



An Experimental Comparison of In-Person and Remote Instruction for Preschoolers with Disabilities

Joseph Peysin¹ · Daniel Fienup¹  · Stavra Romas²

Accepted: 15 August 2022
© Association for Behavior Analysis International 2022

Abstract

Remote instruction is becoming increasingly common, yet few studies have directly compared remote and in-person instruction in a controlled manner. We used a reversal design to compare the effects of in-person and remote instruction for six preschool participants with disabilities learning facts and sight words. Distribution of instruction, methodology, and materials across in-person and remote conditions were equated so that the only difference across conditions was the modality of instruction. Across conditions, we measured (1) the rate of learning; (2) the rate of trial presentation; (3) number of targets mastered; and (4) percentage of correct responses during follow-up assessment. Results indicate that three of six participants reliably met acquisition criteria and completed instruction faster in-person, with mixed results for the other three participants. No consistent difference was observed in response maintenance or generalization across modalities. These findings add to existing literature suggesting that remote instruction should be considered in situations where in-person instruction is unavailable.

Keywords COVID-19 · in-person instruction · telehealth · remote instruction

During the COVID-19 pandemic, governmental orders led to the closure of schools in all 50 states in the United States (Fronapfel & Demchak, 2020). As a result of this, service providers across the nation were presented with an unprecedented challenge of delivering nearly all educational services (both behavior analytic and not) remotely. As the shift to remote service delivery was due to a public health crisis, which had no precedent in the last century, the literature basis for this topic is in its early stages. As such, this context provides an opportunity for additional research to address many unanswered questions.

Applied behavior analysis (ABA) experimenters and practitioners have long used remote service delivery to reach individuals living in geographically remote locations where in-person services were not available (Higgins et al., 2017). There have been a number of studies demonstrating the effectiveness of remote ABA service delivery in domains such as conducting remote brief functional

behavior assessments (FBAs; Barretto et al., 2006), providing a remote training course for ABA technicians (Fisher et al., 2014), and training teachers to use positive reinforcement strategies to decrease problem behavior in a classroom (Knowles et al., 2017). Additional data also suggests the efficiency of remote ABA services for training implementers (Hay-Hansson & Eldevik, 2013). Finally, a systematic review (Tomlinson et al., 2018) of training individuals implementing ABA procedures remotely concluded that the emerging literature suggests that this may be an effective modality and should be considered as an option, given the cost and travel burdens were significantly decreased compared to support provided in-person.

Other disciplines have conducted large scale comparisons between in-person and remote delivery of services. For example, both speech and language pathology (Sutherland et al., 2018) and occupational therapy (Little et al., 2018) have compared modalities. However, few studies in ABA have directly compared the two modalities. One such comparison of ABA was conducted by Lindgren et al. (2016), who compared three service delivery models and their respective outcomes and costs for implementing ABA interventions to reduce problem behavior for 107 participants. The researchers compared in-person/in-home therapy, clinic-based telehealth, and home-based telehealth. The results

✉ Daniel Fienup
fienup@tc.columbia.edu

¹ Teachers College, Columbia University, New York, NY, USA

² The Fred S. Keller School, Yonkers, NY, USA

demonstrated that all three delivery models were successful in reducing problem behavior by more than 90% using an intervention package consisting of coaching parents to conduct a functional analysis (FA) and training parents to accurately implement functional communication training (FCT). Further, the total costs for the telehealth models were significantly less than in-home therapy.

In addition, two studies have examined in-person and remote instructions when the goal of ABA is to promote children's acquisition of novel academic responses—or educational applications of ABA. Pollard et al. (2021) reported 17 cases of clients who transitioned from receiving in-person ABA services to receiving services remotely. Nearly all students maintained or improved correct responses across all programs as well as maintained the same frequency of instructional sessions remotely as they received in-person. Oblak (2021) reported successful transition of service delivery to telehealth across a system-wide model using the Comprehensive Application of Behavior Analysis to Schooling (CABAS®) model. Measures of service delivery reported by Oblak included the total number of trials delivered and number of trials needed to master an objective. In both of these measures, learning outcomes reported during remote service delivery during the pandemic were comparable or even superior to data collected prior to the pandemic. Both of these educational studies suggest that ABA educational services can be maintained when transitioning from in-person to remote delivery (Oblak, 2021; Pollard et al., 2021).

Oblak (2021) and Pollard et al. (2021) both provide strong correlational evidence of the effectiveness of ABA instruction when making the transition from in-person to remote service delivery (Oblak, 2021; Pollard et al., 2021). However, their studies were not experimental. Oblak simply measured client outcomes prior to and following the transition to remote instruction due to COVID-19 lockdowns. Pollard et al. likewise measured client outcomes prior to and following lockdown and also reported separately on clients who received technician-delivered and caregiver-implemented remote instruction. With technician-delivered remote instruction, instruction was delivered exclusively by a qualified ABA therapist. With caregiver-implemented remote instruction, a qualified ABA therapist delivered instructions that provided appropriate consequences and were aided by a caregiver who provided prompts and consequences according to the ABA therapist. Although promising, the available literature base would benefit from experimental work examining the effects of instructional delivery modality to aid clinicians in making determinations about the appropriateness of in-person versus remote instruction. In addition, given that the preponderance of ABA telehealth/remote literature focuses on behavior reduction, additional research is needed to clarify the effects on learning novel academic responses.

In this study, we sought to extend the literature by conducting experimental comparisons of in-person and remote instruction. To further evaluate these modalities of instruction, we held constant as many instructional variables as we could to isolate the effects of delivering instruction in-person and remotely. We controlled for experimenter antecedents, target behaviors, consequences, and delivery of instructional materials on a computer screen. The difference between in-person and remote instruction was whether the experimenter was physically present next to the participant or could be heard through an online platform while the participant was at home. We took advantage of a hybrid instruction schedule during Fall 2020 whereby children attended school in-person for 1 week and remotely for a 2nd week throughout the school semester, which created a naturalistic reversal design. We considered in-person instruction as the baseline condition by which to compare remote instruction. We asked the following questions: (1) Will the rate of learning differ in a remote setting as compared to in-person? (2) Will the rate of trial presentations differ in a remote setting as compared to in-person? (3) Will there be a difference in 14-day and 21-day maintenance measures for target responses in a remote setting as compared to in-person?

Method

Participants

Six preschool-aged boys participated in this study. We selected these participants for three reasons: (1) Each child was participating in a hybrid instructional format across in-person and remote instruction, which was amenable to this study; (2) Each participant had the self-management prerequisite to remain seated and respond throughout a 30-min instructional session delivered over teleconference software; and (3) Each participant's parent committed to completing at least one session per day of virtual instruction during the respective virtual-instruction weeks.

All participants had educational classifications as preschoolers with a disability and had individualized education plans (IEPs). The specific disability diagnosis of each participant was not identified. In the state where the study was conducted, educational classifications gained children access to intensive ABA-based education. In terms of verbal behavior developmental cusps (Greer & Ross, 2008) all participants readily attended to instructors' faces, instructors' voices/directions, and 2D and 3D stimuli according to the Early Learner Curriculum and Achievement Record (Greer et al., 2019). Further, all participants demonstrated advanced listener literacy (i.e., responded to at least 20 vocal directions without the aid of any visual cues), and had independent tact and independent mand cusps (i.e., used vocal

verbal behavior to mediate their environment for social [tact] or tangible [mand] functions) in repertoire. All participants attended classrooms that operated using the CABAS® (Greer, 2002) system of instruction. This model uses the principles and tactics of ABA to guide instructional methods and decisions.

Dente was a 4-year 1-month-old boy and was a fluent listener (i.e., responded as a listener to at least 20 spoken instructions) and speaker (i.e., emitted at least 50 tacts and/or mands) repertoires. He demonstrated incidental bidirectional naming (Inc-BiN) level of verbal behavior, meaning his listener and speaker repertoires were joined and was able to learn word-object relations as a speaker incidentally without direct reinforcement provided by the experimenter. Dente received a full-scale score of 67 and a verbal comprehension score of 76 on the *Wechsler Preschool and Primary Scale of Intelligence* (WPPSI)-IV.

David was a 4-year 8-month-old boy who was a fluent listener (i.e., responded as a listener to at least 20 spoken instructions) and speaker (i.e., emitted at least 50 tacts and/or mands) repertoires and demonstrated Inc-BiN. David received a cognitive score of 96, and communication score of 80 (81 for expressive language and 78 for receptive language) on the *Developmental Assessment of Young Children* scale.

Jales was a 4-year 4-month-old-boy who was a fluent listener (i.e., responded as a listener to at least 20 spoken instructions) and speaker (i.e., emitted at least 50 tacts and/or mands) repertoires. He demonstrated unidirectional naming (Inc-UniN) level of verbal behavior, meaning that the listener and speaker repertoires were not joined. He acquired novel listener responses incidentally, but did not acquire novel speaker responses incidentally. Jales received a full-scale score of 77 and a verbal comprehension score of 71 on the WPPSI-IV.

Nat was a 4-year 6-month-old boy who was a fluent listener (i.e., responded as a listener to at least 20 spoken instructions) and speaker (i.e., emitted at least 50 tacts and/or mands) repertoires. Nat demonstrated Inc-BiN. Nat received a cognitive score of 65, an expressive language score of 74, and receptive language score of 67 on the *Developmental Assessment of Young Children* scale.

Clement was a 3-year 4-month-old boy and was a fluent listener (i.e., responded as a listener to at least 20 spoken instructions) and speaker (i.e., emitted at least 50 tacts and/or mands) repertoires. Clement demonstrated Inc-UniN. Clement received an auditory comprehension score of 73, expressive comprehension score of 71, and total language score of 70 on the *Preschool Language Scale*.

Mike was a 4-year 6-month-old boy and was a fluent listener (i.e., responded as a listener to at least 20 spoken instructions) and speaker (i.e., emitted at least 50 tacts and/or mands) repertoires. Mike demonstrated no degree

of naming, meaning he did not learn listener or speaker responses incidentally and required direct instruction and consequences to learn listener and speaker responses. Although Mike emitted a variety of mands and tacts in the instructional setting, he frequently needed prompts to vocally mand in a noninstructional setting. Mike received a full-scale score of 63 and a verbal comprehension score of 68 on the WPPSI-IV.

All standardized assessments were performed and reported by licensed school psychologists on behalf of the students' school districts.

Setting

The instructional sessions in this experiment were divided across two locations. All in-person instruction was delivered in the students' classroom in the school whereas remote instruction was delivered while the participant was at home. During in-person instructional sessions the experimenter delivered instruction to the student while they were both seated on chairs at a table. Throughout in-person instruction, there were between one and four students present in the room and three or four teachers delivering instruction to other students either face-to-face or remotely over a computer. The experimenters were the respective teachers for each participant and remained the same for each student across in-person and remote instruction.

During remote instruction, the experimenter presented instruction from the classroom (or home) through a teleconferencing software (i.e., Zoom) whereas the participant participated remotely with a caregiver present throughout the session to provide support and redirection, if needed. The caregiver delivered prompts, corrections, or reinforcement based on the guidance of the experimenter (similar to caregiver-implemented instruction in Pollard et al., 2021). For all participants, in the remote setting, the caregiver present to facilitate instruction was an adult (e.g., parent or grandparent), other than Nat, who had an older sister (approximately 13 years of age) who helped facilitate instruction while the parent was present in the house but was not at Nat's side during instruction.

Although there were some minor differences in the home environment across remote sessions (e.g., other siblings being home, changing of participant's location from one room in their house to another), this was not different from similar changes in the school setting (e.g., number of classmates being in the room varied on a given day, particular location in the classroom where the participant received instruction might vary across days) and thus it was not noted.

During remote instruction, each student received between one and three sessions of instruction a week for a duration of 30 min each. These times were established at the beginning of each week depending on the participant's and their

caregiver's schedule. Once established, these times for instruction were held constant throughout the week. During in-person instruction, the participant was in the school setting between 9:00 AM and 2:00 PM and received instruction throughout the day. Across both settings, no specific recurring time was set for conducting instruction included in this experiment, but rather the experimental instruction was interspersed as part of the participant's overall instruction and the total number of learning opportunities was yoked across in-person and remote settings.

Materials

Presentation of instructional materials was identical for in-person and remote sessions. Experimenters prepared instructional materials on PowerPoint presentations prior to the session and presented the materials over the computer screen during both in-person and remote lessons. The stimuli used for tact instruction included pictures that covered between 50% and 75% of the PowerPoint slide. For each individual tact, the experimenter presented four different exemplars of the stimulus (e.g., four pictures of a dog). The stimuli used for sight word instruction varied in font, color, and size within each set and this was held constant across conditions. Font sizes varied between 16- and 30-point font. During in-person sessions, the materials in this study were presented on the screen of a laptop computer. During remote learning, sessions were conducted using the Zoom teleconferencing software without confirming which type of device was being used by the participant in the home setting.

The experimenter implemented a token economy across conditions. During remote sessions, the experimenter created a digital token board that was displayed on the computer screen. During in-person instruction, experimenters used physical tokens and token boards. Backup reinforcers were the same across conditions and included access to watching preferred videos over the computer screen, access to an iPad, edible reinforcement (e.g., cookies, juice, pancakes), preferred toys (e.g., playdough or toy farm animals) and vocal praise. Aside from vocal praise and videos played over teleconference, backup reinforcers were delivered by the caregivers in the home setting.

Dependent Variable and Measurement

Experimenters measured responses to learn unit (LU) instruction (Albers & Greer, 1991). In LU instruction, the experimenter gains the participant's attention and delivers a clear unambiguous antecedent. After the participant's response the experimenter delivers a consequence. For correct responses, the experimenter delivered praise and reinforcement operations. For incorrect responses,

the experimenter provided an error correction procedure (described below).

The experimenters measured accuracy during LU instruction per operant (Wong et al., 2021). A correct response was scored when the participant emitted the target response within 3 s of the experimenter presenting the antecedent. An incorrect response was scored if the participant did not respond within 3 s of the experimenter presenting the antecedent or emitted a response other than the target response. The acquisition criterion (sometimes referred to as "mastery criterion") was seven consecutive correct responses and we counted the number of operants that met the acquisition criterion in a given instructional week (operant analysis; Wong et al., 2021; Wong & Fienup, 2022). We also report the number of LUs required for operants to reach criterion, per condition. During 14- and 21-day maintenance assessments, we report percentage correct responding to operants that met the acquisition criterion by dividing the number of correct responses by the total number of trials multiplied by 100.

Responses to LU instruction were the basis for calculating dependent variables. The dependent variables were the average number of learn units a participant required to master an operant, the average rate (per minute) of learn unit presentation, and the percentage of correct responses to maintenance probes for mastered operants conducted 14 and 21 days following mastery. Criterion for mastery was set at seven consecutive correct responses. Thus, the cumulative number of operants mastered in each condition was reported as the number of operants for which the participant emitted seven consecutive correct responses in each week.

The rate of learning was measured and reported as the total number of learn units delivered within a condition divided by the number of operants where the participant met the acquisition criterion. The rate of learn unit presentation was measured by having the experimenter start a timer as they delivered the first learn unit of a session and turning off the timer upon completion of the final learn unit of a session. The time of the session was divided by the number of learn units delivered to obtain the number of learn units per min.

Procedure

This study was conducted during a portion of ongoing educational services provided to the participants that aligned with IEP goals. We implemented a series of controls across the method and distribution of how instruction was delivered as well as how targets were selected. The purpose of these controls was to hold all variables constant across in-person and remote modalities aside from the modality in which instruction was delivered. This included having the same experimenter work with a single child during both in-person and remote teaching sessions. What follows is an overview of variables controlled for during this experiment.

Precomparison Control Procedures

Controlling Instruction across Weeks We held constant instructional variables across 2-week comparison time periods. For each direct comparison of in-person and remote instruction, we controlled variables across a specific 2-week period and not necessarily across multiple 2-week periods. To ensure equal distribution of instruction, experimenters delivered instructional sessions for targets included in this experiment daily on the maximum number of days or sessions that could be conducted equally across both week in the 2-week comparison. As a rule, sessions were delivered on either 3 or 4 days per week as variables such as instructional days on the school calendar or participant availability needed to be accounted for across a 2-week comparison. If school was in session on days when experimental instructional targets were not taught, other instructional programming was delivered in its place and was not included in this experiment.

Once the number of instructional days was equated across a 2-week comparison, we set a target learn unit goal per session. This number was held constant across each session within that 2-week comparison (i.e., across a week of remote and a week of in-person instruction). This determined the total LU each participant received per week in instructional modality (e.g., if a participant received 21 LU per session across 4 sessions (84 LU total) in 1 week of remote instruction, they would receive the same distribution of instruction in the week of the in-person instructional condition. Throughout instructional sessions in a 2-week comparison, the total number of independent learn units was held constant across conditions. All instruction was delivered on a computer screen—both during in-person and remote instruction.

Controlling Targets across 2-week Comparisons After the distribution of instruction (how many instructional days and how many LU per session) was established, experimenters selected target operants to be taught. The potential target operants to be taught were selected by assembling a list of potential target tacts and sight words limited by a range of parameters including total number of syllables per word, first letter/sound of the word, and category of tact (e.g., actions, animal, food item). Probe trials were then conducted to determine if the participant already had the target response in their repertoire (description of specific procedures are provided below). If the participant emitted a correct response on any of the probe trials, we removed those stimuli from the list of potential target operants. If the participant did not emit any correct responses to any of the probe trials, then target was left in the list. After determining which target operants the participants did not have in their repertoire, the remaining targets were placed into two groups so that each group

had potential targets with equivalent parameters as described above. These groups of targets were then randomly assigned to one of the instructional conditions, either remote or in-person (Cariveau et al., 2021).

Target sets were assigned using the logistical analysis method (Cariveau et al., 2021; Wolery et al., 2014). That is, the experimenters equated operants based on the number of syllables in the target responses and targets that were phonetically or visually similar were not included in the same sets. Sight words were selected from the same Dolch sight word lists across comparisons.

Instructional Procedures

Preexperimental Screening After assembling the list of potential targets, the experimenter conducted unconsequated probes for each participant to determine which responses were not yet in the participant's repertoire by presenting each stimulus three times. Probe trials consisted of the experimenter presenting the target stimulus on the screen and asking, "What is this?" The experimenter did not provide accuracy feedback to the participant; however, the experimenter praised related attending behavior.

Learn Unit Instruction During learn unit instruction, the experimenter established attending behavior by saying the participant's name and presenting a preferred backup reinforcer to establish the motivating conditions to attend to the instructor's antecedent. When the participant oriented toward the experimenter, the experimenter started the timer and delivered the first learn unit. In the learn unit, the experimenter presented the target stimuli on a computer screen (either a picture or a sight word) and presented the vocal antecedent "What is this?" and waited up to 3 s for the child to respond. If the child responded correctly, the experimenter delivered vocal praise or a token (depending on the participant and their history of consequences). If the child emitted an incorrect response, the experimenter conducted a correction procedure that consisted of modeling the correct response and then re-presenting the antecedent in order to allow for the participant to have an independent opportunity to respond to the antecedent. If the participant responded correctly following the correction, the experimenter continued to the next trial. If the participant responded incorrectly or did not respond within 3 s, the experimenter repeated the correction procedure up to three times before continuing to the next trial. Experimenters did not praise or deliver tokens for independent responses following a correction procedure.

When introducing novel targets, the experimenter provided two or three prompted responses before delivering learn units for independent responses. The type of prompt provided by the experimenter varied depending on the level

of verbal behavior of the participant. The experimenter modeled object-names for Nat, Dente, and David prior to presenting trials with an independent response requirement. This was based on previous research demonstrating children with Inc-BiN benefit from such models (Hran-chuk et al., 2019). Experimenters presented three echoics as a response prompt (Billingsley & Romer, 1983) prior to introducing an independent response requirement for Jales, Clement, and Mike.

Each session consisted of 21 trials across three different operants. Thus, experimenters presented each operant for seven trials within a session. Within each session, learn units were rotated across operants so that the same target was not presented two times consecutively. Learn unit instruction was delivered for each target operant until the participant's responding met the acquisition criterion set at seven consecutive responses. Seven consecutive responses were selected as acquisition criterion as that was the maximum number of presentations of a single operant in an instructional session and would thus represent the equivalent of 100% accuracy in a single session (Fuller & Fienup, 2018; Wong et al., 2021; Wong & Fienup, 2022). The acquisition criterion of seven consecutive responses was scored whether the consecutive response occurred within or across sessions. Once a single target response met the acquisition criterion, that target was removed from the session and the next available novel target was introduced to the session.

For all participants, three different tact operants were taught in rotation in each session. For Nat, an additional four sight words were taught in rotation in each session. Nat was the only participant to receive sight word instruction because he was the only participant with sight words included as part of his regular curricular objectives. Thus, Nat had an additional 21 LU of sight word instruction in addition to tact instruction. Data were recorded separately for each potential operant. When a participant emitted seven consecutive independent correct responses (within a session or across sessions) to a target operant that target was scored as mastered and removed from the rotation. A new target was then entered into the rotation of targets to be taught as outlined by Wong et al. (2021).

In-person instruction. All instruction for participants was delivered by the same experimenter that delivered instruction to the participant throughout the school year. In the in-person instructional session the experimenter sat with the participant at a table in the child's classroom and after establishing that a backup reinforcer was in place delivered learn unit instruction until the participant completed the predetermined number of independent learn units for the session.

Remote instruction. During remote instruction, the experimenter delivered instruction from the classroom or home¹ whereas the participant participated from a home setting with a caregiver present to help facilitate the lesson. The caregiver helped in redirecting the participant if they needed redirection to attend to the instructor and to deliver backup reinforcers when the child earned them. Instruction was led by the experimenter with support provided by the caregiver as described above.

Maintenance and Generalization Follow-up probes were conducted 14 and 21 days after the participants mastered objectives. The 14-day maintenance probe was conducted in the same setting as the original operants were mastered, whereas the 21-day maintenance and stimulus generalization probes was conducted in the opposite setting (generalization across modalities). Thus, for example, if an operant was mastered during the in-person condition, the 14-day maintenance probe was conducted in an in-person setting whereas the 21-day probe was conducted in a remote setting.

Experimental Design

We used a naturalistic reversal design (Cooper et al., 2007) to test the efficacy of remote instruction to in-person instruction across all dependent variables. The comparison controls described above were held constant across each 2-week period but not across separate 2-week periods due to differences in schedules, number of LU, or number of days that instruction was delivered across different comparisons. In this design, after ensuring that the target operants were not in the participant's repertoire, the experimenter equated targets so that they would be comparable and taught across in-person or remote instruction conditions. The experimenter assigned target sets using the logistical analysis method (Cariveau et al., 2021; Wolery et al., 2014). Instruction occurred in one context for 3 or 4 days and then in the other context for the same number of days—according to the hybrid educational schedule for the school. The number of days in which the participants received intervention was yoked across each 2-week period such that if there were only 3 days available for an in-person week, we delivered intervention on 3 days of the remote week whereas other instructional objectives not included in this analysis were taught on the other days. The experimenter rotated the modality of learn unit instruction (in-person or remote) each week as

¹ Remote instruction was delivered by the experimenter from the school setting Monday–Thursday. On Fridays the school building was closed so instruction was delivered from the teacher's homes.

part of the school's preexisting schedule.² At the end of each week, the participant's data were collected and included for analysis in this experiment. The number of days on which learn unit instruction was delivered as well as the number of experimental trials was held identical across comparison conditions. Conditions switched at the end of each week as the participant changed their modality of instruction (in-person or remote) based on the school's schedule. It is important to note that although only data collected during the predetermined number of days per week was included for analysis, for ethical considerations any target that was not mastered during the experiment was taught outside of the experiment during standard instruction.

Interobserver Agreement and Treatment Fidelity

Trial-by-trial interobserver agreement (IOA) was conducted throughout this experiment. IOA was collected by the supervisor as part of the Teacher Performance Rate and Accuracy (TPRA) procedure (Ingham & Greer, 1992) or by a trained independent observer who was a graduate student pursuing a master's degree in school psychology. The training of the observer consisted of conducting IOA together with the supervisor until the observer scored two sessions consecutively with 100% agreement. IOA was calculated by comparing the data recorded by the experimenter and that of the observer and dividing the number of agreed-upon trials by the total trials recorded and multiplying by 100%. Trial-by-trial IOA was collected for 36% of sessions for Dente with a 100% agreement, 25% of session for David with a mean agreement of 99% (range: 95%–100%), 25% of sessions for Jales with a mean agreement of 98% (range: 95%–100%), 27% of sessions for Nat with a mean agreement of 97% (range: 96%–100%), 27% of sessions for Clement with a mean agreement of 100%, and 43% of sessions for Mike with a mean agreement of 99% (range: 95%–100%).

An independent observer used the TPRA assessment to evaluate treatment fidelity. Treatment fidelity for the instructor delivering instruction was measured by the supervisor by evaluating the accuracy of each antecedent and consequence presentation made by the experimenter. The percentage of

fidelity was obtained by calculating the number of correct components delivered and dividing that number by the total number of components in the session and multiplying by 100%. Treatment fidelity was conducted for 14% of sessions with 100% fidelity for Dente, 16% of sessions with 98% fidelity (range: 95%–100%) for David, 10% of sessions for Jales with a 100% fidelity, 9% of sessions for Nat with a mean of 97% fidelity (range: 96%–97%), 27% of sessions for Clement with a mean of 100% fidelity, and 21% of sessions for Mike with a mean of 93% fidelity (range: 82%–100%).

Results

To establish a benchmark for determining differences between conditions, we adopted a standard of 20% difference. The 20% difference was selected as it is slightly more than one standard deviation above the mean of a normal distribution. In figures, we note differences that were 20% or greater with a *.

Number of Targets Meeting Criterion

Figure 1 displays the number of operants with criteria met for each participant per comparison for each participant per week for in-person (black bars) and remote (grey bars) instruction. Dente acquired at least 20% more operants during in-person instruction across both comparisons. David acquired at least 20% more operants during in-person instruction for one of three comparisons while acquiring at least 20% more operants remotely in the other two of three comparisons. Jales acquired at least 20% more operants during in-person instruction across both comparisons. Nat showed less than a 20% difference in number of operants learned across both comparisons. Clement acquired at least 20% more operants during in-person instruction across both comparisons. Finally, Mike acquired at least 20% more operants remotely across both comparisons.

Thus, three participants learned at least 20% more operants during in-person instruction across all comparisons, one participant learned at least 20% more operants remotely across all comparisons, one participant showed no difference, and one participant learned 20% more operants during one in-person comparison and learned 20% more operants during the remaining two remote comparisons.

Across all participants, participants acquired at least 20% more operants during in-person instruction for 7 of 13 (54%) comparisons, while acquiring 20% more operants during remote instruction in 4 of 13 (31%) comparisons with less than a 20% difference in the remaining 2 of 13 (15%) comparisons. Overall, during in-person instruction, participants acquired 20% more or an equivalent number

² Due to restriction in building capacity due to COVID-19, only half the student population was permitted to be present in the school building at a time. Therefore, half of the student body received in-person instruction each week while the remaining half received remote instruction. These groups then switched the following week with the students who had received remote instruction in Week 1 switching to in-person instruction during Week 2. On the other hand, the students who had received in-person instruction in Week 1 switched to remote instruction for Week 2. This rotation happened each week. In-person instruction was available for students Monday–Thursday whereas all students received virtual instruction on Fridays due to the school building being closed for disinfection.

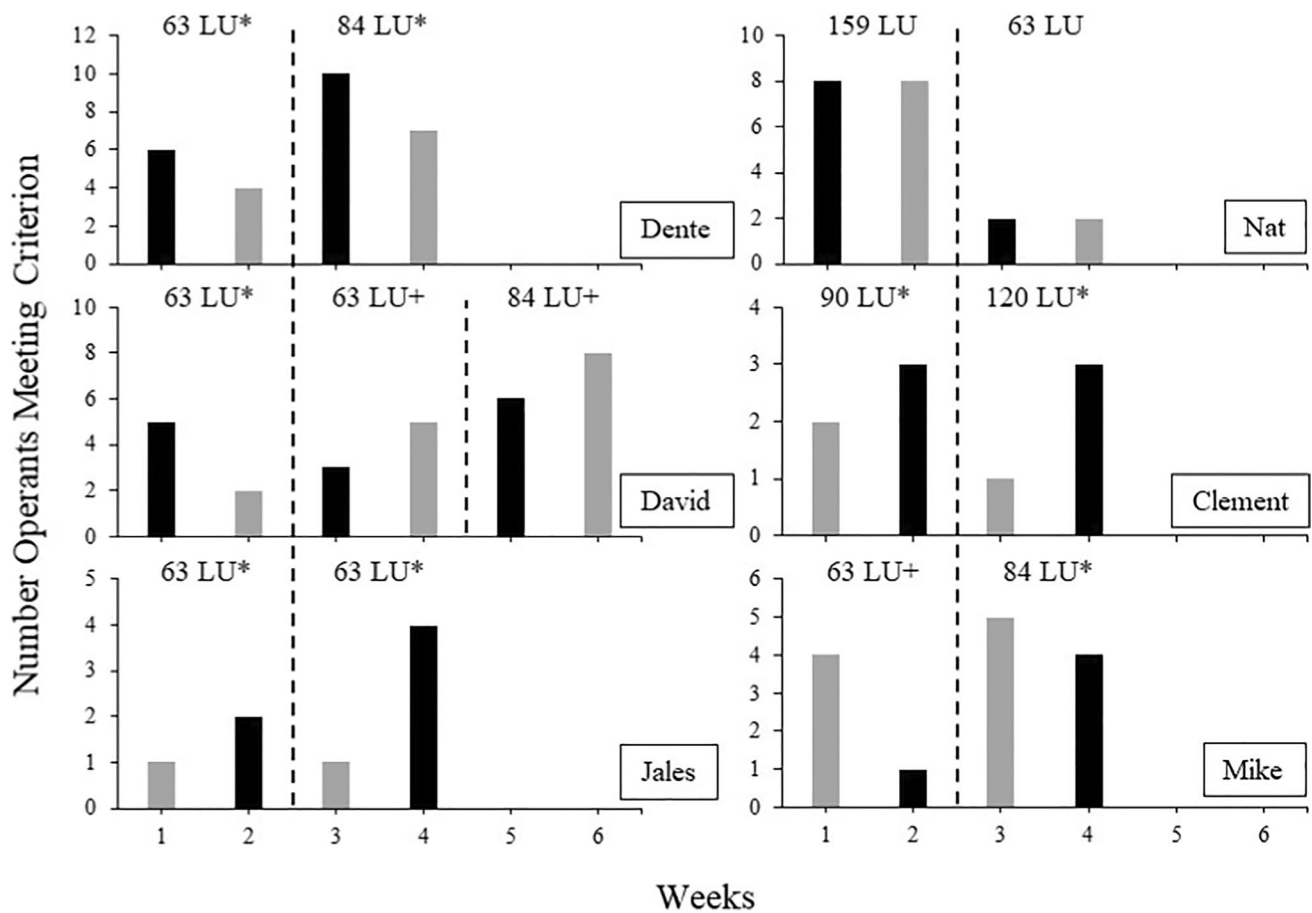


Fig. 1 Number of operants meeting criterion per comparison condition. Black bars represent in-person instruction while gray bars represent remote instruction. An asterisk (*) on top of the condition indicates a difference of 20% or more during in-person instruction whereas a plus symbol (+) indicates 20% or more learning during

remote instruction. The condition labels indicate the total number of LU delivered during each week within that comparison. All variables were held constant during each comparison and is thus represented by the phase change lines

of operants for 69% of all comparisons when compared to remote instruction.

Rate of Learning

Figure 2 displays the rate of learning (i.e., mean number of LU delivered per operant meeting the acquisition criterion) for each participant per week for in-person (black bars) and remote (gray bars) instruction. Dente learned at least 20% faster (i.e., 20% LU or fewer LU per STO) during in-person instruction across both of his comparisons. David learned at least 20% faster for one of three comparisons while learning in-person and at least 20% faster during remote instruction in the other two comparisons. Jales learned at least 20% faster during in-person instruction across both of his comparisons. Nat's rate of learning across in-person and remote instruction was undifferentiated during both comparisons according to the 20% difference criterion. Clement learned at least 20% faster during

in-person instruction across both comparisons. Finally, Mike learned at least 20% faster during remote instruction in one comparison while showing less than a 20% difference in the other. Thus, three participants learned faster during in-person instruction across all comparisons, one participant learned faster during remote instruction across both comparisons, one participant did not show a difference in rate of learning or rate of instruction across either of two comparisons, and one participant showed a faster rate of learning in one in-person comparison while learning faster remotely in his other two comparisons.

In total, using a minimum of 20% as a measure of difference, participants learned faster during in-person instruction in 7 of 13 (54%) comparisons, while learning faster in 3 of 13 (23%) comparisons during remote instruction, and showing less than a 20% difference in 3 of 13 (23%) of comparisons. It should be noted that the rate of learning during remote instruction was 20% faster or equivalent to in-person instruction across 6 of 13 (46%) of comparisons.

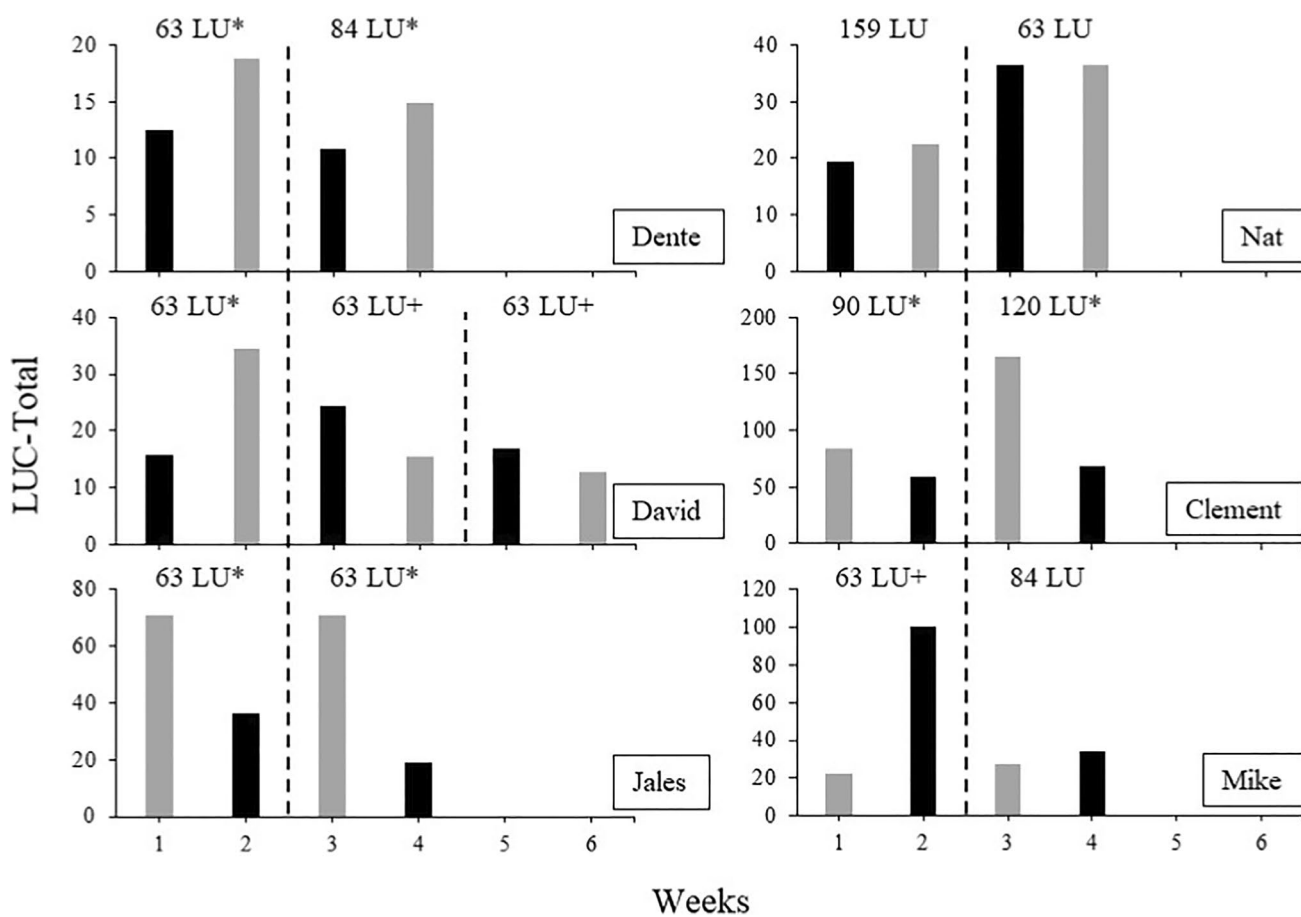


Fig. 2 Average number of learn units to meet a criterion per comparison condition. Black bars represent in-person instruction while gray bars represent remote. An asterisk (*) on top of the condition indicates a difference of 20% or more during in-person instruction while a plus symbol (+) indicates 20% or more learning during

remote instruction. The condition labels indicate the total number of LU delivered during each week within that comparison. All variables were held constant during each comparison and is thus represented by the phase change lines

Rate of LU Presentation

Figure 3 displays the rate of instructional completion measured as number of LU completed per min for each participant per week for in-person (black bars) and remote (grey bars) instruction. Dente completed all LU during instructional sessions at least 20% faster (i.e., 20% more LU per minute) during in-person instruction in one comparison while showing no difference in the other. David completed all LU during instructional sessions at least 20% faster during in-person instruction across one of three comparisons while completing instruction in the other two of three comparisons with less than a 20% difference. Jales completed all LU during instructional sessions at least 20% faster during in-person instruction across both comparisons. Nat completed all LU during instructional sessions at least 20% faster during in-person instruction for one comparison while showing less than 20% difference in the other. Clement completed all LU during instructional sessions at least 20% faster during

in-person instruction for one comparison while showing no difference in the other. Mike completed all LU during instructional sessions with less than a 20% difference across conditions in both comparisons.

In total, using a minimum of 20% as a measure of difference, participants completed all LU during in-person instruction faster in 6 of 13 (46%) comparisons while showing less than a 20% difference in the remaining 7 of 13 (54%) comparisons. Of note, no participants completed all LU during instructional sessions 20% faster remotely when compared to in-person across any comparisons.

Response Maintenance and Stimulus Generalization

Figure 4 displays the rate of the percentage of correct responses to follow-up probes conducted 14 days (dark bars) and 21 days (white bars) after meeting criterion for mastery of an operant for each participant per comparison. Dark bars (14-day maintenance) are differentiated across

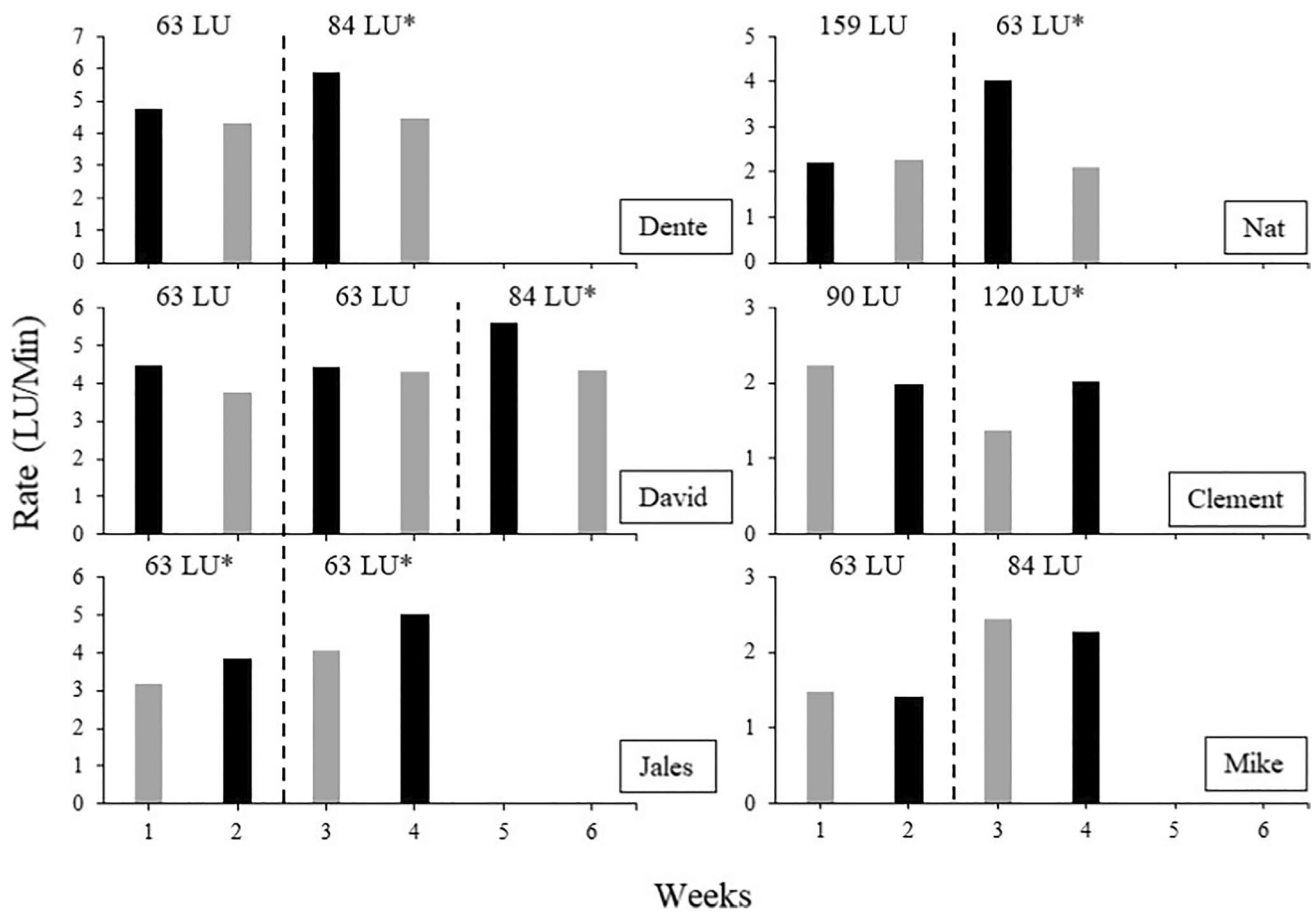


Fig. 3 Average number of learn units per minute across in-person and remote comparisons. Black bars represent in-person instruction while gray bars represent remote. An asterisk (*) on top of the condition indicates a difference of 20% or more during in-person instruction while a plus symbol (+) indicates 20% or more learning during

remote instruction. The condition labels indicate the total number of LU delivered during each week within that comparison. All variables were held constant during each comparison and is thus represented by the phase change lines

in person and remote sessions with maintenance following in-person instruction being represented as black bars and maintenance following remote instruction being represented as grey bars. Probes measured 14 days following meeting the acquisition criterion represent a measure of response maintenance (testing in same context as teaching) and 21-day probes represent a measure of maintenance and stimulus generalization (testing in other context than teaching). In the following section, we report each 14-day pair and each 21-day pair within a phase as a separate comparison.

To best interpret the Fig. 4, comparisons should be conducted between bars within the same phase lines as instructional conditions were held constant for all instruction within a phase line but not across phase lines. Within each phase, 14-day maintenance (solid fill) and 21-day maintenance and generalization (no fill) for targets mastered during in-person instruction are demarcated with black fill and outline whereas follow up data for 14-day maintenance and 21-day

maintenance and generalization are demarcated with grey fill and grey outline for the above measures respectively.

Several different comparisons can be studied in the follow up data. To compare response maintenance (14-day probe) to response maintenance and stimulus generalization (21-day probe) within targets mastered in the same modality, one would compare the solid fill bar to the white bar with an outline of the same color as the solid fill bar (e.g. a black bar [14-day] would be compared to a white bar with black outline [21-day] whereas a gray bar [14-day] would be compared to a white bar with a gray outline [21-day] within the same phase). Of note, white bars with a black outline are always placed immediately next to the black bars whereas white bars with a gray outline are always placed immediately next to the gray bars.

To compare response maintenance (14-day) measures across modalities, one would compare bars with solid fill within the same phase (e.g., black bar [in person] to a gray bar [remote]). To compare response maintenance along with

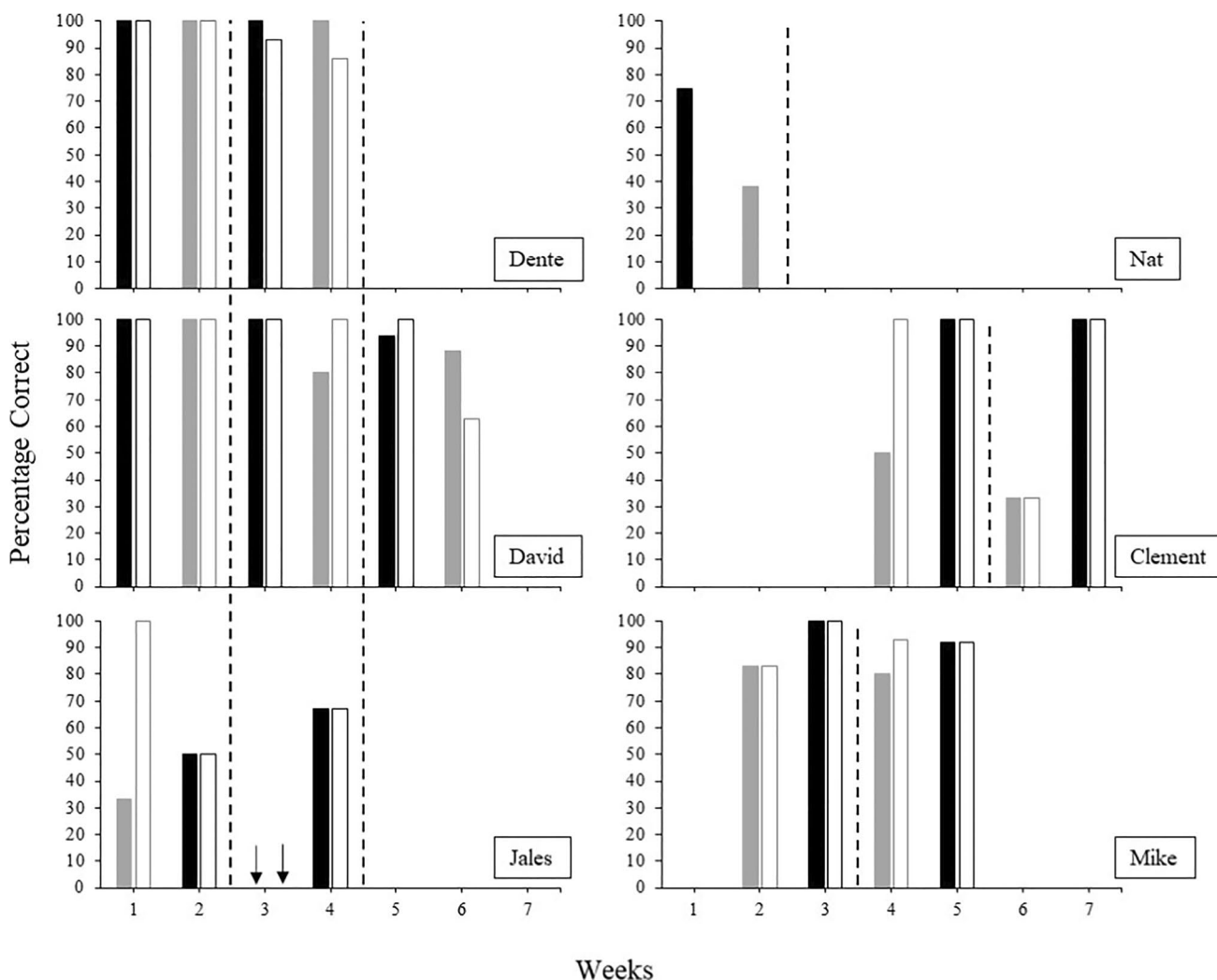


Fig. 4 Percent of correct responses for 14-day and 21-day follow-up probes. Dark bars represent follow-up from operants mastered during in-person (black) and remote (gray) instruction during the 14-day assessments (maintenance). White bars represent follow up probes

of operants during the 21-day (white bars) follow up probes (maintenance and generalization). The x-axis reflects the temporal order of weeks during intervention. Complete follow-up data were not available for Nat

stimulus generalization across modalities, one would compare white bars with black outline to the white bars with the gray outline within the same phase.

Dente demonstrated less than a 20% difference in correct responding in all follow-up probes across the four comparisons (14-day in first comparison/phase, 21-day in first comparison/phase, 14-day in second comparison/phase, 21-day in second comparison/phase). David responded with 20% greater accuracy in follow-up probes following in-person instruction for two of six comparisons (14-day in second comparison/phase and 21-day in third comparison/phase) while demonstrating no difference for the remaining four of six comparisons. Jales responded with 20% greater accuracy following remote instruction for one of four comparisons (21-day in first comparison/phase) while demonstrating 20% greater accuracy in two of four comparisons following

in-person instruction (14-day and 21-day in second comparison/phase), and no difference for the remaining comparison. Nat responded with 20% greater accuracy following in-person instruction during the 14-day maintenance assessment (only comparison with maintenance data). No additional maintenance and generalization data were collected due to lapses in attendance. Clement responded with 20% greater accuracy following in-person instruction for three of four comparisons while demonstrating less than a 20% difference for the remaining comparison (14-day in first comparison/phase). Mike demonstrated less than a 20% difference in correct responding in all follow-up probes.

In total, participants responded with 20% or more correct responses following in-person instruction for 8 of 23 (35%) comparisons, 20% more correct responses following remote instruction for 1 of 23 (5%) of comparisons, and less than a

20% difference across the remaining 14 of 23 (64%) comparisons. This indicates that in-person instruction resulted in an equal or greater number of correct responses across 96% of comparisons.

When subdividing the 23 comparisons to maintenance (14-day) and maintenance with generalization (21-day) probes, there were 20% more correct responses in 5 of 12 (42%) maintenance (14-day) probes following in-person instruction whereas there was no difference in the remaining 7 of 11 (64%) probes. For the 21-day (maintenance with generalization) probes, there were 20% more correct responses in 3 of 11 (27%) probes following in-person instruction, 1 of 11 (9%) probes following remote instruction, and no difference in the remaining 7 of 11 (64%) of probes. Overall, the outcomes showed comparable maintenance and generalization across in-person and remote instruction, with a small advantage for learning that occurred in-person.

Discussion

In order to test the effects of instructional modality on educational dependent variables, we used a naturalistic reversal design to teach a series of operants across the in-person and remote instructional modality while controlling many other confounding variables. Using a minimum difference of 20% to be considered different, the results of this experiment indicate that, when all else is held constant, completing instruction in-person resulted in acquiring more operants for 54% of comparisons, faster instructional delivery for 46% of comparisons and better maintenance for 35% of comparisons (42% of maintenance probes at 14 days, and 27% of maintenance with generalization probes at 21 days). This is compared to 20% greater learning in remote instruction for 31% of comparisons for number of acquired operants, 0% for rate of instructional delivery, and 5% of 14- and 21-day follow-up maintenance measures (the only difference was in one 21-day maintenance with generalization probe). The remainder of comparisons had showed less than a 20% difference.

The number of comparisons that resulted in near-equivalent learning outcomes (36% for operants mastered, 54% for rate of instructional completion, and 69% of maintenance measures) demonstrates that learning still occurs and remote instructional presentation should be considered as a viable approach where other factors support its consideration. Although this study's findings indicate that in-person instruction resulted in equal or better learning outcomes when compared to remote instruction in many cases, learning consistently did occur during remote instruction. The findings in our comparison between in-person and remote instructional delivery should help encourage further research to identify the conditions under which remote instruction

should be considered and selected if in-person instruction is not readily available.

These results add nuance to the findings of Oblak (2021) and Pollard et al. (2021) who reported on the effectiveness of remote instruction following the transition from in-person instruction due to the pandemic lockdown. In our experiment we looked at clinician delivered intervention targeting acquisition of novel academic responses. We found that although in-person instruction often resulted in more robust outcomes, approximately half of remote instructional comparisons resulted in near-equivalent outcomes. This supports the findings from the literature above that remote instructional delivery can be considered a viable alternative to in-person instruction when in-person instruction is not readily available.

The findings reported herein support existing literature reporting on the effectiveness of teaching in a remote modality (e.g., Hay-Hansen [2013] reporting on discrete trial training, and Pellegrino and DiGennaro Reed [2020] reporting on effective training using total-task chaining). However, our experimental results differ slightly from the findings of Pollard et al. (2021), who reported that there was a slight increase of correct independent responding across all targets during remote instruction compared to in-person instruction. Although our measures reported herein differ from the measures recorded by Pollard et al. because our measures imply a slight quicker learning outcomes to in-person instruction, the implication that the viability of remote instruction should be further investigated is shared by both reports.

The above studies provide some initial data suggesting that individuals with disabilities can benefit from remote service delivery over teleconference. In the conclusion of a systematic review of ABA delivered remotely, Unholz-Bowden et al. (2020) concluded that the limited evidence currently available seems to support the effectiveness of ABA service delivery remotely. The current study extends the above research in that it also demonstrates effectiveness of remote service delivery as do Ferguson et al. (2020) and Pellegrino and DiGennaro Reed (2020) but extends those findings by evaluating the relative effectiveness of remote instruction compared to identical instruction delivered in-person. With results indicating similar levels of correct responding on follow-up probes for objectives mastered during in-person and remote instruction (similar to the topic of analysis by Pollard et al., 2021), this study found that the participants maintained similar levels of correct responding demonstrated during in-person instruction when there is a transition of service delivery to remote. However, this study diverges from Pollard et al. (2021) in that it provides a detailed analysis of variables such as number of LU per minute, objectives mastered, and rate of learning for each participant whereas Pollard et al. (2021) reports on aggregate data across participants on overall levels of correct responding. Thus, it

can be argued that this study provides direct measures of operant acquisition or educational responses whereas the data reported by Unholz-Bowden et al. (2020) and Pollard et al. (2021) either emphasize therapy/reduction of problem behavior, or aggregated measures that do not provide detailed analysis of learning.

Finally, it is important to note that a faster rate of LU presentation has been related to improved participant learning (Ingham & Greer, 1992). There was less than a 20% difference in rate of LU presentations in 54% of comparisons and in the remaining 46% of comparisons, participants receiving instruction in-person demonstrated an increase of at least 20% in number of learn units completed per minute. Of note, that during our experiment, the LU delivered per minute was not 20% faster during remote instruction across a single comparison. Thus, during remote instruction, the number of LU delivered per minute was able to be maintained at a near-equivalent rate in approximately half of the comparisons compared to in-person instruction.

Implications for Practice

This experiment was conducted at a time when much of the world was utilizing remote instruction due to the COVID-19 pandemic. However, the implications of such a program of research reach beyond emergency pandemic measures and can inform service provisions in early childhood settings beyond pandemic times. Thus, the above research adds to the early research in supporting remote delivery of ABA services. If further research provides more data supporting similar findings as ours, this can provide stronger support for removing barriers for access to ABA services due to geographical distance. Further, as our findings demonstrate that substantial learning can occur in remote settings, this can inform service provisions in early childhood settings where training and support of staff can be accomplished by clinicians providing services remotely.

Despite the potential effectiveness of remote instruction, it is important to note that remote instruction can create new barriers as well. New barriers can include limitations relating to access to internet, fluency in using computer software needed for remote instruction, as well as limited access to necessary hardware may hamper some students from benefitting from remote instruction. Finally, because caregiver availability is needed to facilitate remote instruction, especially in the early childhood setting, this can pose a limitation for students who do not have available caregivers. Thus, if remote instruction were to be more widely adopted, it will be important to account for new barriers that may arise.

Limitations and Implications for Future Research

It is important to mention some considerations when interpreting the results from this comparison between in-person and remote instruction. One particular area to highlight is the effect of establishing operations and their influence on results from the home setting compared to an in-person setting. For example, during in-person instruction, the participant was in the classroom throughout a 5-hr window. During remote instruction, the timing of instructional sessions was fixed at the beginning of each week. This added flexibility in the in-person setting, might allow the experimenter to respond to possible establishing operations that might influence responding. For example, if the experimenter noticed the participant was hungry, they might first provide lunch before continuing instruction. In a similar situation during remote instruction, the caregiver in the home setting may try to complete the instructional session before providing lunch as the time for the instructional session is preestablished.

In addition, having a caregiver present throughout instruction in the remote setting may present a setting event that may alter the instructional control established by an experimenter in an in-person setting. Further, given that in-person instructional sessions took place in the participants' school setting, where a primary goal is to teach academic responding, it is likely that the physical environment of the instructional sessions may support academic responding more in the school setting compared to the home setting where formal academic responding is usually not encouraged as consistently.

There were some limitations to be considered when interpreting the results of this experiment. One area to consider is the characteristics of the participants which met criteria for inclusion. First, all participants were under instructional control and required minimal prompts in order to remain seated and attend to the instructional sessions. For some children, a period of time to establish instructional control for the child to participate throughout the duration of an instructional session may be needed before achieving results similar to those reported herein.

Further, it is important to emphasize that these findings were achieved with participants with similar levels of verbal behavior. All participants in this study emitted at least 25 independent tacts or mands and responded as listeners to at least 25 spoken instructions. Thus, this research should be extended to include participants who do not yet have similar levels of listener and speaker behaviors in their repertoire to test if the modality of instruction will relate to changes in educational outcomes. Further, the target operants that were taught in this study included tacts and whole-word sight words, which required, for these participants, relatively low levels of response-effort compared to other programs that

require greater response effort. An example of a more effortful educational response that may prove to be more challenging in a remote modality might include a multiword sentence or multistep math problem. It might be the case that target operants that require greater level of response effort might produce different outcomes than what was demonstrated here. Further research could address this limitation by replicating these effects across a wider variety of target operants.

Another limitation is that measures of caregiver responding (e.g., rate of approvals, prompting, involvement in instruction) were not included in this investigation. As the caregiver helped facilitate instruction during remote sessions, it is reasonable to assume that variables of caregiver behavior will affect the participant's learning (Kim & Fienup, 2022). In addition to identifying prerequisite skills for children to benefit from remote instruction, it would be meaningful to identify which caregiver repertoires are necessary to strengthen to better facilitate child learning.

Despite these limitations, these findings are significant in that they demonstrate with a tightly controlled experiment where all else is held constant, in-person instruction results in more robust learning outcomes more consistently than remote instruction whereas remote instruction also results in consistent positive learning outcomes. Replicating such an investigation on a wider scale will be important because it would improve external validity. Such data can help guide policy over remote instruction as well as helping identify potential prerequisites that may allow participants to benefit more from remote instruction.

Acknowledgments This study was conducted by the first author in partial fulfillment of a PhD in Applied Behavior Analysis at Teachers College Columbia University. We thank Dr. R. Douglas Greer, Dr. Jessica Dudek, Dr. Grant Gautreaux, and Dr. Laudan Jahromi for comments on an earlier version of this manuscript.

Funding This research was conducted as part of a doctoral dissertation for the first author. No direct funding was received for the completion of this experiment.

Declarations

Conflicts of Interest We have no known conflicts of interest to disclose

Ethics Approval This research was approved by the IRB board at Teachers College, Columbia University and the Fred Keller School.

Consent to Participate Written consent to participate and publish anonymous collected data was obtained from the parents of all participants.

Consent for Publication Written consent to participate and publish anonymous collected data was obtained from the parents of all participants.

References

- Albers, A. E., & Greer, R. D. (1991). Is the three-term contingency trial a predictor of effective instruction? *Journal of Behavioral Education, 1*(3), 337–354. <https://doi.org/10.1007/bf00947188>
- Barretto, A., Wacker, D. P., Harding, J., Lee, J., & Berg, W. K. (2006). Using telemedicine to conduct behavioral assessments. *Journal of Applied Behavior Analysis, 39*(3), 333–340. <https://doi.org/10.1901/jaba.2006.173-04>
- Billingsley, F. F., & Romer, L. T. (1983). Response prompting and the transfer of stimulus control: Methods, research, and a conceptual framework. *Journal of the Association for the Severely Handicapped, 8*(2), 3–12. <https://doi.org/10.1177/154079698300800201>
- Cariveau, T., Batchelder, S., Ball, S., & La Cruz Montilla, A. (2021). Review of methods to equate target sets in the adapted alternating treatments design. *Behavior Modification, 45*(5), 695–714. <https://doi.org/10.1177/0145445520903049>
- Cooper, J. O., Heron, T. E., & Heward, W. L. (2007). *Applied behavior analysis*. Pearson.
- Ferguson, J. L., Majeski, M. J., McEachin, J., Leaf, R., Cihon, J. H., & Leaf, J. B. (2020). Evaluating discrete trial teaching with instructive feedback delivered in a dyad arrangement via telehealth. *Journal of Applied Behavior Analysis, 53*(4), 1876–1888. <https://doi.org/10.1002/jaba.773>
- Fisher, W. W., Luczynski, K. C., Hood, S. A., Lesser, A. D., Machado, M. A., & Piazza, C. C. (2014). Preliminary findings of a randomized clinical trial of a virtual training program for applied behavior analysis technicians. *Research in Autism Spectrum Disorders, 8*(9), 1044–1054. <https://doi.org/10.1016/j.rasd.2014.05.002>
- Fronapfel, B. H., & Demchak, M. (2020). School's Out for COVID-19: 50 Ways BCBA trainees in special education settings can accrue independent fieldwork experience hours during the pandemic. *Behavior Analysis in Practice, 13*(2), 312–320. <https://doi.org/10.31234/osf.io/cr3uv>
- Fuller, J. L., & Fienup, D. M. (2018). A preliminary analysis of mastery criterion level: Effects on response maintenance. *Behavior Analysis in Practice, 11*(1), 1–8. <https://doi.org/10.1007/s40617-017-0201-0>
- Greer, R. D. (2002). *Designing teaching strategies: An applied behavior analysis systems approach*. Elsevier.
- Greer, R. D., & Ross, D. E. (2008). *Verbal behavior analysis: Inducing and expanding complex communication in children with severe language delays*. Allyn & Bacon.
- Greer, R. D., Speckman, J., Dudek, J., Cahill, C., Weber, J., Du, L., & Longano, J. (2019). *Early learner curriculum and achievement record (ELCAR): A CABAS® developmental inventory*. CABAS® and the Fred S. Keller School.
- Hay-Hansson, A. W., & Eldevik, S. (2013). Training discrete trials teaching skills using videoconference. *Research in Autism Spectrum Disorders, 7*(11), 1300–1309. <https://doi.org/10.1016/j.rasd.2013.07.022>
- Higgins, W. J., Luczynski, K. C., Carroll, R. A., Fisher, W. W., & Muford, O. C. (2017). Evaluation of a telehealth training package to remotely train staff to conduct a preference assessment. *Journal of Applied Behavior Analysis, 50*(2), 238–251. <https://doi.org/10.1002/jaba.370>
- Hranchuk, K., Greer, R. D., & Longano, J. (2019). Instructional demonstrations are more efficient than consequences alone for children with naming. *Analysis of Verbal Behavior, 35*(1), 1–20. <https://doi.org/10.1007/s40616-018-0095-0>
- Ingham, P., & Greer, R. D. (1992). Changes in student and teacher responses in observed and generalized settings as a function of supervisor observations. *Journal of Applied Behavior Analysis, 25*(1), 153–164. <https://doi.org/10.1901/jaba.1992.25-153>

- Kim, J. Y., & Fienup, D. M. (2022). Increasing access to online learning for students with disabilities during the COVID-19 pandemic. *Journal of Special Education, 55*(4), 213–221. <https://doi.org/10.1177/0022466921998067>
- Knowles, C., Massar, M., Raulston, T. J., & Machalicek, W. (2017). Telehealth consultation in a self-contained classroom for behavior: A pilot study. *Preventing School Failure: Alternative Education for Children and Youth, 61*(1), 28–38. <https://doi.org/10.1080/1045988x.2016.1167012>
- Lindgren, S., Wacker, D., Suess, A., Schieltz, K., Pelzel, K., Kopelman, T., Lee, T., Romani, P., & Waldron, D. (2016). Telehealth and autism: Treating challenging behavior at lower cost. *Pediatrics, 137*(Suppl. 2), S167–S175. <https://doi.org/10.1542/peds.2015-2851o>
- Little, L. M., Wallisch, A., Pope, E., & Dunn, W. (2018). Acceptability and cost comparison of a telehealth intervention for families of children with autism. *Infants & Young Children, 31*(4), 275–286. <https://doi.org/10.1097/iyc.000000000000126>
- Oblak, M. K. (2021). An analysis of a system under pandemic conditions. *Journal of Applied Behavior Analysis, 54*(2), 530–546. <https://doi.org/10.1002/jaba.836>
- Pellegrino, A. J., & DiGennaro Reed, F. D. (2020). Using telehealth to teach valued skills to adults with intellectual and developmental disabilities. *Journal of Applied Behavior Analysis, 53*(3), 1276–1289. <https://doi.org/10.1002/jaba.734>
- Pollard, J. S., LeBlanc, L. A., Griffin, C. A., & Baker, J. M. (2021). The effects of transition to technician-delivered telehealth ABA treatment during the COVID-19 crisis: A preliminary analysis. *Journal of Applied Behavior Analysis, 54*(1), 87–102. <https://doi.org/10.1002/jaba.803>
- Sutherland, R., Trembath, D., & Roberts, J. (2018). Telehealth and autism: A systematic search and review of the literature. *International Journal of Speech-Language Pathology, 20*(3), 324–336. <https://doi.org/10.1080/17549507.2018.1465123>
- Tomlinson, S. R., Gore, N., & McGill, P. (2018). Training individuals to implement applied behavior analytic procedures via telehealth: A systematic review of the literature. *Journal of Behavioral Education, 27*(2), 172–222. <https://doi.org/10.1007/s10864-018-9292-0>
- Unholz-Bowden, E., McComas, J. J., McMaster, K. L., Girtler, S. N., Kolb, R. L., & Shipchandler, A. (2020). Caregiver training via telehealth on behavioral procedures: A systematic review. *Journal of Behavioral Education, 29*(2), 246–281. <https://doi.org/10.1007/s10864-020-09381-7>
- Wolery, M., Gast, D. L., & Ledford, J. R. (2014). Comparison designs. In D. L. Gast & J. R. Ledford (Eds.), *Single case research methodology* (pp. 297–345). Routledge. <https://doi.org/10.4324/9780203521892-12>
- Wong, K. K., & Fienup, D. M. (2022). Units of analysis in acquisition-performance criteria for “mastery”: A systematic replication. *Journal of Applied Behavior Analysis*. Advance online publication. <https://doi.org/10.1002/jaba.915>
- Wong, K. K., Bajwa, T., & Fienup, D. M. (2021). The application of mastery criterion to individual operants and the effects on acquisition and maintenance of responses. *Journal of Behavioral Education, 1*–23. <https://doi.org/10.1007/s10864-020-09420-3>

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.