Journal of Behavioral Education (2022) 31:265–280
https://doi.org/10.1007/s10864-020-09397-z

ORIGINAL PAPER

Establishing Requesting with Children Diagnosed with Autism Using Embedded Instruction in the Context of Academic Activities

Yvonne Cheung[1] · Christy Oi Yee Lai[1] · Joseph H. Cihon[2,3] · Justin B. Leaf[2,3] · Toby Mountjoy[4]

Published online: 1 July 2020

Abstract

Embedded instruction offers a potentially effective, non-disruptive, and socially acceptable intervention approach for individuals diagnosed with autism spectrum disorder (ASD) in general education settings. However, the literature using embedded instruction has not frequently provided data on embedded instruction targets and targets within the ongoing lessons or maintenance of the acquired skills. This study evaluated the effectiveness of embedded instruction to teach three individuals diagnosed with ASD communication skills during the course of existing lessons. Data were collected on embedded instruction targets, academic targets (i.e., targets within existing lessons), and maintenance of mastered targets. The results of a non-concurrent multiple baseline design indicated embedded instruction was effective for all three participants and the acquired skills maintained. The results are discussed with respect to future research and clinical application of the methods evaluated.

Keywords Embedded instruction · Communication · Autism · Discrete trial teaching

Introduction

Individuals diagnosed with autism spectrum disorder (ASD) commonly display social communication deficits (American Psychiatric Association 2013). As such, interventions commonly include targeting various communication skills (Barbera and Rasmussen 2007; Greer and Ross 2008; Petursdottir and Carr 2011; Sundberg

✉ Joseph H. Cihon
jcihon@autismpartnership.com

1 Aoi Pui School, Hung Hom, Hong Kong

2 Autism Partnership Foundation, 200 Marina Dr., Seal Beach, CA 90740, USA

3 Endicott College, Beverly, MA, USA

4 Autism Partnership Hong Kong, Kowloon, Hong Kong

and Partington 1998). When these intervention programs are based upon the science of behavior analysis, language is classified and targeted based upon Skinner's (1957) analysis of language and description of the four elementary verbal operants (i.e., echoic, tact, mand, and intraverbal). Given the social communication deficits common with a diagnosis of ASD, some have recommended targeting the development of mand relations as a strong focus during early stages of these interventions (e.g., Sundberg and Michael 2001). Mand relations "…may be defined as a verbal operant in which the response is reinforced by a characteristic consequence and is therefore under the functional control of relevant conditions of deprivation or aversive stimulation" (Skinner 1957, pp. 35–36). For example, when a child has not had access to juice for an extended period of time says, "I want juice" to their parents and the parents then provide the child juice.

It is not uncommon for interventions targeting the development of mand relations for individuals diagnosed with ASD to occur within home and clinical settings (Leaf et al. 2018). However, over the last few decades, there has been an increased emphasis on the inclusion of individuals diagnosed with developmental disabilities in general education settings (Polychronis et al. 2004). The push for an inclusive educational environment with respect to special education was sparked by Public Law 94-142 (Education for All Handicapped Children Act 1975) and its revisions (e.g., Individuals with Disabilities Education Improvement Act 1990). The push for inclusion may be most evident for individuals diagnosed with ASD. However, individuals diagnosed with ASD often engage in problematic behavior (e.g., stereotypy, aggression, and self-injury; Ala'i-Rosales et al. 2019) that can create a myriad of challenges within a general education setting. Given the increase in the number of individuals diagnosed with ASD entering general education settings and the potential challenges this can create, the need for effective, non-disruptive instructional strategies are essential.

One potential instructional strategy that has been documented to be effective within more structured educational settings is embedded instruction (e.g., Neef et al. 1984). Embedded instruction involves inserting learning opportunities into existing routines as the context for instruction (Neef et al. 1984; Sigafoos et al. 2009). For example, when an individual is engaged in a task related to the development of a pie chart, the teacher may insert learning opportunities to request materials by not providing all of the required materials. Using this approach, a teacher can embed learning opportunities targeting core deficits of ASD (e.g., communication) in an effective, non-disruptive manner within lesson plans designed for the whole class (Sutton et al. 2019). The use of embedded instruction has been well documented within the research as an effective approach to teach communication skills for individuals diagnosed with ASD.

In an early study, Neef et al. (1984) compared embedded instruction to tutoring when teaching four individuals diagnosed with ASD to respond to yes/no questions. Tutoring was similar to conventional discrete trial teaching (DTT; Leaf et al. 2016a) and consisted of contracting for a purported reinforcer, asking yes/no questions about five items (e.g., "Is this a fork?"), providing access to the reinforcer following correct responses, and modeling the correct response following incorrect responses. Embedded instruction consisted of a participant independently initiating a request which was

followed by the teacher asking a yes/no question (e.g., "Do you want ___?"). Correct responses resulted in access to the requested item, and incorrect responses resulted in a model of the correct response after a short delay. The results of a multiple baseline across students indicated all four participants acquired the yes/no responses with embedded instruction and none acquired them with the tutoring condition.

Sigafoos et al. (2009) also compared the effects of embedded instruction to a conventional DTT model on the self-injury, correct responding, and mood for a 12-year-old boy diagnosed with ASD. The hypothesis was that embedded instruction would result in less self-injury and overall mood than DTT. Embedded instruction involved opportunities to respond every 30 s within three different activities (i.e., swinging, walking, and music). Each activity was paused and resumed if the participant pressed a switch within 10 s of the pause. DTT involved presenting a trial every 10 s and providing praise and a pat on the back following correct responses and prompting following incorrect trials. The results indicated that embedded instruction was more effective with respect to less self-injury, more correct responding, and better mood ratings.

While the aforementioned studies, as well as many others, have demonstrated the effectiveness of embedded instruction in isolation and compared to other approaches, literature reviews on embedded instruction have cited some gaps within the current literature. Rakap and Parlak-Rakap (2011) reviewed 16 studies that evaluated the effectiveness of embedded instruction for preschool age children diagnosed with disabilities in inclusive preschool programs. Within their review, Rakap and Parlak-Rakap cited that future research is necessary with children with different abilities while also including maintenance data. It should also be noted that while some researchers have evaluated the effectiveness of embedded instruction to develop communication skills (e.g., Christensen-Sandfort and Whinnery 2013), none of the studies Rakap and Parlak-Rakap reviewed targeted mand relations (often referred to as requesting), and only one of the studies mentioned above (i.e., Sigafoos et al. 2009) targeted requesting. As such, research is necessary to address the gaps within the current literature on embedded instruction. It is possible that methods used to develop mand relations within clinical settings, such as the interrupted chain procedure (e.g., Albert et al. 2012; Lorah et al. 2014), could be embedded within existing lessons within the school setting.

The purpose of the present study was to examine the effectiveness of an embedded instruction approach to teach three school-aged individuals diagnosed with ASD communication skills (i.e., requesting) during the course of existing lessons. To address the limitations within the current literature base, data were collected on embedded instruction targets, academic targets (i.e., targets within existing lessons), and maintenance of mastered targets.

Method

Participants

Three children participated in the study. Each participant had an independent diagnosis of ASD and was currently enrolled in a private school in Hong Kong that

provided behavioral intervention for students diagnosed with ASD. Danny was an 8-year-old male with a Vineland-3 Adaptive Behavior Scales (VABS-3; Sparrow et al. 2005) Composite Score of 79, Walker–McConnell Scale of Social Competence and School Adjustment Elementary Version total scale score of 144, Social Responsiveness Scale-2 T-score of 60, and Social Skills Improvement System (SSiS; Gresham and Elliott 2008) social skill scale standard score of 94, problem behavior scale standard score of 106, academic competence skill standard score of 85. Danny communicated using full sentences (e.g., "I want iPad," "Can you help me to open my lunch box?"). Danny could also count quantities up to 100 and compare quantities within the range of 1–100. Danny displayed deficits in clarifying when someone misunderstood his requests (e.g., when he requested an item but was given a different item) and using superlatives (e.g., a lot, a few, and no) to express numbers and quantities.

Archie was a 6-year-old male with a VABS-3 Composite Score of 84, the Walker–McConnell Scale of Social Competence and School Adjustment Elementary Version total scale score of 155 (percentile rank: 39), Social Responsiveness Scale-2 T-score of 61 (mild range of deficiencies in reciprocal social behavior) and SSiS social skill scale standard score of 79, problem behavior scale standard score of 105, academic competence skill standard score of 81. Archie communicated using full sentences (e.g., "I want to play iPad," "Where is my water bottle?"). Archie could count quantities up to 100 and compare quantities within the range of 1–50.

Timothy was an 8-year-old male with a VABS-3 Composite Score of 59, the Walker–McConnell Scale of Social Competence and School Adjustment Elementary Version total scale score of 61, Social Responsiveness Scale-2 T-score of 85 and SSiS social skill scale standard score of 50, problem behavior scale standard score of 138, academic competence skill standard score of 88. Timothy communicated using single words (e.g., "iPad") or simple sentences (e.g., "I want blocks"). Timothy could count quantities up to 100 and compare quantities within the range of 1–100. Timothy could not request a specific quantity of an item (e.g., "I want 10 goldfish") or use superlatives to express numbers and quantities (e.g., a lot, a few, and no).

Each participant was paired with two other children during all sessions across all conditions. Two children were included because all the participants required intensive intervention to acquire communication and academic skills when the study took place. Their teachers therefore taught in small groups of three to provide them with intensive intervention on the one hand and opportunities to acquire group learning skills on the other hand.

Teachers

Given the use of flexible prompt fading (described later), the repertoires of the teachers may be important for application and replication. There were a total of three adults that served as teachers for all intervention sessions. Jamie served as the teacher for Danny. Jamie was a 26-year-old female with a bachelor's degree of Science in Psychology. At the time of the study, she had been working in the school

in which the study took place for 1.5 years. She was responsible for implementing behaviorally based intervention in a class of eight students (ages ranging from 7 to 8 years) diagnosed with ASD.

Cheryl served as the teacher for Archie. Cheryl was a 26-year-old female with a Bachelor ‘s degree of Arts in Psychology. At the time of the study, she had been working in the school in which the study took place for 3 years and was pursuing a master’s degree in Applied Behavior Analysis. She was responsible for implementing behaviorally based intervention in a class of eight students (ages ranging from 5 to 6 years) diagnosed with ASD.

Michelle served as the teacher for Timothy. She was a 25-year-old female with a bachelor’s degree in Linguistics and Modern Languages. At the time of the study, she had been working in the school in which the study took place for 2 years. She was responsible for implementing behaviorally based intervention in a class of eight students (ages ranging from 7 to 8 years) diagnosed with ASD.

All three teachers had received ongoing training on behaviorally based interventions for teaching skills including learning how to learn, communication, and social. This training primarily consisted of identifying the skill, modeling the skill, role playing, and feedback. The three teachers also participated in a 7-week intensive induction training when they began to work in the school. The induction training included didactic and supervised hands-on learning. In addition, all teachers were provided with coaching by a Board Certified Behavior Analyst (BCBA) or class supervisor on an average of 6 to 8 h per week. The coaching was tailored to the individual training need of each teacher.

Setting and Materials

All sessions were conducted in a school in Hong Kong that specializes in teaching individuals diagnosed with ASD. The school primarily employs educational methods based upon applied behavior analysis. Sessions occurred in the participants’ classroom, and all intervention was provided by the participants’ classroom teachers. The classroom included a white board, four tables, and eight chairs. Sessions occurred up to 5 times per week and lasted approximately 20 min.

The materials were selected based on each student’s target skill. Materials for Danny and Timothy included a whiteboard, white board marker, pictures of food items in McDonalds (e.g., hamburger, French fries, chicken nuggets) and Pizza Hut (e.g., pizza, pasta, coke), and three index cards reading “A Lot,” “A Few,” and “No.” Materials for Archie included pie charts, worksheet, blue tag, stickers, stamps, whiteboard, marker, and eraser.

Dependent Measures

The primary dependent variable was an individualized communicative response following the teacher desired evocative event (e.g., not providing access to an item required to complete a task while presenting a demand to complete the task). The term *evocative event* was adopted from the functional analysis literature for events

that set the occasion for a specific behavior (for example, see Slaton et al. 2017). The target skills were selected based on observations of each participant engaging or not engaging in communication and each participant's age. Danny displayed deficits in clarifying requests when the listener misunderstood his request. Therefore, the communicative response for Danny was stating, "Not this one, I want [object]" following the teacher providing him the wrong item. Archie displayed deficits in alerting the teacher when he was missed. As such, the communicative response for Archie was stating, "What about me?" following the teacher providing each student a material except him. During observations of Timothy, he displayed deficits in requesting a specific quantity of an item. Therefore, the communicative response for Timothy was stating "I want [quantity] [object]." The evocative event was providing Timothy fewer items than required to complete a task. For example, if Timothy was asked to put eight pizzas on the white board, he was only given two pizzas.

Correct responses were defined as the participant engaging in the communicative response within 5 s of the evocative event. For instance, Timothy stating "I want six pizzas" following the previous example of his evocative event. Incorrect responses were defined as not engaging in the communicative responses within 5 s of the evocative event or engaging in a communicative response different from the target response. For instance, Timothy stating "pizzas" or nothing following the previous example of his evocative event. Responding was converted into a percentage by dividing the number of correct responses by the total number of correct and incorrect responses multiplying by 100%. The mastery criterion was set at a participant engaging in correct responses during at least 80% of trials across three consecutive daily probes.

The second dependent variable was the number of correct academic responses, which were also individualized for each participant. The school in which the study took place followed the UK curriculum (see National Curriculum 2014 for a complete description). The targeted academic skills were selected in reference to the participants' skill level, grade level, and the functionality of the skills. Danny and Timothy were in grade 2, and Archie was in grade 1. The targeted academic skill for Danny and Timothy was describing quantities of items using the descriptors "a lot," "a few," and "no." The teacher presented the students two items of different quantities (e.g., ten pizzas and two hotdogs) and asked the students to "describe the amount/quantity of [object]." Correct responses were defined as using the descriptor that correlated with the quantity (i.e., "a lot of" to describe the larger quantity, "a few" to describe the fewer quantity, and "no [object]" if an item was not present) within 5 s of the teacher's instruction. Incorrect responses were defined as using a descriptor that did not correlate with the quantity or failing to respond within 5 s of the teacher's instruction.

The targeted academic skill for Archie was identifying the most and the least selected option and the number of students who selected a particular option on a pie chart. The teacher showed Archie a pie chart and asked him, "Which one is the most popular?", "Which one is the least popular?", or "How many people chose [option]?" in a random order. Correct responses were defined as engaging in a vocal–verbal response that correlated with the question within 5 s of the teacher's question. Incorrect responses were defined as engaging in a vocal–verbal response

that did not correlate with the question or failing to respond within 5 s of the teacher's question. The mastery criterion was set at a participant engaging a correct response on the first-trial probe across three consecutive sessions.

Baseline

Baseline sessions occurred once per weekday up to five times a week. Each session consisted five trials for the communicative response and three trials for the academic response for a total of eight trials per session. The trials for communicative response and academic response occurred in random order. On each communicative response trial, the teacher presented an evocative event that set an occasion for the participant to engage in the targeted communicative response, for example, asking everyone in the group to color a segment of the pie chart but Archie, giving Danny a pizza following his request for a hot dog, or asking Timothy to put six hamburgers on the whiteboard when he only had two. On each academic response trial, the teacher issued an instruction based upon the participants' individual academic skill (e.g., "Which one is the least popular?"). No corrective feedback or programmed reinforcement occurred during probe trials regardless of the response.

Daily Probes

Daily probes occurred prior to each intervention session. They were conducted in the same manner as baseline. The purpose of daily probes was to provide a repeated measure throughout intervention in the absence of programmed reinforcement to assess mastery. Each probe consisted of five trials for communicative responses and three trials for academic responses. The trials for communicative and academic responses occurred in random order. No corrective feedback or programmed reinforcement occurred during probe trials regardless of the response. Following the daily probe, the teacher started teaching the group.

Intervention

General

All intervention sessions occurred with the participants and two other children (previously described). Sessions occurred once per day up to five times a week and lasted approximately 20 min.

Academic Task

The academic tasks were the same for Danny and Timothy. The room was arranged to resemble a fast food restaurant. Food items were arranged on the table or a white board. The participants were required to prepare an order of food by obtaining the required items for the order. To begin, the therapist provided an instruction for the children in the group to obtain specific quantities of items. After the children obtained the items,

the teacher provided an instruction to describe the quantity of each of the items (i.e., "a lot of" to describe the larger quantity, "a few" to describe the fewer quantity, and "no [object]" if an item was not present). The teacher used flexible prompt fading (FPF; Soluaga et al. 2008) to teach the correct responses to the instruction. Prompts generally consisted of written (e.g., a notecard with "a lot" written on it), gestural (e.g., pointing to an item), and verbal models (e.g., "There are a lot of pizzas."). The prompts were faded based on the teacher's assessment of the participant's responding on current and previous trials.

The academic task for Archie consisted of a math lesson. To begin the lesson, the teacher arranged numbers that correlated with a survey on a table (e.g., favorite foods of the teachers). The survey numbers differed each day. The children were then instructed to arrange the survey results into a pie chart. During this instruction, the teacher showed Archie a pie chart and asked him, "Which one is the most popular?", "Which one is the least popular?," or "How many people chose [option]?" in a random order. The teacher used FPF to teach the correct responses to the instruction. Prompts generally consisted of written (e.g., a notecard with "most" written on it), within-stimulus (e.g., emphasizing the word "most" when asking Archie "Which one is the most popular?"), and verbal prompt (e.g., "most means biggest"). The prompts were faded based on the teacher's assessment of the participant's responding on current and previous trials.

Embedded Instruction

The teacher embedded the teaching of the targeted communication skills in the academic activity. For Danny and Timothy, this consisted of leaving out items that they were required to obtain for the fast food order. When Danny requested the missing item from the teacher, the teacher provided him with the wrong item (i.e., and item that did not correlate with the requested item). Danny was then required to engage in the targeted communicative response (e.g., stating, "Not this one, I want [object]"), while Timothy was required to request the specific quantity of the items missing. For Archie, the teacher either provided an instruction for the children to complete as task related to the pie chart (e.g., telling two of the children, but not Archie, what colors to use when coloring the pie chart) or provided items to each of the children and not to Archie. Archie was then required to engage in the targeted communicative response (e.g., stating "What about me?").

Similar to the academic targets, the teacher used FPF to teach the correct responses. Prompts generally consisted of full verbal models (e.g., "I want four pizzas."). The prompts were faded based on the teacher's assessment of the participant's responding on current and previous trials. Anytime the participants did not respond following the evocative event and a prompt was not provided, the teacher provided feedback (e.g., "You didn't say anything") and moved on to the next trial.

Maintenance

Maintenance was assessed at least 5 days after the completion of the intervention phase. Each participant had three maintenance sessions at least 5 days apart from each other. Maintenance sessions were conducted in the same manner as baseline.

Design

A non-concurrent multiple baseline design (Watson and Workman 1981; Kazdin 2011) across participants was used to evaluate the effects of the intervention on participant responding. Non-concurrent multiple baseline designs provide flexibility when conducting research in applied settings, such as this study, that a concurrent multiple baseline design may not. Within this design, baseline phases are typically selected a priori, and participants are randomly assigned to each baseline length as they become available. Participants in this study progressed from baseline to the intervention condition once a stable level of commenting was observed during baseline; however, in an attempt to improve the strength of this design, an additional criterion common within multiple baseline logic was used. If necessary, we extended baseline sessions for the next participant until intervention effects were observed with the previous participant. As such, experimental control was demonstrated when the intervention resulted in changes in a participant's behavior without changes in the remaining participants' behavior during baseline sessions (Baer et al. 1968; Carr 2005). Participants progressed to the maintenance condition when they had reached the mastery criterion (i.e., engaging in correct communicate responses during at least 80% of trials and engaging a correct academic response on the first-trial probe across three consecutive daily probes).

Interobserver Agreement

The experimenter and an independent observer recorded the participants' responses during 35.71%, 35.29%, and 50% of sessions for Danny, Archie, and Timothy, respectively. The independent observer was trained on the operational definitions of the targeted communicative responses and academic responses. Interobserver agreement (IOA) was calculated by totaling the number of times the experimenter, and the independent observer agreed on the scoring of a response divided by the total number of agreements plus disagreements and multiplying by 100%. IOA between the experimenter and the independent observer was 100% for baseline, intervention, and maintenance across all participants.

Social Validity

The three teachers responsible for conducting sessions completed a questionnaire about their perception of the effectiveness and acceptability of embedded

instruction as a teaching strategy upon the completion of the maintenance phase. The questionnaire was adapted from Johnson et al. (2004). The questionnaire consisted of eight questions. The teacher rated each question using a seven-point Likert scale with 1 representing the lowest score (e.g., poor, very costly, and not effective) and 7 representing the highest score (e.g., excellent, not costly, and very effective). The questions included: (1) how well did embedded instruction procedure address the needs of the student?, (2) how costly (in terms of resources) was it to carry out embedded instruction?, (3) how likely is embedded instruction to make permanent improvement in learning for the student?, (4) How likely is it that you would use embedded instruction again?, (5) how disruptive was embedded instruction to ongoing classroom instruction?, (6) How effective was embedded instruction in teaching the student academic skills?, (7) how effective was embedded instruction in teaching the student spontaneous communication skill?, and (8) how serious were the learning problems of the student participating in the study compared to other students in the class?.

Results

Figure 1 displays the results for all three participants. Each panel represents a different participant's results. The top panel depicts the results for Danny. During baseline, Danny did not engage in the targeted communication response during any session. Although he did engage in the targeted academic task during one session, responding was low across the three baseline sessions. Following intervention, Danny began to use the target spontaneous communication skill on 100% of trials with an exception of the first probe. Danny reached the mastery criterion for the communicative response on the fourth session; however, his performance on the academic response remained variable. Danny eventually reached mastery criterion in 11 sessions. Danny continued to engage in the targeted communication response on 100% of trials following the conclusion of intervention, as depicted in the maintenance condition. He also engaged in the targeted academic task on 100% of trials during the first two maintenance probes and decreased to 66% of trials during the last maintenance probe.

The middle panel depicts the results for Archie. During baseline, Archie did not engage in the targeted communication response during any session and engaged in variable responding with respect to the academic task. Following intervention, Archie quickly reached the mastery criteria within four and five sessions for the targeted academic and communication responses, respectively. Archie continued to engage in the targeted academic and communication responses at similar percentages during the assessment of maintenance.

The bottom panel depicts the results for Timothy. During baseline Timothy did not engage in the targeted academic response during any session and only engaged in the targeted communication response in one session. Following intervention, Timothy reached the mastery criteria within seven and five sessions for the targeted academic and communication responses, respectively. Timothy continued to engage

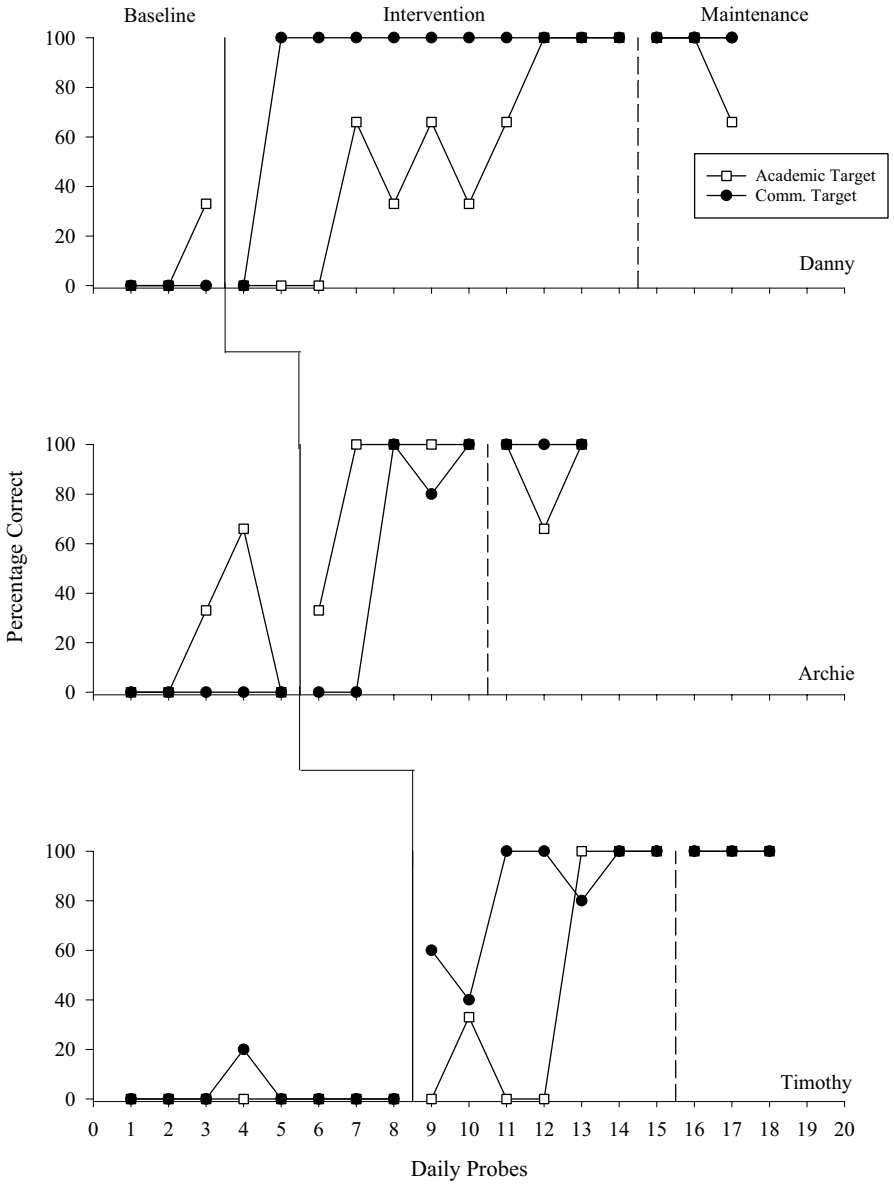

Fig. 1 Results for all three participants. Closed circles represent the individualized communicative responses. Open squares represent the academic targets

in the targeted academic and communication responses during 100% of trials during the assessment of maintenance.

Tau-U

Tau-U, which combines analyses of trend within phases and overlapping data points between phases, was calculated for each of the participants' targeted communicative response during baseline and intervention (Parker et al. 2011). Maintenance data were not used in the Tau-U calculation. The Tau scores for baseline trend for Danny, Archie, and Timothy were 0, 0, and −0.0357, respectively. As such, no corrections for baseline trends were required. Phase contrasts were then conducted for each participant. Comparing baseline and intervention phases for all three participants yielded Tau-U scores of 0.9091, 0.6, and 1 for Danny, Archie, and Timothy, respectively. These scores represent a large (i.e., .60 to .80) to very large (i.e., .80 to 1) change (Parker et al. 2011). When combining contrasts into a weighted average, the overall effect across participants indicated a very large change (Tau-U = 0.8482).

Social Validity

With respect to questions related to useful and effectiveness (i.e., questions 1, 3, 6, and 7), all of the teachers rated the intervention using embedded instructions as useful with (i.e., average scores of 6, 6, 5, and 6.33, respectively). The teachers also indicated they were likely to use the intervention in the future (i.e., questions 4, 5) with average scores of 6.67 and 5.67, respectively. However, it should be noted that, on average, the teachers rated the intervention as costly with respect to the resources involved (i.e., an average score of 4.33 on question 2).

Discussion

The purpose of the present study was to examine the effectiveness of an embedded instruction approach to teach three school-aged individuals diagnosed with ASD communication skills (i.e., requesting) during the course of existing lessons. All three of the participants acquired the individualized communicative responses as well as the academic targets. These results are consistent with the previous research that has demonstrated the effectiveness of embedded instruction for students diagnosed with ASD (e.g., Neef et al. 1984; Polychronis et al. 2004; Sigafoos et al. 2009). The individualized communicative responses as well as the academic targets also maintained for all three participants.

This study expands upon the literature on embedded instruction in at least three ways. First, data were collected on embedded instruction targets and academic targets (i.e., targets within existing lessons). Many studies within the current literature only report data on embedded instruction targets and not on targets of the existing, or ongoing, lessons (e.g., Neef et al. 1984). Second, maintenance of the acquired individualized communicative responses and academic targets was collected for all

three participants. Literature reviews of embedded instruction studies (e.g., Rakap and Parlak-Rakap 2011) have noted that many do not report the maintenance of mastered targets. Taken together, these two contributions provided evidence that embedded instruction targets and targets of the existing lesson can not only be acquired using this method, but they are also likely to maintain.

A third contribution of this study to the literature on embedded instruction are the methods used to develop mand relations. The methods embedded in the academic tasks may not be novel, but they are within the embedded instruction literature. For instance, the methods closely resemble those used within interrupted chain procedures (e.g., Albert et al. 2012; Lorah et al. 2014). However, it should be noted that there are differences in the methods employed in the current study and interrupted chain procedures (e.g., providing the wrong item and providing access to others but not the participant). Nonetheless, this study demonstrated that procedures for the development of mand relations could be embedded within ongoing academic activities.

This study did no go without limitations that warrant discussion. First, the generalization of the acquired skills was not assessed. As such, the extent to which the participants generalized the individualized communication responses and academic skills remains uncertain. Future researchers should attempt to replicate the results of this study and include an assessment of generalization across different people and settings. Second, the teaching method employed (i.e., FPF) for teaching the individualized communication responses and academic skills required clinical judgement. While approaches that involve clinical judgement likely resemble actual practice, they create replicability problems for future researchers (Cihon et al. 2019a, c). Third, and relatedly, given the flexibility and clinical judgement permitted in this intervention, treatment fidelity data were not collected. That is, there was no predetermined protocol for when to prompt, what prompt to provide, or how to fade prompts. As such, comparing performance during intervention to a predetermined protocol was not possible to determine treatment fidelity. It should be noted, however, FPF has a growing literature base indicating its effectiveness (e.g., Cihon et al. 2019b; Ferguson et al. 2020; Leaf et al. 2019; Soluaga et al. 2008). Given the environment in which the study took place and the flexibility in the intervention, the results should be interpreted with caution. Future researchers could help to address these limitations by evaluating agreement with an independent observer with respect to the interventionists' choice as to whether to prompt, when to prompt, and how to prompt. This would provide data on how likely an independent observer would be to respond similarly in similar situations.

Despite the aforementioned limitations, there are at least two important clinical implications of the present study. First, the increase in the number of individuals diagnosed with ASD entering general education settings creates potential challenges for instructors to individualize intervention, which is commonly recommended for individuals diagnosed with ASD (Leaf et al. 2016b). This study demonstrated an effective, non-disruptive instructional strategy that can be used in traditional educational settings and permit individualization. Second, previous researchers have suggested that embedded instruction strategies could mitigate concerns regarding the social invalidity of other traditional methods (e.g., discrete trial teachings; Geiger

et al. 2012). It should be noted, however, there is general disagreement on the methods that should be employed within discrete trial teaching (see Grow and LeBlanc 2013; Leaf et al. 2016a). Nonetheless, the results of the assessment of social validity within this study support the assertion that embedded instruction is a socially acceptable intervention.

To conclude, this study demonstrated that embedded instruction is an effective and socially valid method of instruction without sacrificing individualization for individuals diagnosed with ASD in a traditional educational setting. All three participants acquired and maintained individualized communicative responses as well as academic targets. We hope that this study, in combination with the literature documenting the effectiveness of embedded instruction, will result in practitioners integrating embedded instruction when working in traditional educational settings.

Compliance with Ethical Standards

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

Ethical approval 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.

References

Ala'i-Rosales, S., Cihon, J. H., Currier, T. D. R., Ferguson, J. L., Leaf, J. B., Leaf, R., et al. (2019). The big four: Functional assessment research informs preventative behavior analysis. *Behavior Analysis in Practice, 12*(1), 222–234. https://doi.org/10.1007/s40617-018-00291-9.

Albert, K. M., Carbone, V. J., Murray, D. D., Hagerty, M., & Sweeney-Kerwin, E. J. (2012). Increasing the mand repertoire of children with autism through the use of an interrupted chain procedure. *Behavior Analysis in Practice, 5*(2), 65–76. https://doi.org/10.1007/bf03391825.

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

Barbera, M. L., & Rasmussen, T. (2007). *The verbal behavior approach: How to teach children with autism and related disorders*. London: Kingsley.

Baer, D. M., Wolf, M. M., & Risley, T. R. (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis*, *1*(1), 91–97. https://doi.org/10.1901/jaba.1968.1-91.

Christensen-Sandfort, R. J., & Whinnery, S. B. (2013). Impact of milieu teaching on communication skills of young children with autism spectrum disorder. *Topics in Early Childhood Special Education, 32*(4), 211–222. https://doi.org/10.1177/0271121411404930.

Cihon, J. H., Ferguson, J. L., Leaf, J. B., Leaf, R., McEachin, J., & Taubman, M. (2019a). Use of a level system with flexible shaping to improve synchronous engagement. *Behavior Analysis in Practice, 12*(1), 44–51.

Cihon, J. H., Ferguson, J. L., Leaf, J. B., Milne, C. M., Leaf, R., & McEachin, J. (2019b). A randomized clinical trial of three prompting systems to teach tact relations. *Journal of Applied Behavior Analysis*. https://doi.org/10.1002/jaba.617.

Cihon, J. H., Ferguson, J. L., Milne, C. M., Leaf, J. B., McEachin, J., & Leaf, R. (2019c). A preliminary evaluation of a token system with a flexible earning requirement. *Behavior Analysis in Practice, 12*(3), 548–556.

Education for All Handicapped Children Act of 1975, Pub. L. No. 94-142, 89 Stat. 773 (1975) (codified as amended at 20 U.S.C. §§ 1400–1461 (1976)).

Ferguson, J. L., Milne, C. M., Cihon, J. H., Dotson, A., Leaf, J. B., McEachin, J., et al. (2020). An evaluation of estimation data collection to trial-by trial data collection during discrete trial teaching. *Behavioral Interventions, 35*(1), 178–191. https://doi.org/10.1002/bin.1705.

Greer, R. D., & Ross, D. E. (2008). *Verbal behavior analysis: Inducing and expanding complex communication in children with severe language delays*. Boston, MA: Allyn & Bacon.

Gresham, F. M., & Elliott, S. N. (2008). *Social skills improvement system: Rating scales*. Bloomington, MN: Pearson Assessments.

Grow, L., & LeBlanc, L. (2013). Teaching receptive language skills: Recommendations for instructors. *Behavior Analysis in Practice, 6*(1), 56–75.

Individuals with Disabilities Education Act of 1990, Pub. L. No. 104-476, 104 Stat. 1103, 1141–42 (1990) (codified as amended at 20 U.S.C. §§ 1400-1482).

Carr, J. E., (2005). Recommendations for reporting multiple-baseline designs across participants. *Behavioral Interventions*, *20*(3), 219–224.

Johnson, J. W., McDonnell, J., Holzwarth, V. N., & Hunter, K. (2004). The efficacy of embedded instruction for students with developmental disabilities enrolled in general education classes. *Journal of Positive Behavior Interventions, 6*(4), 214–227.

Geiger, K. B., Carr, J. E., LeBlanc, L. A., Hanney, N. M., Polick, A. S., & Heinicke, M. R. (2012). Teaching Receptive Discriminations to Children With Autism: A Comparison of Traditional and Embedded Discrete Trial Teaching. *Behavior Analysis in Practice*, *5*(2), 49–59.

Kazdin, A. E. (2011). *Single-case research designs: Methods for clinical and applied settings*. New York, NY: Oxford University Press.

Leaf, J. B., Cihon, J. H., Ferguson, J. L., Leaf, R., & McEachin, J. (2019). Comparing no-no prompt to flexible prompt fading to teach expressive labels to individuals diagnosed with autism spectrum disorder. *Education and Training in Autism and Developmental Disabilities, 54*(3), 274–287.

Leaf, J. B., Cihon, J. H., Leaf, R., McEachin, J., & Taubman, M. (2016a). A progressive approach to discrete trial teaching: Some current guidelines. *International Electronic Journal of Elementary Education, 9*(2), 361–372.

Leaf, J. B., Leaf, R., McEachin, J., Cihon, J. H., & Ferguson, J. L. (2018). Advantages and challenges of a home- and clinic-based model of behavioral intervention for individuals diagnosed with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 48*(6), 2258–2266. https://doi.org/10.1007/s10803-017-3443-3.

Leaf, J. B., Leaf, R., McEachin, J., Taubman, M., Ala'i-Rosales, S., Ross, R. K., et al. (2016b). Applied behavior analysis is a science and therefore, progressive. *Journal of Autism and Developmental Disorders, 46*(2), 720–731. https://doi.org/10.1007/s10803-015-2591-6.

Lorah, E. R., Gilroy, S. P., & Hineline, P. N. (2014). Acquisition of peer manding and listener responding in young children with autism. *Research in Autism Spectrum Disorders, 8*(2), 61–67. https://doi.org/10.1016/j.rasd.2013.10.009.

National Curriculum (2014). Retrieved from: https://www.gov.uk/government/collections/national-curriculum.

Neef, N. A., Walters, J., & Egel, A. L. (1984). Establishing generative yes/no responses in developmentally disabled children. *Journal of Applied Behavior Analysis, 17*(4), 453–460. https://doi.org/10.1901/jaba.1984.17-453.

Parker, R. I., Vannest, K. J., & Davis, J. L. (2011). Effect size in single-case research: A review of nine nonoverlap techniques. *Behavior Modification, 35*(4), 303–322. https://doi.org/10.1177/0145445511399147.

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.

Polychronis, S. C., McDonnell, J., Johnson, J. W., Riesen, T., & Jameson, M. (2004). A comparison of two trial distribution schedules in embedded instruction. *Focus on Autism and other Developmental Disabilities, 19*(3), 140–151. https://doi.org/10.1177/10883576040190030201.

Rakap, S., & Parlak-Rakap, A. (2011). Effectiveness of embedded instruction in early childhood special education: A literature review. *European Early Childhood Education Research Journal, 19*(1), 79–96.

Sigafoos, J., O'Reilly, M., Ma, C. H., Edrisinha, C., Cannella, H., & Lancioni, G. E. (2009). Effects of embedded instruction versus discrete-trial training on self-injury, correct responding, and mood in a child with autism. *Journal of Intellectual and Developmental Disability, 31*(4), 196–203. https://doi.org/10.1080/13668250600999160.

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

Slaton, J. D., Hanley, G. P., & Raftery, K. J. (2017). Interview-informed functional analyses: A comparison of synthesized and isolated components. *Journal of Applied Behavior Analysis, 50*(2), 252–277. https://doi.org/10.1002/jaba.384.

Soluaga, D., Leaf, J. B., Taubman, M., McEachin, J., & Leaf, R. (2008). A comparison of flexible prompt fading and constant time delay for five children with autism. *Research in Autism Spectrum Disorders, 2*(4), 753–765. https://doi.org/10.1016/j.rasd.2008.03.005.

Sparrow, S. S., Cicchetti, D. V., & Balla, D. A. (2005). *Vineland adaptive behavior scales 2nd edition (Vineland-II)*. Circle Pines: American Guidance Service.

Sundberg, M. L., & Michael, J. (2001). The benefits of Skinner's analysis of verbal behavior for children with autism. *Behavior Modification, 25*(5), 698–724. https://doi.org/10.1177/0145445501255003.

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

Sutton, B. M., Webster, A. A., & Westerveld, M. F. (2019). A systematic review of school-based interventions targeting social communication behaviors for students with autism. *Autism, 23*(2), 274–286. https://doi.org/10.1177/1362361317753564.

Watson, P. J., & Workman, E. A. (1981). The non-concurrent multiple baseline across-individuals design: an extension of the traditional multiple baseline design. *Journal of Behavior Therapy and Experimental Psychiatry, 12*(3), 257–259.