, Dee, and Dennis, respectively. Mean agreement was 100% for all.

Treatment integrity data were recorded by trained observers during treatment sessions. As all participants received treatment simultaneously, scores were not separated according to each participant for group 1. In group 2, Maureen had 3 additional sessions of 1:1 instruction so her scores are reported separately from Dee and Dennis. A checklist detailing each step of the treatment procedures was created and observers scored a "+" if the step was followed with fidelity on each trial. If a step was not followed with integrity, a "−" was scored. For group 1, data were collected during teaching for 31% of sessions with 99% fidelity (range 99–100%). For group 2, data were collected during teaching for 30% of sessions with 100% fidelity for Dee and Dennis. For Maureen, data were collected during teaching for 23% of sessions with 100% fidelity.

Experimental Design

A multiple probe design across stimulus sets (Horner and Baer 1978) was used to assess the effects of the group treatment sessions on the occurrence of all evaluated responses (e.g., tact: B–D, tact feature: B–E, name-feature intraverbal: A–E, and feature-name intraverbal: C–D).

Across Sets Probes

All responses for all sets were evaluated during Across Sets Probe (ASP) sessions, resembling procedures employed by Shillingsburg et al. (2018). ASP sessions were conducted individually with each participant, so responding would not be influenced by peers. ASP were broken up into four blocks of nine trials in which each operant for each set could be specifically evaluated. The trial blocks were always conducted in the same order: tact name (B–D) for all sets, tact feature (B–E) for all sets, feature-name intraverbal (C–D) for all sets, and name-feature intraverbals (A–E) for all sets. Within each 9-trial block, three trials corresponded to each of the three targets within the set. For example, tact (B–D) set 1-trial block consisted of three trials corresponding to Benjamin Franklin, three trials corresponding to William Shakespeare, and three trials corresponding to Leonardo DaVinci. The order of the trials was pre-randomized for each session to avoid patterning. The total number of trials for an ASP session was 108 for Group 1 [e.g., 27 tact (B–D) trials, 27 tact feature (B–E) trials, 27 name-feature intraverbals (A–E), and 27 feature-name intraverbals (C–D)] and 72 for Group 2 [e.g., 18 tact (B–D) trials, 18 tact feature (B–E) trials, 18 name-feature intraverbals (A–E), and 18 feature-name intraverbals (C–D)].

During ASP probe sessions, the instructor presented one to two mastered demands to promote attending. Mastered demands varied across participants (e.g., motor imitation, one-step instructions, tacts, listener skills); however, they were not related to the targets in the evaluated sets. Following correct responses for mastered demands, praise was provided, and then a probe trial was presented. Neutral statements followed all responses to probe trials (e.g., "Ok"); differential consequences

were not provided. Following between three and nine probe trials, a mastered demand was presented and a correct response to the mastered demand was reinforced with praise and a point on each participant's points board. This procedure was used to prevent extinction conditions within the session. The same point board was used throughout the rest of the participants' clinical services; mastery of skills in clinical services would suggest that the points functioned as reinforcement.

Treatment

Treatment sessions were conducted when all three participants in the group were present. Prior to a treatment trial, the instructor required an observing response (Grow and Leblanc 2013) in the presence of the target stimulus by each participant. The instructor presented the picture to each participant and issued a command that required him or her to engage with the picture in some form. The commands varied per trial but included responses such as waving at the picture, pointing to the picture, tapping the picture, or blowing a kiss to the picture. Following the observing response the instructor presented a tact trial (B–D) to the participant assigned to that target. The instructor held the picture in front of the designated participant and asked, "Who is it?" If a correct response occurred, praise, a point, and the corresponding IF was provided. For example, once the participant tacted, "Benjamin Franklin," the instructor said, "You're right! And, he discovered electricity." If an incorrect response occurred, the instructor began an error correction sequence. The instructor repeated the question and provided an immediate echoic prompt (e.g., "Say, Benjamin Franklin."). Once the participant responded to the prompt, the instructor repeated the question again and provided another immediate echoic prompt. Once the participant responded to this second prompt, the instructor provided the question again but did not provide a prompt (i.e., an independent opportunity). If the participant emitted a correct response, praise, a point, and IF were provided. If an incorrect response occurred, the question was presented a final time and an immediate echoic prompt was provided. If this occurred, no IF was provided. For group 2, trials during the first intervention session for each set did not include an independent opportunity at the beginning of the teaching sequence (i.e., errorless teaching). This change was made for group 2 to bring these procedures in closer alignment with the typical errorless teaching procedures used in their clinical programming. For all additional sessions, the procedures were as described above.

Each treatment session consisted of 15 tact (B–D) trials, five target trials directed to each participant. Participants were only presented with their primary targets and were not required to respond to peer targets (see variation for Ronald below). Mastered demands were presented to the whole group following three to five target trials to promote attending and to increase the resemblance to typical group-learning sessions. Correct responses to group instructions were intermittently reinforced with points on their point boards, which was an established class-wide strategy. Once all participants earned enough points to fill their boards (10 for group 1 and 5 for group 2), the participants were asked what they wanted and allowed to mand for preferred item or activity. No

instruction took place while any participant had access to reinforcement. The session would resume when all participants had finished engaging with selected items or activities.

Note that, data were not collected on the percentage of correct responses during treatment sessions to allow the instructor to present targets more rapidly, increasing the resemblance to a typical group instruction session. The instructor checked a box after each trial sequence, similar to the procedures used in Frampton et al. (2017). This procedure ensured that the number of trial presentations was identical across targets and participants, and thus, any differential results by participants would not be due to uneven trial distribution in treatment sessions.

To determine mastery of the targets presented in treatment sessions, daily probe sessions were conducted following treatment sessions. At minimum, the probes were conducted 30 min after the treatment session, though they were typically conducted on the following day. Daily probe session was always conducted individually so responding could not be influenced by peers. During daily probe sessions, only the tact (B–D) and tact feature (B–E) targets within the current set were presented. Each target was presented once in randomized order. As in ASP, neutral responses followed all responses and mastered demands were interspersed and reinforced. Mastery for each participant was determined based on responding to their primary tact (B–D) target only; mastery criteria were correct responses across three consecutive daily probe sessions. Group mastery criteria were met when all participants met the individual mastery criteria.

Observation Trial

For Ronald, a change was made to teaching procedures for sets 2 and 3 to increase his observation of other participants' responses. Following a trial presented to Frank, the instructor would immediately ask Ronald, "What did Frank just say?" If Ronald emitted the correct response (i.e., repeated what Frank said during his tact (B–D) trial), Ronald earned an additional point on his board. If he did not emit a correct response, no point was earned. No additional prompts or feedback were provided. This procedure was only in place for trials presented to Frank, no changes were made for trials presented to Charlie to determine if the modification led to generalized improvements for Ronald's responding across all targets or if the effects were isolated to only Frank's targets designated targets.

Results

For all participants, results of ASP sessions are displayed in Figs. 3 and 4. To clearly illustrate response patterns, the ASP are broken down into the four, nine-trial blocks in which each operant was evaluated specifically (e.g., B–D, B–E, A–E, and C–D). In these figures, responses are not distinguished based on method of instruction (i.e., whether the tact was their primary target or observed). In Tables 3 (Group 1) and 4 (Group 2), responses are differentiated into more precise categories: (1) primary target (B–D), (2) observed target (B–D), (3) secondary target (B–E), (4) observed
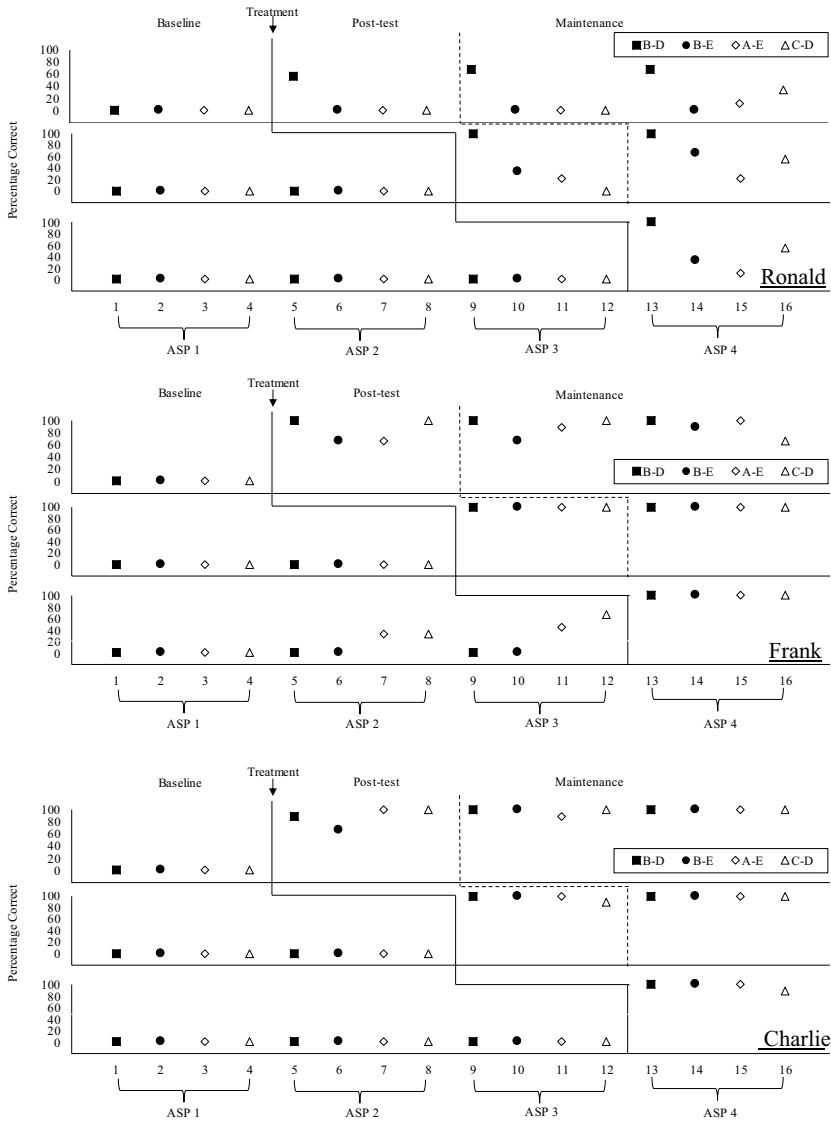
Fig. 3 Performance during Across Set Probes (ASP) is shown for Group 1 across trial blocks. Tact (B–D) responses are displayed with black squares. Tact feature (B–E) responses are displayed with black circles. Name-feature intraverbal (A–E) responses are displayed with white diamonds. Feature-Name intraverbal responses (C–D) are displayed with white triangles

secondary target (B–E), (5) intraverbal responses related to primary targets (A–E and C–D), and (6) intraverbal responses related to observed targets (A–E and C–D). Results of daily probes are not displayed but are available upon request.
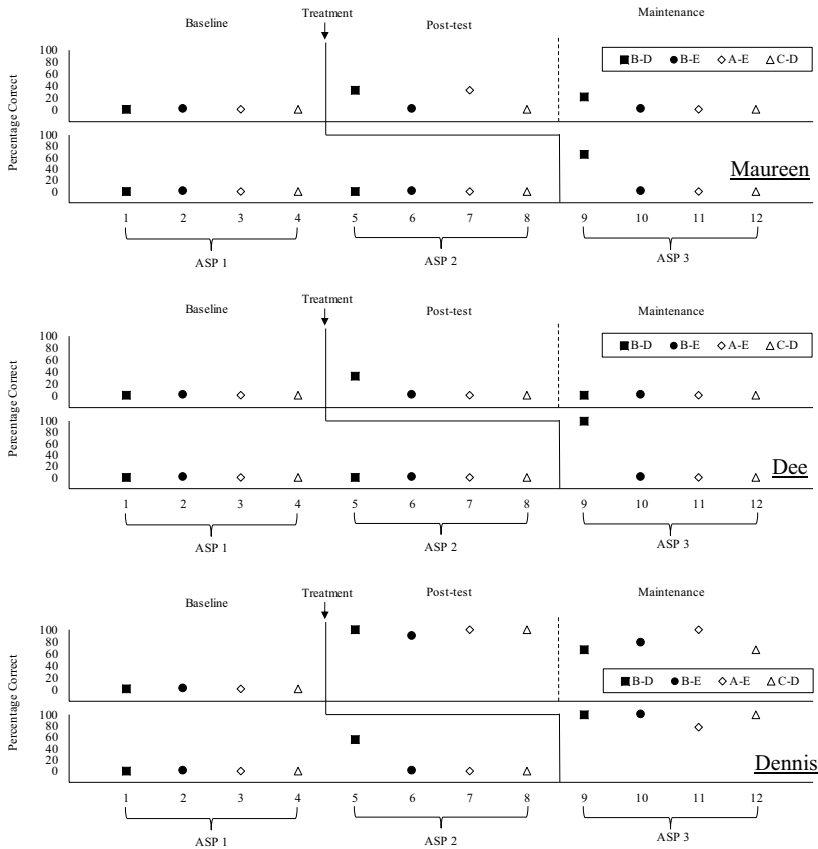
Fig. 4 Performance during Across Set Probes (ASP) is shown for Group 2 across trial blocks. Tact (B–D) responses are displayed with black squares. Tact feature (B–E) responses are displayed with black circles. Name-feature intraverbal (A–E) responses are displayed with white diamonds. Feature-name intraverbal responses (C–D) are displayed with white triangles

Group 1

Ronald emitted no correct responses in ASP 1 (Baseline; Fig. 3). Ronald's B–D responding for his primary target reached mastery criteria for set 1 after four sessions. Results of ASP 2 showed that for set 1 Ronald emitted correct responses on 55% of B–D trials, though no correct responses were observed for any other trial type. Responding for all targets in sets 2 and 3 remained at 0%. Ronald's B–D responding for his primary target reached mastery criteria for set 2 after four treatment sessions, though a total of seven sessions were conducted to allow Charlie's responding to meet mastery criteria. Observation trials were added during treatment sessions for set 2. Results of ASP 3 showed that for set 2 Ronald emitted correct responses for 100% of B–D targets, 33% of B–E targets, 22% of A–E responses, and 0% of C–D responses. Ronald's B–D responding reached mastery criteria for

set 3 after three treatment sessions, which included the observation trials. Results of ASP 4 showed that for set 3 Ronald emitted correct responses for 100% of B–D targets, 33% of B–E targets, 11% of A–E, and 55% of C–D responses. For Ronald, responding improved over subsequent maintenance probes for sets 1 and 2 with delayed emergence of A–E and C–D responses. Generally, correct responses for Ronald were distributed across both targets assigned to him and those assigned to both peers (Table 3).

Frank emitted no correct responses in ASP 1 (Baseline; Fig. 3). Frank's B–D responding for his primary target in set 1 met mastery criteria after three sessions, though a total of four sessions were conducted to allow Charlie's and Ronald's responding to reach mastery criteria. Results of ASP 2 showed that for set 1 Frank emitted moderate to high levels of correct responses for all trial types (e.g., 00% B–D, 66% B–E, 66% A–E, and 100% C–D). No correct responses were observed for set 2 though low levels of correct responses were observed for A–E and C–D responses in set 3. Frank's B–D responding met mastery criteria for set 2 after three treatment sessions, though a total of seven sessions were conducted to allow Charlie's responding to meet mastery criteria. Results of ASP 3 showed that for set 2 Frank emitted correct responses on all trials for all trial types. Moderate levels of A–E and C–D responses continued for set 3, likely due to practice effects and exclusion-based learning. Frank's B–D responding met mastery criteria for set 3 after three treatment sessions. Results of ASP 4 showed that Frank emitted correct responses on all trials for all trial types. Responses maintained at high levels for sets 1 and 2. Correct responses in set 1 were more consistent with peer targets than those assigned to him, though by sets 2 and 3 correct responses were evenly distributed (Table 3).

Charlie emitted no correct responses in ASP 1 (Baseline; Fig. 3). Charlie's B–D responding for his primary target reached mastery criteria for set 1 after 4 treatment sessions. Results of ASP 2 showed moderate to high levels of correct responses across all trial types for set 1; no correct responding was observed for sets 2 and 3. Charlie's B–D responding reached mastery criteria for set 2 after seven treatment sessions. Results of ASP 3 showed that for set 2 Charlie emitted high levels of correct responses across all trial types. No correct responses were observed for set 3. Charlie's B–D responding reached mastery criteria for set 3 after three treatment sessions. Results of ASP 4 showed that Charlie emitted high levels of correct responses across all trial types for set 3. Responses maintained at high levels for sets 1 and 2. Correct responses were evenly distributed across targets assigned to him and his peers (Table 3).

Group 2

Maureen emitted no correct responses in ASP 1 (Baseline; Fig. 4). Maureen's B–D responding reached mastery criteria for set 1 after four treatment sessions. Results of ASP 2 showed that Maureen emitted low levels of correct B–D and A–E response for set 1. When analyzed further, the correct B–D responses occurred exclusively on trials for her primary target and the correct A–E responses related only to peer

Table 3 Distribution of correct responses according to trial type, Group 1

Participant	Set	Primary target (Tact; B–D)	Peer targets (Tact; B–D)	Secondary target (Tact feature; B–E)	Peer targets (Tact feature; B–E)	Related Responses (Intraverbal; A–E and C–D)	Peer related responses (Intraverbal; AE and CD)
Ronald	1	3/3	2/6	0/3	0/6	0/6	0/12
	2	3/3	6/6	1/3	2/6	1/6	1/12
	3	3/3	6/6	0/3	3/6	2/6	4/12
Frank	1	3/3	6/6	0/3	6/6	3/6	12/12
	2	3/3	6/6	3/3	6/6	6/6	12/12
	3	3/3	6/6	3/3	6/6	6/6	12/12
Charlie	1	3/3	5/6	3/3	3/6	6/6	12/12
	2	3/3	6/6	3/3	6/6	5/6	12/12
	3	3/3	6/6	3/3	6/6	5/6	12/12

Performance evaluated in posttest (i.e., ASP immediately following intervention)

Table 4 Distribution of correct responses according to trial type, Group 2

Participant	Set	Primary target (Tact; B–D)	Peer targets (Tact; B–D)	Secondary target (Tact feature; B–E)	Peer targets (Tact feature; B–E)	Related responses (Intraverbal; A–E and C–D)	Peer related responses (Intraverbal; AE and CD)
Maureen	1	3/3	0/6	0/3	0/6	0/6	3/12
	2	3/3	3/6	0/3	0/6	0/6	0/12
Dee	1	3/3	0/6	0/3	0/6	0/6	0/12
	2	3/3	6/6	0/3	0/6	0/6	0/12
Dennis	1	3/3	6/6	3/3	5/6	6/6	12/12
	2	3/3	6/6	3/3	6/6	6/6	10/12

Performance evaluated in posttest (i.e., ASP immediately following intervention)

targets (Table 4). Responding for set 2 remained at 0%. Maureen's B–D responding reached mastery criteria for set 2 after six treatment sessions. Results of ASP 3 showed that Maureen emitted correct responses on 66% of B–D trials for set 2. Responses were evenly distributed between her primary target and those of her peers. Set 1 A–E responses did not maintain, though some maintenance of set 1 B–D responses was observed (22.2%).

Dee emitted no correct responses in ASP 1 (Baseline; Fig. 4). During B–D treatment for set 1, Dee did not emit any correct responses by the time both her peers' B–D responding reached mastery criteria. Rather than continuing with group intervention sessions, one-on-one intervention sessions were conducted with Dee (presenting only her primary and secondary targets). Mastery criteria for B–D were met after three sessions of one-on-one treatment. Results of ASP 2 showed that Dee emitted low levels of correct responses on B–D trials (33.3%), all responses in the presence of her primary target only (Table 4). Responding for set 2 remained at 0%. Dee's B–D responding reached mastery criteria for set 2 after four treatment sessions (now back in the group). Results of ASP 3 showed that Dee emitted correct responses on 100% of tact trials. B–D responses demonstrated in set 1 did not maintain from ASP 2 to ASP 3.

Dennis emitted no correct responses in ASP 1 (Baseline; Fig. 4). Dennis' B–D responding reached mastery criteria for set 1 after four treatment sessions. Results of ASP 2 showed high levels of correct responding across all trial types for set 1. No correct responses occurred for set 2. Dennis' B–D responding met mastery criteria for set 2 after three treatment sessions. Results of ASP 3 showed high levels of correct responding across all trial types. Set 1 responses maintained at moderate levels from ASP 2 to ASP 3. Correct responses were evenly distributed across targets assigned to him and to his peers (Table 4).

Discussion

In the current study, the effect of a teaching package that included IF and EBI was evaluated with six learners with ASD. During baseline, all participants engaged in low levels of responding to instructional stimuli. After teaching was introduced, three (Charlie, Frank, and Dennis) of the six participants demonstrated high levels of correct responding regardless of instructional delivery. Though Ronald demonstrated lower levels of correct responses, the response pattern that he exhibited may be indicative of early acquisition of targeted stimuli. During teaching for set 1, he engaged in low levels of responding to all stimuli that were not directly taught, but when the attentional cue was introduced (i.e., "what did ___ say?"), responding increased. This effect was replicated with set 3 and, interestingly, improvements were subsequently seen with set 1 targets though no programmed maintenance was conducted. This pattern of responding may indicate that a different teaching procedure (e.g., multiple exemplar instruction) may be necessary for the occurrence of secondary and emergent targets.

The poor responding observed for Dee and Maureen may be the result of their instructional histories, and a potential lack of contact with reinforcement for

learning targets that are not directly taught. Previous commentaries have hypothesized that one mechanism underlying IF procedures is observational or incidental learning (Werts et al. 1995; Nottingham et al. 2015), which require attention and responding to stimuli that is not always directly reinforced. From the current data, it seems that Dee and Maureen did not have the prerequisite skills to learn either observationally or incidentally in a group-learning context. The extent to which they may be taught to do so was limited by their discharge from the clinic. Given the success of an attentional cue modification for Ronald, this may be one avenue that can be explored by future investigations. Dee and Maureeen may have benefited from explicit teaching of the observational and secondary targets similar to the teaching implemented in Tullis et al. (2017). In the Tullis et al. investigation, learners that did not acquire secondary targets were directly taught one target set to mastery prior to IF procedures being implemented in subsequent sets. The more targeted teaching procedure resulted in acquisition of secondary targets across all participants for which IF procedures alone were initially ineffective.

In comparison with the previous literature related to IF and EBI, the current study provides three meaningful extensions. First, the current investigation consisted of groups of three learners instead of dyads that have been found in previous group EBI. For example, Rehfeldt et al. (2003) and MacDonald et al. (1986) implemented procedures in dyads, which may have resulted in increased stimulus salience. Although dyad instruction is still considered a group instructional arrangement (Collins et al. 1991), this type of setting may not result in the same or similar results when compared to instructional arrangements with more learners (e.g., groups of three or more). In dyad instruction, learners may attend to stimuli more readily because only one other learner was included in the experimental arrangement, and proximity to teaching stimuli was lessened. Additionally, dyad instruction may allow for the instructor to hold the learner's attention longer because it is easier to keep a high rate of reinforcement. The addition of a third learner may reflect a more naturalistic form of instruction encountered in educational contexts where staffing typically does not allow for dyad instruction as a group format.

Second, in the group instruction literature, several recommendations have been made related to how to structure the teaching of stimuli (Collins et al. 1991; Collins 2012). Although these recommendations have been present in the literature (e.g., using the same materials and teaching different skills), few studies have precisely analyzed a method of delivering instructional content with the goal of demonstrating the occurrence of secondary, emergent, and observed skills. Several studies have demonstrated the occurrence of secondary targets after primary target instruction in a group context (Leaf et al. 2017; Ledford et al. 2008, 2012), but these investigations may be limited in that they did not document the occurrence of either observed, or emergent targets, either singly or in combination. Additionally, little data are available on the effects of structuring stimuli and responses in a group context according to guidelines suggested by Collins et al. (1991) and others. In the current investigation, instructional stimuli were topographically dissimilar, as were the responses required by each learner. This explicit programming presents one empirically validated method of presenting instructional stimuli that is in line with previous recommendations (e.g., Collins et al. 1991).

Last, the current investigation contains an instructional arrangement consisting of both IF and EBI procedures. Previous work has supported the effectiveness of these interventions in group arrangements (e.g., Leaf et al. 2017), but to our knowledge, none have combined both procedures. In the current investigation, the combination of IF and EBI procedures resulted in acquisition of a large number of targets for four participants, which may indicate that this type of structured arrangement may be one way of increasing instructional efficiency in settings where educational or training resources are limited (Stahmer et al. 2005).

Although the current investigation was effective for four of the six participants, further replication may be necessary to refine the current procedures. First, from the data recorded, instructional efficiency can be at least partially concluded. Participants were taught a small number of targets directly, and a number occurred without explicit teaching. The overall number of targets learned without direct teaching is encouraging, but that measure does not yield a precise measure of instructional efficiency. Yaw et al. (2014), and Black et al. (2016) suggest calculating time per target to mastery as a more precise method of depicting instructional efficiency. This type of measure takes into account the both the number of targets mastered, and the amount of time that was needed to reach mastery. Future replication work should include these more precise measures in an effort to depict not only raw numbers of acquired skills that were not directly taught, but also the instructional time that was required for those skills to be observed.

Second, the current sample of learners presented with similar skill repertoires, but only a portion acquired directly taught, secondary, observational, and emergent responses. For responders in particular (i.e., those that acquired all relations), the potential exists that they were able to form three-member equivalences classes prior to the study. The procedures in the current study may have been more representative of teaching learners to form novel classes when the component skills for deriving were already present. Although consistent with previous research on IF (e.g., Delmolino et al. 2013), and EBI (Rehfeldt et al. 2003), the learner repertoires that are required to acquire skills that are not directly taught are somewhat elusive (Nottingham et al. 2015). Future research may benefit from further assessment of the component skills that may be necessary to acquire skills without explicit teaching in group contexts.

Some of the possible prerequisites have been hypothesized for IF procedures by Nottingham et al. (2015), and for stimulus EBI procedures by McLay et al. (2016), but these considerations may not take into account the skills that are necessary for a learner to fully benefit from group instruction. A more comprehensive analysis may be necessary similar to that of MacDonald and Ahearn (2015), where specific component skills were tested and taught prior to procedures that involved observational learning as an outcome (e.g., group instruction).

Third, in the current study, each participant was responsible for responding to only one element of the class presented. In group instructional contexts, peers may serve as a critical element in facilitating the acquisition of skills that are not explicitly taught. For example, a peer may provide correction for an incorrect answer or provide hints or additional information that may lead to acquisition. The current procedures were not specifically arranged to test the efficacy of

introducing peer-assisted learning (e.g., Petursdottir et al. 2007) into IF or EBI. Procedures were arranged such that learners were required to engage in similar attending responses across all trials to insure some level of attention toward stimuli presented regardless of the instructional trial. Although this aspect is present in the current investigation, it may not fully demonstrate the supports necessary to facilitate group instruction in a manner that equivalence classes may form. Future research may benefit from the addition of peer-assisted learning strategies to determine if the addition of peer interaction results in the emergence of novel stimulus classes.

Last, all of the targets in the current study were vocal responses (e.g., intraverbals). Although the current participants were vocally verbal, the extent to which the current results may be applied to learners with ASD that do not present with strong vocal repertoires is unknown. Previous research has demonstrated that learners without a well-developed vocal repertoire may acquire targets in a group instructional context (e.g., Griffen et al. 1992), but these demonstrations were limited in that they focused on observational, or secondary target acquisition. Future research may benefit from investigating the current procedures with listener responses, or by including participants that rely on an aided, augmentative, or alternative communication device.

With regard to feasibility, the current study was conducted in the course of regular services, in the participants' typical classroom, with members of their clinical teams. While the ASP sessions were lengthier, daily probe sessions lasted approximately 1–2 min for each participant (i.e., 3–6 min total). Treatment sessions lasted approximately 15 min, longer during earlier sessions when error corrections were required more frequently and shorter once targets were reaching mastery. Even for the participants that did not respond optimally (Dee and Maureen), mastery of directly taught targets was eventually achieved in the context of group learning. Thus, it is fairly safe to say that loss of instructional opportunities, by participating in group rather than one-on-one instruction, was minimal. Ultimately, for all participants, valuable information about their readiness for group learning was gained allowing the clinical team to tailor recommendations for their educational services across settings. We hope that this study offers a practical and clinician-friendly model of both designing instructional sessions and assessing readiness for group learning.

Compliance with Ethical Standards

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*. Philadelphia: American Psychiatric Pub.

Black, M. P., Skinner, C. H., Forbes, B. E., McCurdy, M., Coleman, M. B., Davis, K., et al. (2016). Cumulative instructional time and relative effectiveness conclusions: Extending research on response intervals, learning, and measurement scale. *Behavior Analysis in Practice, 9,* 58–62.

Cihak, D., Alberto, P. A., Taber-Doughty, T., & Gama, R. I. (2006). A comparison of static picture prompting and video prompting simulation strategies using group instructional procedures. *Focus on Autism and Other Developmental Disabilities, 21,* 89–99.

Collins, B. C. (2012). *Systematic instruction for students with moderate and severe disabilities*. Baltimore, MD: Brookes.

Collins, B. C., Gast, D. L., Ault, M. J., & Wolery, M. (1991). Small group instruction: Guidelines for teachers of students with moderate to severe handicaps. *Education and Training in Mental Retardation, 26,* 18–32.

Delmolino, L., Hansford, A. P., Fiske, K. E., & Larue, R. H. (2013). The use of instructive feedback for teaching language skills to children with autism. *Research in Autism Spectrum Disorders, 7,* 648–661.

Frampton, S. E., Robinson, H. C., Conine, D. E., & Delfs, C. H. (2017). An abbreviated evaluation of the efficiency of listener and tact instruction for children with autism. *Behavior Analysis in Practice, 10,* 131–144.

Griffen, A. K., Wolery, M., & Schuster, J. W. (1992). Triadic instruction of chained food preparation responses: Acquisition and observational learning. *Journal of Applied Behavior Analysis, 25,* 193–204.

Grow, L., & LeBlanc, L. (2013). Teaching receptive language skills. *Behavior Analysis in Practice, 6,* 56–75.

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

Jimenez, B. A., Browder, D. M., Spooner, F., & Dibiase, W. (2012). Inclusive inquiry science using peer-mediated embedded instruction for students with moderate intellectual disability. *Exceptional Children, 78,* 301–317.

Kleinert, H., Towles-Reeves, E., Quenemoen, R., Thurlow, M., Fluegge, L., Weseman, L., et al. (2015). Where students with the most significant cognitive disabilities are taught: Implications for general curriculum access. *Exceptional Children, 81,* 312–328.

Leaf, J. B., Cihon, J. H., Alcalay, A., Mitchell, E., Townley-Cochran, D., Miller, K., et al. (2017). Instructive feedback embedded within group instruction for children diagnosed with autism spectrum disorder. *Journal of Applied Behavior Analysis, 50,* 304–316.

LeBlanc, J. M., & Ruggles, T. R. (1982). Instructional strategies for individual and group teaching. *Analysis and Intervention in Developmental Disabilities, 2,* 129–137.

Ledford, J. R., Gast, D. L., Luscre, D., & Ayres, K. M. (2008). Observational and incidental learning by children with autism during small group instruction. *Journal of Autism and Developmental Disorders, 38,* 86–103.

Ledford, J. R., Lane, J. D., Elam, K. L., & Wolery, M. (2012). Using response-prompting procedures during small-group direct instruction: Outcomes and procedural variations. *American Journal on Intellectual and Developmental Disabilities, 117,* 413–434.

Ledford, J. R., & Wolery, M. (2015). Observational learning of academic and social behaviors during small-group direct instruction. *Exceptional Children, 81,* 272–291.

MacDonald, J., & Ahearn, W. H. (2015). Teaching observational learning to children with autism. *Journal of Applied Behavior Analysis, 48,* 800–816.

MacDonald, R. P. F., Dixon, L. S., & LeBlanc, J. M. (1986). Stimulus class formation following observational learning. *Analysis and Intervention in Developmental Disabilities, 6,* 73–87.

McLay, L. K., Church, J., & Sutherland, D. (2016). Variables affecting the emergence of untaught equivalence relations in children with and without autism. *Developmental Neurorehabilitation, 19,* 75–87.

McLay, L. K., Sutherland, D., Church, J., & Tyler-Merrick, G. (2013). The formation of equivalence classes in individuals with autism spectrum disorder: A review of the literature. *Research in Autism Spectrum Disorders, 7,* 418–431.

Morningstar, M. E., & Kurth, J. A. (2017). Status of inclusive educational placement for students with extensive and pervasive support needs. *Inclusion, 5,* 83–93.

Morningstar, M. E., Kurth, J. A., & Johnson, P. E. (2017). Examining national trends in educational placements for students with significant disabilities. *Remedial and Special Education, 38,* 3–12.

Nottingham, C. L., Vladescu, J. C., & Kodak, T. M. (2015). Incorporating additional targets into learning trials for individuals with autism spectrum disorders. *Journal of Applied Behavior Analysis, 48,* 1–6.

Petursdottir, A. L., McComas, J., McMaster, K., & Horner, K. (2007). The effects of scripted peer tutoring and programming common stimuli on social interactions of a student with autism spectrum disorder. *Journal of Applied Behavior Analysis, 40,* 353–357.

Plavnick, J. B., & Hume, K. A. (2014). Observational learning by individuals with autism: A review of teaching strategies. *Autism, 18,* 458–466.

Rehfeldt, R. A., Latimore, D., & Stromer, R. (2003). Observational learning and the formation of classes of reading skills by individuals with autism and other developmental disabilities. *Research in Developmental Disabilities, 24,* 333–358.

Shillingsburg, M. A., Frampton, S. E., Cleveland, S. A., & Cariveau, T. (2018). Clinical applications of procedures to promote the emergence of untrained intraverbal relations with children with autism. *Learning and Motivation, 62,* 51–66.

Sidman, M. (1994). *Equivalence relations and behavior: A research story*. Boston: Authors Cooperative.

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

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

Stahmer, A. C., Collings, N. M., & Palinkas, L. A. (2005). Early intervention practices for children with autism: Descriptions from community providers. *Focs on Autism and Other Developmental Disabilities, 20,* 66–79.

Stanley, C. R., Belisle, J., & Dixon, M. R. (2018). Equivalence-based instruction of academic skills: Application to adolescents with autism. *Journal of Applied Behavior Analysis, 51,* 352–359.

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

Sundberg, M. L., & Sundberg, C. A. (2011). Intraverbal behavior and verbal conditional discriminations in typically developing children and children with autism. *The Analysis of Verbal Behavior, 27,* 23–44.

Tullis, C. A., Frampton, S. E., Delfs, C. H., & Shillingsburg, M. A. (2017). Teaching problem explanations using instructive feedback. *The Analysis of Verbal Behavior, 33,* 64–69.

Werts, M. G., Wolery, M., Holcombe, A., & Gast, D. L. (1995). Instructive feedback: Review of parameters and effects. *Journal of Behavioral Education, 5,* 55–75.

Yaw, J., Skinner, C. H., Delisle, J., Skinner, A. L., Maurer, K., Cihak, D., et al. (2014). Measurement scale influences in the evaluation of sightword reading interventions. *Journal of Applied Behavior Analysis, 47,* 360–379.