Descriptive and Experimental Analyses of In-person and Remote Instruction

Joseph M. Peysin

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2021

Abstract

Descriptive and Experimental Analyses of In-Person and Remote Instruction

Joseph M. Peysin

Due to the COVID-19 pandemic many schools were forced to interrupt in-person delivery of educational services and switched to delivery of instruction in a remote setting. The educational impacts of school closures and remote instructional delivery have become a concern for the impact on an entire generation of students. Although delivery of behavioral interventions remotely is a topic that has been reported on in behavior analytic literature for over 15 years before this pandemic, few studies directly compared delivery of equivalent services across in-person and remote settings. Further, no studies included in recent literature reviews included comparative analyses between acquisition of novel instructional objectives across in-person and remote settings. Additionally, no studies are reported using verbal behavior developmental measures to identify potential prerequisites for benefitting from remote instruction. I therefore present a series of experiments to investigate the relative effectiveness of in-person and remote instruction as well as comparing outcomes for students grouped by level of verbal behavior development. In Experiment I, I conducted a carefully controlled experiment to compare rate of learning, rate of instructional presentation, and maintenance of objectives mastered across in-person and remote settings. I used a reversal design across 6 preschool aged participants with disabilities. The results indicate that some participants reliably mastered objectives and completed instruction faster in-person for 3 of 6 participants while the results for the other 3 participants were mixed. Overall, participants mastered objectives and completed instruction faster in-person in approximately half of the comparisons while showing no difference or learning faster remotely in the remaining comparisons. No consistent difference was shown in 14 and 21-day follow up maintenance measures. No consistent difference in the outcomes of students who demonstrated Naming compared to those who did not demonstrate any Naming. In Experiment II, I extend the findings of the previous experiments by comparing system wide educational outcomes of a hybrid in-person and remote educational model to the outcomes of the same model in a pure in-person setting in terms of fidelity of instruction, educational outcomes, and a cost analysis to determine how much the transition to remote provision of instruction costs stakeholders. Further, I compared educational outcomes across students categorized by level of verbal behavior development. The results indicate that the total number of learning opportunities and objectives mastered are significantly higher during a fully in-person model when compared to a hybrid educational model containing a remote instructional component. Further, when comparing educational outcomes across groups of students categorized by level of verbal behavior development, the results indicate that the rate of learning and objectives masted are significantly increased once students demonstrate joining of the listener and speaker repertoires as indicated by the presence of the Naming capability. The implication of the results are discussed in terms of feasibility of remote instruction as an alternative to in-person instruction as well as the importance of identifying and establishing the Naming capability for students to best benefit from remote instruction.

Keywords: Telehealth, COVID-19, remote instruction, virtual instruction, verbal behavior development

Table of Contents

List of Tables

List of Figures

Acknowledgements

It truly takes a village. This milestone achievement is a celebration of the dozens of individuals who have made this possible. Thank you Chana for supporting me through my long absences. Thank you to my sponsor, Dr. Fienup for guiding me, mentoring me, and keeping me human throughout this past year. Thank you to my dissertation committee members: Dr. Greer, Dr. Dudek, Dr. Jahromi, and Dr. Gautreaux for generously giving of your insight, time, and wisdom. Thank you to those who helped me gather the enormous amount of data and individual student responses which are the center of this dissertation: Dr. Jeanne Speckman, Dr. Stavra Romas, Ms. Brenda and Jashley. Thank you Dr. Yitzhak Chakiris who, after using *Designing Teaching Strategies* to guide instructional design at a school in Brooklyn sent me to learn from the author himself. Thank you to the many teacher mentors who have taught me and made me a better scientist and teacher: Lenah Alshowaiman, Brittany Bly, Ginger Harms, and Kate Loomis. Thank you Robin Nuzzolo for the work you do to so that Fred S Keller school can continue changing lives of all of its students - the children students as well as the adult students who come to teach and don't stop learning. Thank you to the many Masters students who have graciously learned with me and taught me so much: Patricia, Song, Sonja, Lauren, Jashley, Cass, and Danielle. Finally, thank you to all of my students from the past five years who have taught me all I know what it means to teach.

Dedication

This paper is dedicated to the Lubavitcher Rebbe, Rabbi Menachem Mendel Schneerson, who brought me to work in the field of education.

Chapter I:

Introduction

Across various disciplines, clinicians have used telehealth as a service delivery model for over 50 years (American Telemedicine Association, 2020). The term telehealth is used to describe a variety of service-delivery forms including use of communication technology to enhance health care, public health, and health education (Center for Connected Health Policy, 2020). As it relates to ABA, live videoconferencing has been the most common form of telehealth for use by providers to provide consultation for direct assessment and treatment of children emitting problem behaviors (Tomlinson et al. 2018).

In applied behavior analysis (ABA), remote service delivery has been reported on as early as 2006 when Barreto et al. provided two case examples illustrating the use of remote service delivery to train caregivers in a center to conduct brief functional analyses (FA) for children with severe problem behavior. The participants included a 5-year-old boy with autism and a 1-year-old girl who were referred for treatment due to disruptive behavior, noncompliance, object destruction, and severe self-injury. The experimenters trained the local caregivers to conduct a brief FA for each participant. The functional analysis helped identify the escape functions of the target behaviors for both participants – which were different functions than had been hypothesized by the local teams. Although the experiment reported on the successful implementation of a brief FA, no data were reported on any subsequent intervention of reduction of problem behavior.

Developments in Remote Service Delivery

The need for effective telehealth-delivered educational services was highlighted in March of 2020 as the COVID-19 pandemic led to governmental orders the closure of schools in all 50 states in America (Fronapfel & Demchak, 2020). As there are widespread school closures, there is a greater emphasis on comprehensive, provider-delivered, remote ABA service delivery. As service providers across the nation are being called on to deliver ABA services altogether remotely, strategies are needed for comprehensive application of remote ABA services – including problem behavior reduction, training of caregivers, and direct instruction in acquisition of new behaviors. This differs to what literature in remote ABA delivery has emphasized historically.

Research on the remote service delivery in ABA was initially conducted to evaluate means to deliver services to individuals living in geographically distant areas where in-person services were not available (Higgins et al. 2017). For example, the seminal article by Barretto and colleagues (2006) reported service delivery to children who lived in rural areas of Iowa without local ABA services available. Similarly, Fischer et al. (2014) reported the effects of a virtual training course for ABA technicians delivered to behavioral technicians employed at military bases throughout the United States – including bases in areas without local ABA services available.

Remote service delivery has proliferated with increased internet and technological access. Along with increased access to the relevant technologies, an increase in scholarly investigation of telehealth's effectiveness has appeared and as of 2016, 86% of publications relating to telehealth had been published in the preceding 15 years (Wacker et al. 2016). At the time of this writing, due to school closures and inclusion of remote instruction nationwide, a large expansion of remote service delivery of ABA is occurring (Fronapfel & Demchak, 2020). There is an exponential increase in the utilization of remote service delivery and despite limited data supporting the effectiveness of remote ABA service delivery. Notably, some government funded insurance plans have allowed for reimbursement of remote ABA service delivery due to the national health crisis (e.g. South Carolina Medicaid, 2020). More recently, a systematic review identified that although there was not sufficient research available for conclusive evidence, the available evidence suggests that remote service delivery of ABA technologies can produce similar results to direct service delivery (Unholz-Bowden et al. 2020)

Current Needs in Telehealth Research

Due to the transfer of much academic instruction to distance learning models, the emphasis in prior literature does not reflect the needs that have risen in the current situation. Recent literature reviews of ABA services delivered over telehealth for children diagnosed with autism spectrum disorder (ASD) found that the research has primarily focused on training other individuals to implement behavior analytic procedures (Ferguson et al. 2019). With ABA providers now transitioning to providing services remotely for a greater part of their work, there is a need for a shift in emphasis to reflect the current situation of practitioner-delivered ABA service.

The current situation, with many schools nationwide including distance learning as part of their educational model, means that more behavior analysts and teachers will be implementing clinical sessions directly to a child through a videoconferencing platform. This contrasts with the topic emphasized in prior literature – that of training parents or other caregivers to implement the behavior analytic procedures as was reported until now (Rodriguez, 2020). As such, investigations of topics relating to practitioner-delivered services are needed.

Finally, no investigations identifying prerequisites for students to benefit from remote instruction in terms of verbal behavior developmental cusps are available in the published literature. Conducting an empirical analysis of outcomes for students of differing levels of verbal behavior development can be an important first step in helping to identify potential prerequisites for students to benefit from remote instruction. The benefit from such an analysis in terms of verbal behavior development is twofold. First, such an analysis can help identify who is most likely to benefit from remote instruction. Second, the use of verbal behavior developmental cusps as a measure of prerequisites can be helpful as there is much published literature on interventions needed to build those cusps for students who do not develop them without external interventions (e.g. see Greer & Ross, 2008).

Past Research

There is considerable available data supporting the use of remote delivery of ABA services. For example, Lindgren et al. (2016) conducted a study with 107 children and found that center-based and home-based telehealth models were significantly less costly than an in-home model and produced similar outcomes in terms of reduction of problem behavior. Additionally, parents rated telehealth services as acceptable. More recently, Unholz-Bowden et al. (2020) conducted a systematic review of 30 studies investigating the effectiveness of delivery of ABA services of telehealth. Among many, the findings reported across these studies include direct measurement of effective training of caregivers and teachers to implement behavioral interventions to increase habilitative behaviors while decreasing problem behaviors.

One such evaluation is described in Knowles et al. (2017) who reported the effects of a teacher training package to decrease problem behavior in a classroom setting. This study used a multiple baseline across behaviors design to test the use of a telehealth consultation model to increase evidence-based teaching behaviors (e.g., vocal praise, prompts, opportunity to respond) by teachers to reduce problem behavior. The participants included one teacher and four children (8-9 years old) in a self-contained classroom. The telehealth consultation model consisted of a treatment package including performance feedback, modeling, review of data, and video self-modeling. The results demonstrated increased occurrences of target teaching behavior and substantial decreases of problem behavior across all students.

In another example, Benson and colleagues (2018) report the effectiveness of remote parent training in a home setting. The experimenters used a reversal design to test the effects of training parents to conduct a functional analysis (FA) and implement a functional communication training (FCT) intervention to decrease rates of self-injurious behavior (SIB) for two individuals with disabilities. Training in the FBA consisted of a pre-session review of the experimenter with the parent of target procedures and responses to different possible behaviors of the child. FCT training consisted of the experimenter reviewing target procedures pre-session. The results indicate that the parent-implemented FA and FCT was effective in reducing the rate of SIB for both students.

Finally, Meadan et al. (2016) report the effectiveness of training parents to increase the social communication behavior of their children in a home setting. The experimenters used a multiple baseline across strategies design to test the effects of naturalistic teaching strategies (e.g. modeling, mand-model, time-delay, and environmental arrangements) on social communication behaviors by three preschool aged children with ASD. Specifically, interventionists trained the mothers of child participants in implementation of the abovementioned behavioral strategies. Interventionists used a treatment package including performance-feedback, within-session instruction, modeling, pre-session instruction, caregiver assessment, and video self-modeling to train the parents in implementation of the intervention. Results demonstrated increases in social verbal behavioral initiations and increased percentage of intervals with social verbal behaviors for all participants. All parents demonstrated increases in rate and fidelity of target strategies following coaching by experimenters.

Overall, it is important to highlight that the overwhelming majority of published literature investigates clinical applications of ABA – with an emphasis on problem behavior reduction.

Gaps in the Available Research

Despite the success of interventions reported thus far, one gap that still exists in available literature is testing the effects of therapist-directed interventions on students acquiring novel behavior over a remote instructional model. As reported by Unholz-Bowden et al.’s (2020) systematic review, the scope of the research was focused on parent or caregiver training and most articles reviewed interventions focusing on behavior reduction strategies. The majority of articles reviewed investigated the effects of caregivers who were trained by the experimenter and then delivered interventions – with the majority of articles discussing behavior reduction strategies, The authors noted that for investigations on establishing novel communicative behaviors, only one study included discrete trial training (Hay-Hansson & Eldevik, 2013) and one study included mand and echoics training (Simacek et al. 2017)..

Although not included in the systematic review, Ferguson et al. (2020) reported an evaluation of discrete trial teaching in dyad arrangements over telehealth and found that all participants acquired primary and secondary targets and five of six participants acquired observational primary and secondary targets without direct teaching. In conclusion, the current literature demonstrates initial success of remote behavior analytic service delivery, especially for participants who learn by observation of others. Further, there seems to be an emphasis on remote application of ABA to decrease problem behaviors (i.e., clinical applications of ABA) while there is little research investigating using principles of ABA for educational ends to teach novel behavior

As remote learning for students and service delivery for practitioners becomes more widespread due to COVID-19, it is important to evaluate more closely the viability and effectiveness of remote service delivery with greater precision and across more comprehensive forms of service delivery. Specifically, as many schools across the nation have started to incorporate some degree of remote instruction across their entire educational service delivery, it would be appropriate to evaluate, a) teacher-delivered acquisition strategies, and b) the effectiveness of a school-wide behavior analytic service package. Finally, no literature is available in terms of verbal behavior developmental prerequisites to best benefit from remote instruction. Thus, identification of verbal behavior developmental cusps that allow for a student to benefit most from remote instruction is an important area of investigation.

CABAS®

The Comprehensive Application of Behavior Analysis to Schooling (CABAS®) model is a comprehensive educational model that incorporates tactics from across the science of behavior and applies them to teaching, supervision, and parent involvement (Singer-Dudek, Speckman, & Nuzzolo, 2010). At the heart of the CABAS® model lies the students and their outcomes. Some important components in the CABAS® system include the learn unit (Albers & Greer, 1991) as the fundamental unit of instruction. The Teacher Performance Rate and Accuracy (TPRA) procedure (Ingham & Greer, 1992) is used to ensure treatment fidelity and help diagnose any issues interfering with the student learning. The decision-making algorithm (Keohane & Greer, 2005) is used to analyze data and optimize instruction to be responsive to the performance of the child. The identification of the verbal behavior development (Ross & Greer, 2008) of children is used to structure appropriate instructional sequences and finally, the *Early Learner Curriculum and Achievement Record* (Greer et al. 2020) is a comprehensive curriculum aligned with national and state preschool standards from which target objectives are drawn that has been shown to be related to long-term beneficial effects for students who complete instruction using this curriculum (McGarrell et al., 2009; Waddington & Reed, 2009).

This model has been shown to be effective through repeated investigations. Specifically, a functional analysis (Selinske et al., 1991) a comparative analysis of the program after 20 years (Singer-Dudek et al., 2010), a sustained analysis of the above model (Greer et al.,1989), and an empirical descriptive analysis (Reed et al., 2006) have demonstrated that this program is successful in producing learning outcomes for students enrolled. Further, successful replications of the above schooling model across programs in Italy (Lamm & Greer, 1991) and South Korea (Park et al., 2020) have all consistently shown effective outcomes in this treatment package. However, despite these repeated successful replications, this model has not yet been tested across a remote learning medium of service delivery.

There have been several concerns raised as to the barriers that exist which limit the effectiveness of remote instruction. For example, the results of a survey of 1,181 participants conducted in California during the summer of 2020 (Partnership for Los Angeles Schools, 2020) found that 25% of families have no fixed internet access at home. Further, despite the schools providing digital education platforms (e.g. Google Classroom), 47% of parents/guardians reported never having visited the school's digital education platform. Even for the students who do not have difficulty accessing materials, the home environment resulted in added difficulties with 53% of respondents citing issues maintaining motivation to complete schoolwork as a major barrier to student success in virtual learning. Thus, the transition of educational service delivery from an in-person to an online format can encounter barriers in, a) students' access to materials provided, as well as b) the effectiveness of the delivery. Further, findings (Heppen et al. 2017) examining the 1,224 ninth graders who had previously failed a high school algebra course delivered in a face-to-face setting found that those who participated in an online recovery course reported the course to be more difficult, were less likely to recover credit, and scored lower on an algebra posttest. The authors concluded that students who struggle in-person may struggle even more online.

An Extension of CABAS® to Remote Provision

As many school systems have incorporated some form of distance learning as a way to follow local and national health guidelines while continuing the education for the students in their care, there is a concern that the education of many students is being disrupted. Specifically, some researchers have expressed concern that the school closures could have deleterious effects on an entire generation of students – including those with special needs (Psacharopoulos et al. 2020). Given the robust literature that the components of CABAS® emerge from, as well as the literature demonstrating effectiveness of CABAS® as a system across replications from different decades and continents, it would be worthwhile to investigate if this success in educating preschool students is continuing in a remote learning model that it has come to incorporate due to the national health crisis of COVID-19. In this series of experiments, I will investigate comparisons of in-person and remote instruction which use all the different components of the CABAS® model. In addition to evaluating the outcomes of the CABAS® model of instruction for different students classified by their level of verbal behavior development, it would be worthwhile to conduct a carefully controlled comparison between in-person and virtual instruction in which as many variables as possible are held constant and the only change in intervention would be the medium of service delivery.

The survey results cited above (Partnership for Los Angeles Schools, 2020) highlight that barriers to success in delivering educational services in a distance learning format can result from lack of access, technical fluency, or effectiveness of instruction. That is, even for students who have the necessary technological access and fluency, it may be that the actual medium of remote instruction will present with other challenges. Thus, it is possible that these barriers result in different outcomes from in-person models when replicating effective educational models in a remote service-delivery model.

Thus, to investigate the effectiveness of remote instruction it would be prudent to conduct tests of closely controlled comparisons of remote versus in-person instruction. The data from such an experiment would be valuable in helping identify if the distance learning medium of instruction, on its own, serves as a barrier to effective instruction. Further, it would be prudent to compare the effectiveness of a system-wide model which incorporates remote instruction compared to a purely in-person model. The system-wide comparison would help look at actual outcomes on a larger scale to determine if outside factors (e.g. parent availability, technological access/fluency, limitation on which objectives can be taught) from pure instructional delivery can influence student outcomes.

The findings of these experiments are educationally and socially significant in that they can be used to extend the basic science of verbal behavior development and see if certain cusps/capabilities might serve as prerequisites for students to benefit from remote instruction. As mentioned above, a core component of the CABAS® model involves the use of identification and establishment of verbal behavior developmental cusps to identify appropriate forms of instruction (e.g. teaching students who demonstrate bidirectional Naming (BiN) using model learn units as opposed to direct learn units as reported by Greer and colleagues (2011)) as well as identifying which instructional targets may be appropriate for students given their level of verbal behavior development. A more in-depth explanation of the progressive stages of verbal behavior development is included in the introduction to Experiment II.

Extending this literature can help inform educational policy during the ongoing COVID-19 pandemic and beyond. Further, the data collected in a between-student analysis while classifying students by level of VB development can serve to extend the basic literature on VB development. If some students show that they learn well in an online setting while others do not, the data collected can help form a foundation for further investigation as to which sources of stimulus control, which instructional histories, or what instructional sequences are necessary for students to benefit from remote instruction. Finally, if I am able to identify which characteristics can help predict that a student will benefit from remote delivery of behavior analytic services, this can help practitioners provide services for individuals needing support in geographically remote areas – even when there is no pandemic forcing schools to close. Based on the literature reviewed above, it seems that the CABAS® model of instruction is a robust educational model with a strong base of evidence for its effectiveness in teaching preschool-aged students with and without disabilities. As the COVID-19 pandemic has resulted in nearly all schools nationwide offering some distance learning, it is important to find a way to test the effectiveness of remote instructional delivery. Also, beyond the COVID-19 pandemic, remote delivery of behavior analytic services is an important area of research in that it can be used to provide support for individuals living in geographically remote locations.

A carefully designed comparison of in-person and remote instruction can be helpful in identifying whether certain cusps and capabilities serve as prerequisites for students to benefit from remote instruction. Ultimately, this program of research can help identify which prerequisite or instructional history is necessary to establish so that students can be prepared to benefit from remote instruction. Moreover, a system-wide analysis and comparison of CABAS® in a remote instruction model to student outcomes from CABAS® in an in-person model can be helpful to extend the existing literature supporting the use of the CABAS® model of instruction as a means to help teach preschool-aged children with and without disabilities.

Research Questions

Experimental Analysis (Study 1)

In this experiment I will conduct an experimental analysis across in-person and remote instruction. I will hold all variables constant except the modality of instruction, be it in-person or remote. When all variables are held constant except for the medium of service delivery, I will seek to answer the following questions:

1) Will the rate of learning differ in a remote setting when compared to in-person?
2) Will the rate of learn unit presentation differ in a remote setting when compared to in-person?
3) Will there be a difference in 14-day and 21-day maintenance measures for objectives mastered in a remote setting when compared to in-person?
4) Will the presence of the Naming capability predict a difference in rate of learning, LU presentation, or maintenance measures?

System-Wide Comparative Analysis (Study 2)

In this study I will review data collected across an entire school campus to try and identify variables which will predict the outcomes of success in terms of in-person or remote instruction. I will conduct a statistical analysis to answer the following questions:

1) Will the total number of learn units delivered in a hybrid model differ from a model delivering 100% of instruction in-person?
2) Will the rate of learning differ in a hybrid model with one delivering 100% of instruction in-person?
3) Will the number of supervisor observations to ensure instructional fidelity conducted in the hybrid educational model differ from the number of supervisor observations conducted during a fully in-person model?
4) Will the cost of educational achievement measures differ in a hybrid model from a model delivering 100% of instruction in-person?
5) What is the reliability of the measurement and fidelity of teaching?
6) Will students of differing levels of VB development demonstrate a difference in attendance and instruction received in the hybrid educational model?
7) Will students of differing levels of VB development demonstrate a difference in rate of learning and objectives mastered in the hybrid educational model?
8) Will the cost per measure of educational achievement differ for students of differing level of VB in a hybrid educational model?
9) What is the reliability of the measurement and fidelity of teaching?

Chapter II

An Experimental Analysis of In-Person and Remote Instruction

Abstract

I used a reversal design to compare the effects of in-person and remote instruction on three dependent variables for six preschool-aged participants with disabilities. Three participants demonstrated either the listener half of naming or bidirectional Naming while the other three participants did not demonstrate any Naming. Instruction was delivered to teach objectives across novel tacts and whole word sight words. The variables measured included (a) the rate of learning, (b) the rate of learn unit presentation, (c) number of criteria met, and (d) percentage of correct responses to mastered responses on 14- and 21-day unconsequated follow up measures. Instruction was delivered using learn unit instruction. Distribution of instruction and all instructional methods and materials across in-person and remote conditions were equated so that the only difference in intervention across conditions was the medium of instruction. The results indicate that 3 of 6 participants reliably mastered objectives and completed instruction faster in-person while with the other 3 participants the results were mixed. Overall, participants mastered objectives and completed instruction faster in-person in approximately half of the comparisons while showing no difference or learning faster in the remaining comparisons. No consistent difference was shown in maintenance measures. Further, no difference in overall outcomes was apparent for students with Naming compared to students who did not demonstrate Naming. Results are discussed in terms of extending these findings across larger sample sizes to identify possible prerequisites and predictors of which individuals will most benefit from remote instruction.

Keywords: COVID-19, telehealth, remote instruction, virtual instruction

Introduction

During the COVID-19 pandemic, governmental orders led to the closure of schools in all 50 states in America (Fronapfel & Demchak, 2020). As a result of this, service providers across the nation are presented with an unprecedented challenge of delivering nearly all educational services (both behavior analytic and not) remotely. As the shift to remote service delivery is due to a public health crisis which has no precedent in the last century, the research previously conducted on this topic provides a good foundation but leaves some questions still unanswered.

Remote service delivery in ABA was used to evaluate means to deliver services to individuals living in geographically remote locations where in-person services were not available (Higgins et al., 2017). In situations in which service delivery occurs due to geographical remoteness, the services delivered remotely will often have a different emphasis than the current situation during the COVID-19 pandemic. For example, a recent literature review of remote delivery of ABA services (Ferguson et al., 2019) found that research on telehealth in ABA emphasized training other individuals to implement behavior analytic procedures. This contrasts with the current situation where ABA service providers are providing comprehensive remote service to replace in-person services.

There have been a number of studies demonstrating 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 (Fischer et al., 2014), and training teachers to use positive reinforcement strategies to decrease problem behavior in a classroom (Knowles et al., 2017). Few studies conducted direct comparisons between similar in-person and remote instruction.

One study that compared outcomes from in-person to remote intervention was Lindgren et al. (2016) who compared outcomes and costs for implementing ABA interventions to reduce problem behavior for 107 participants by using three different service delivery models: in-home therapy, clinic-based telehealth, and home-based telehealth. The results demonstrated that all three delivery models were successful in reducing problem behavior by more than 90% using an intervention package consisting of training parents to conduct functional analysis (FA) and functional communication training (FCT). Further, the total costs for the telehealth models were significantly less than in-home therapy.

To compare outcomes from training in-person instruction to training conducted remotely, Hay-Hansson and Eldevik (2013) used a group comparison with fourteen participants who were randomly assigned to two groups. One group received training on implementation of discrete trials training in-person while the other received the training via videoconference. Following the training, results showed both groups improved significantly following the training with no significant differences between either group. These results indicate that training technicians on components of service delivery can be done equally effectively remotely when compared to in-person training. Notably, outcome measures focused on technician delivery of the target intervention without reports of student learning outcomes.

Finally, 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 maintaining the same frequency of instructional sessions remotely as they were receiving in-person.

The studies cited above found that delivery of ABA services remotely resulted in similar results to delivery of ABA services in-person. However, both the Lindgren et al., (2016) and

Hay-Hansson and Eldevik (2013) studies investigated interventions to reduce problem behavior which consisted of the behavior analyst training others in implementing an ABA-based intervention. Pollard et al., (2021) did compare effectiveness of clinician-delivered instruction but did not report a detailed comparison of direct measures of learning but rather reported on the average percentage of correct responses across acquisition, generalization, and maintenance targets. There are few studies which conduct an analysis of direct measures of learning (i.e., number of learning opportunities provided, objectives mastered, and rate of learning) of a clinician-delivered intervention on the acquisition of novel behavior targets.

However, there are some concerns relating to the effectiveness of remote learning (see Partnership for Los Angeles Schools, 2020; Psacharopoulos, Collis, Patrino, & Vegas, 2020). These findings suggest that some students may benefit from remote instruction in comparable ways when provided with a comprehensive educational model that incorporates behavior analytic tactics and methodology.

Holding all variables constant aside from modality of instruction, the data in this experiment will seek to answer the following research questions:

1) Will the rate of learning differ in a remote setting differ when compared to in-person?
2) Will the rate of learn unit presentation differ in a remote setting when compared to in-person?
3) Will there be a difference in 14-day and 21-day maintenance measures for objectives mastered in a remote setting when compared to in-person?
4) Will the presence of the Naming capability predict a difference in rate of learning, LU presentation, or maintenance measures?

Method

Participants

Six preschool-aged boys participated in this study. I 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 and, 2) each participant had the self-management prerequisite to remain seated and respond throughout a 30-minute instructional session delivered over teleconference, and 3) and their parents were willing to commit to completing at least one session per day of virtual instruction. Thus, I selected thee participants because they were under strong instructional control for receiving instruction prior to the onset of the study.

All participants had classifications as preschoolers with a disability and had individualized education plans (IEPs). In terms of verbal behavior developmental cusps (Greer & Ross, 2008) all participants demonstrated conditioned reinforcement for faces, voices, and attending to 2D and 3D stimuli. Further, all participants demonstrated advanced listener literacy (i.e., responded to at least 20 vocal directions without the aid of any visual cues), independent tacts and independent mands (i.e., used vocal verbal behavior to mediate their environment for social (tact) or tangible (mand) functions) in repertoire. Three of six participants demonstrated some joining of the listener and speaker cusps (i.e., Uni or BiDirectional Naming) while three participants did not demonstrate any degree of Naming. All participants attended classrooms which operated using the Comprehensive Application of Behavior Analysis to School (CABAS®) system of instruction. This is a model which uses the principles and tactics of ABA to guide all instructional methods and decisions. A review of key components of this model are presented in the introduction.

Dente was a 4-year 1-month-old boy operating on a listener/speaker level of verbal behavior and demonstrated BiN (i.e., his listener and speaker repertoires were fully joined at the onset of the study and was able to learn word-object relations as a speaker incidentally without direct reinforcement provided by the teacher). Dente responded consistently to vocal directions and had over 50 independent mands and tacts in repertoire. 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.

Jales was 4-year 4-month-old-boy operating on a listener/speaker pre-UniN level of verbal behavior meaning that the listener and speaker repertoires were not fully joined at the onset of the study. Jales responded to a variety of vocal instructions and had over 50 independent mands and tacts in repertoire. Jales a full-scale score of 77 and a verbal comprehension score of 71 on the WPPSI – IV.

David was a 4-year 8-month-old boy operating on a listener/speaker level of verbal behavior and demonstrated BiN (i.e., his listener and speaker repertoires were fully joined at the onset of the study and was able to learn word-object relations as a speaker incidentally without direct reinforcement provided by the teacher). David responded consistently to vocal directions and had over 50 independent mands and tacts in repertoire. 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.

Clement was a 3-year 4-month-old boy operating on a listener/speaker level of verbal behavior and demonstrated UniN (i.e., his listener and speaker repertoires were partially joined at the onset of the study and he was able to learn word-object relations as a listener incidentally without direct reinforcement provided by the teacher). Clement responded fluently to vocal directions, had over 50 independent mands and tacts in repertoire. Clement received an auditory comprehension score of 73, expressive comprehension score of 71, and total language score of 70 on the *Preschool Language Scale*.

Nat was a 4-year 6-month-old boy operating on a listener/speaker level of verbal behavior and demonstrated BiN. Nat demonstrated a fluent listener repertoire and emitted over 50 independent mands and tacts in repertoire. 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.

Mike was a 4-year 6-month-old boy operating on a listener/speaker level of verbal behavior who demonstrated NiN (no degree of Naming). Mike demonstrated a fluent listener repertoire and a variety of independent mands and tacts in an instructional setting. Although Mike emitted a variety of mands and tacts in the instructional setting, he would frequently need 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 while remote instruction was delivered while the participant was at home with the teacher delivering instruction from their home or from the classroom. During in-person instructional sessions the teacher delivered instruction to the student while they were both seated on child-size chairs at a child-sized 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. During remote instruction, the teacher presented instruction from the classroom (or home) through a teleconferencing application (i.e., Zoom) while the participant participated remotely with a caregiver present throughout the session to provide support and redirection if needed. The interventionists delivering instruction during this experiment were the classroom teachers for each student.

Materials

Presentation of instructional materials was identical for in-person and remote sessions. In terms of instructional materials, instructional materials were prepared on PowerPoint presentations prior to the session and presented over the computer screen during both in-person and remote lessons. The stimuli used for tact targets presented on the PowerPoint presentation included pictures which covered between 50% and 75% of the slide. The position of the picture on the slide varied between stimuli. For each set of tacts, four exemplars of each target were used during the instructional sessions rotated across each learn unit. The sight words taught varied in font, color, and size within each set. The sizes of the sight words varied between 16- and 30-point font. The materials in this study were presented on the screen of a laptop computer or a phone on which the session was conducted using the Zoom teleconferencing software.

On the computer the teacher created a digital token board which was used as a token economy to reinforce student responding. During in-person instruction, physical tokens were used for the token economy. Backup reinforcers were used and included access to watch 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 for vocal praise and videos played over teleconference, backup reinforcers were delivered by the caregivers in the home setting.

Dependent Variable and Measurement

Researchers measured responses to learn unit (LU) instruction. In learn unit instruction, the teacher gains the student's attention and delivers a clear unambiguous antecedent. After the student's response the teacher delivers a consequence. A reinforcement procedure is delivered immediately following a correct response while a correction procedure is delivered if the student does not respond within 3s or emits an incorrect response.

If the participant emitted a correct response after the teacher's first presentation of the antecedent a correct response was scored. If the participant did not respond within 3s or emitted an incorrect after the teacher's first presentation of the antecedent an incorrect response was scored. Targets were taught until the student emitted seven consecutive correct responses for that operant at which point that operant was removed from rotation and a new operant was taught in its place. Seven consecutive responses were selected as a criterion for mastery 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.

Responses to LU instruction were the basis for calculating dependent variables. The dependent variables were the cumulative number of operants mastered in each condition, the average number of learn units a participant required to master an operant, the rate of learning, 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 to in each week.

The rate of learning was measured and reported as the total number of learn units delivered across a condition divided by the number of criteria met. The rate of learn unit presentation was obtained 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 a rate of learn units per minute.

Independent Variable

The independent variable in this study was the modality across which the instruction was delivered. Dependent variable measures were compared across targets taught in an in-person setting and a remote setting.

Procedure

This study was conducted during a portion of normal educational services provided to the participants. 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 medium in which instruction was delivered. What follows is an overview of variables controlled for during this experiment.

Pre-Comparison Control Procedures:

Controlling Instruction Across Weeks. To ensure equal distribution of instruction, instructional sessions for targets included in this experiment were delivered daily on the maximum numbers of days or sessions that could be conducted equally across both week in a 2-

week comparison. Typically, sessions were delivered on either three or four 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 sessions 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, I set a target learn unit goal per session. This number was held constant across each session with that 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 one week of remote instruction, they would receive the same distribution of instruction in the week of the in-person instructional condition. Throughout instructional sessions, the total number of independent learn units was held constant across conditions. All instruction was held delivered on a computer screen – both during in-person and remote instruction.

Controlling Targets Across Weeks. After the distribution of instruction (how many instructional days and how many LU per session) was established I 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 etc.). Probe trials were then conducted to determine if the student already had the target response in repertoire. If the student emitted a correct response on any of the probe trials, then that stimulus was removed from the list of potential target operants. If the participant did not emit any correct responses to any of the probe trials, then that target was left in the list. After determining which target operants the participants did not have in repertoire, the remaining targets were placed into two groups so that each group would have 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.

Target sets were assigned using the logistical analysis method (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. A full list of operants mastered by each participant is included in Table 1.

Instructional Procedure.

Pre-Experimental Screening. After assembling the list of potential targets, I conducted unconsequated probes for all participants to determine which responses were not yet in the student'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?" No consequence was provided for the participant's response. Throughout the probe trials, reinforcement was delivered for other behaviors to maintain the student's 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 operation 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 5s for the child to respond. If the child responded correctly within 5s the experimenter delivered vocal praise or generalized reinforcement (i.e., token). If the child emitted an incorrect response or did not respond within 5s the experimenter conducted a correction procedure which 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 learn unit trial. If the participant responded correctly the experimenter continued to the next trial. If the participant responded incorrectly or did not respond within 5s the experimenter repeated the correction procedure up to three times before continuing to the next trial.

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. Participants with the Bidirectional Naming (BiN) capability (Greer & Ross, 2008) received an instructional demonstration learn unit (IDLU) (Hranchuk et al., 2019) while participants who did not demonstrate BiN received three echoics as a response prompt (Billingsley & Romer, 1983) for a new operant when it was first introduced.

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 demonstrated mastery level of responding (seven consecutive correct responses) after which the mastered target was replaced with a novel target. The total number of independent learn units delivered per day remained constant across telehealth and in-person.

In-person instruction. All instruction for participants was delivered by the same teacher that delivered instruction to the student throughout the school year. In the in-person instructional session the teacher 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 complete the predetermined number of independent learn units for the session.

Remote Learning. During remote instruction, the teacher delivered instruction from the classroom or home[1] while the student participated from a home setting with a caregiver present to help facilitate the lesson. The caregiver helped in redirecting the student if they needed redirection to attend to the instructor and to deliver backup reinforcers when the child earned them. Instruction was led by the teacher with support provided by the caregiver as described above.

Maintenance. In the maintenance probes, the target stimuli were presented to the participant with the experimenter's antecedent "what is this?" A correct response was scored if the participant emitted the target response within 3s of the instructor's antecedent. An incorrect response was scored if the participant did not respond within 3s or emitted a response other than the target response. No consequence was provided after correct or incorrect responses. Three trials using different exemplars for each operant being probed were presented per maintenance probe. The percentage of correct responses was reported as correct responses out of total opportunities.

Operant Analysis Mastery. 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 as he was the only participant with that sight words included as part of his regular curricular objectives. 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

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

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., (2020).

Experimental Design

I used a naturalistic reversal design (Cooper et al., 2007) to test the relative efficacy of in-person and remote instruction across all dependent variables. In this design, after ensuring that the target operants were not in the student's repertoire, the researcher equated targets so that they would be comparable and taught across in-person or remote instruction conditions. The researcher assigned target sets using the logistical analysis method (Wolery et al., 2014). Instruction occurred in one context for three or four days and then in the other contexts 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 two-week period such that if there were only three days available for an in-person week, I delivered intervention on three days of the remote week while other instructional objectives not included in this analysis were taught on the other days. The researcher rotated the modality of learn unit instruction (in-person or remote) each week as part of the school's pre-existing schedule[2]. At the end of each week, the student'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 student changed their medium of instruction (in-person or remote) based on the school's schedule. It is important to note that although only data collected during the pre-

[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. Conversely, 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 while all students received virtual instruction on Fridays due to the school building being closed for disinfection.

determined 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 normal 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% accuracy. IOA was calculated by comparing the data recorded by the teacher and that of the observer and dividing the number of agreed-upon trials by the total trials recorded and multiplying by 100%. Point-by-point IOA was collected for 36% of sessions for Dente with a 100% agreement, 25% of session for David with a mean agreement of 99% agreement (range 95%-100%), 25% of sessions for Jales with a mean of 98% accuracy (range 95%-100%), 27% of sessions for Nat with a mean of 97% accuracy (range 96%-100%), 27% of sessions for Clement with a mean of 100% accuracy, and 43% of sessions for Mike with a mean of 99% accuracy (range 95%-100%).

An independent observer used the TPRA assessment to evaluate treatment fidelity. Treatment fidelity for the instructor delivering instruction was ensured through the supervisor conducting an observation and ensuring intact delivery of learn units using a TPRA where the accuracy each component of the learn unit is measured and recorded by the supervisor. The percentage of fidelity was obtained by calculating the number of intact learn units delivered and dividing that number by the total number of learn units in the session and multiplying by 100%.

Treatment fidelity was conducted for 14% of sessions with 100% accuracy for Dente, 16% of sessions with 98% accuracy (Range 95%-100%) for David, 10% of sessions for Jales with a mean of 100% accuracy, 9% of sessions for Nat with a mean of 97% accuracy (range 96% - 97%), 27% of sessions for Clement with a mean of 100% accuracy, and 21% of sessions for Mike with a mean of 93% accuracy (range 82%-100%).

Results

Rate of Learning

Figure 1 displays the rate of learning (i.e., mean number of LU delivered per STO mastered) for each student per week for in-person (black bars) and remote (grey bars) instruction. Each condition denoted with a * indicates that the higher number was 20% higher than the lower number – or a 20% difference. The 20% difference was selected as it is slightly more than one standard deviation above the mean of a normal distribution. 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 1 of 3 comparisons while learning at least 20% faster during remote instruction in the other 2 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 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 across both comparisons. 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 comparison while learning faster remotely in his other two comparisons.

In total, using a minimum of 20% as a measure of difference, students learned faster during remote instruction in 7 of 13 (54%) comparisons, while learning faster in 3 of 13 (23%) of comparisons, and showing less than a 20% difference in 3 of 13 (23%) of comparisons.

Rate of LU Presentation

Figure 2 displays the rate of instructional completion measured as number of LU completed per min for each student per week for in-person (black bars) and remote (grey bars) instruction. Each condition denoted with a * indicates that the higher number was 20% higher than the lower number – or a 20% difference. The 20% difference was selected as it is slightly more than one standard deviation above the mean of a normal distribution. Dente completed instruction 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 instruction at least 20% faster during in-person instruction across 1 of 3 comparisons while completing instruction in the other 2 of 3 comparisons with less than a 20% difference. Jales completed instruction at least 20% faster during in-person instruction across both comparisons, Nate completed instruction at least 20% faster during in-person instruction for one comparison while showing less than 20% difference in the other. Clement completed instruction at least 20% faster during in-person instruction for one comparison while showing no difference in the other. Mike completed instruction with less than a 20% difference across conditions in both comparisons.

In total, using a minimum of 20% as a measure of difference, students completed 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 students completed instruction 20% faster remotely when compared to in-person across any comparisons.

Number of Targets Meeting Criterion

Figure 3 displays the number of criteria met for each student per comparison for each student per week for in-person (black bars) and remote (grey bars) instruction. Each condition denoted with a * indicates that the higher number was 20% higher than the lower number – or a 20% difference. The 20% difference was selected as it is slightly more than one standard deviation above the mean of a normal distribution. Dente met at least 20% more criteria during in-person instruction across both comparisons, David met at least 20% more criteria during in-person instruction across 1 of 3 comparisons while meeting at least 20% more criteria remotely in the other 2 of 3 comparisons. Jales met at least 20% more criteria during in-person instruction across both comparisons, Nate showed less than a 20% difference in number of criteria met across both comparisons. Clement met at least 20% more criteria during in-person instruction for one comparison while showing no difference in the other. Mike met at least 20% more criteria remotely across both comparisons.

In total, using a minimum of 20% as a measure of difference, students met at least 20% more criteria for 7 of 13 (54%) comparisons while showing less than a 20% difference in 4 of 13 (31%) comparisons with less than a 20% difference in the remaining 2 of 13 (15%) comparisons.

Naming Cusp and Instructional Variables

The participants who had demonstrated some degree of Naming (either UniN or BiN) were Clement, Dente, David, and Nat. These participants learned faster during in-person instruction in 55% (5 of 9) of comparisons compared to faster learning remotely for 22% (2 of 9) of comparisons and no difference in 22% (2 of 9). Further, students who demonstrated UniN or

BiN completed instruction faster when delivered in-person for 44% of comparisons (4 of 9) compared to no difference in the remaining 55% (5 of 9) of comparisons. Thus, it seems that even students who demonstrated some degree of Naming did learn faster during in-person instruction but showed less of a difference in rate of instructional completion.

In the 14-day and 21-day follow up probes, the students who demonstrated BiN demonstrated 20% or more correct responding for 29% (4 of 14) of comparisons while showing less than a 20% difference across the remaining 71% (10 of 14) of comparisons.

14-Day and 21-Day Follow-up Probes

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 student per comparison for in-person (black bars) and remote (grey bars) instruction. Dente demonstrated less than a 20% difference in correct responding in all follow-up probes across six comparisons. David responded with 20% greater accuracy in follow-up probes following in-person instruction for 2 of 6 comparisons while demonstrating no difference for the remaining 4 of 6 comparisons. Jales responded with 20% greater accuracy following remote instruction for 1 of 4 comparisons while demonstrating no difference for the remaining 3 of 4 comparisons. Clement responded with 20% greater accuracy following in-person instruction for 1 of 4 comparisons while demonstrating less than a 20% difference for the remaining 3 of 4 comparisons. Mike demonstrated less than a 20% difference in correct responding in all follow-up probes.

In total, students responded with 20% or more correct responses following in-person instruction for five of 22 (23%) comparisons, 20% more correct responses following remote instruction for one of 22 (5%) of comparisons, and less than a 20% difference across the remaining 16 of 22 (73%) comparisons.

Discussion

In order to test the effects of instructional modality (i.e., in-person vs remote) on rate of learning and number of targets mastered, we used a reversal design to teach a series of potential operants across the in-person and remote instructional modality while controlling many other confounding variables, Data were collected on the number of instructional opportunities delivered throughout the sessions, number of objectives mastered, average number of learn units delivered per minute of instruction, as well as percentage of correct responding on 14- and 21-day follow-up probes. 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 a faster rate of learning for 54% of comparisons, faster instructional delivery for 46% of comparisons and better maintenance for 23% of comparisons. This is compared to either no difference or faster in remote instruction for 46% of comparisons for rate of learning, 54% of rate of instructional completions, and 73% of 14 and 21-day day follow-up maintenance measures.

Notable Findings

One important finding of this experiment is that completing instruction in-person results in faster learning and faster instructional completion in approximately half of the comparisons while there is no difference or faster learning and instructional completion occurs in the other half. For maintenance measures, there was no difference across 73% of comparisons. As in-person instructional delivery is more effective and efficient for half of the comparisons while no change is shown in the other half, this indicates that delivery of empirically validated educational tactics delivered with high fidelity can produce equivalent learning outcomes during remote instruction. Although half of the participants showed consistently better outcomes during in-person instruction, the other half did not. This lends support to the potential effectiveness of remote instruction building on the findings of Pollard et al. (2021) and Unholz-Bowden et al., (2021) as the results demonstrating equal or greater effectiveness of remote instructional delivery for half of the comparisons reported suggest that, in many cases, remote instructional delivery is a potentially effective option to consider, especially if in-person instructional services are unavailable.

Another important outcome relates to the difference in instructional outcomes for students who demonstrate some degree of Naming. This finding is important as students who demonstrate the Naming capability could have been predicted to learn equally well remotely and in-person. This assumption may have been made based on previous research which has demonstrated that the presence of Naming allows individuals to learn incidentally through instructional demonstrations even more efficiently than direct consequences (Hranchuk et al., 2019). The stimulus control needed to learn incidentally, as demonstrated through the presence of Naming, has also been suggested as an important prerequisite for children to benefit from whole group instruction (Greer et al., 2011). Yet, the students with some degree of Naming in this study still learned faster when instruction was delivered in-person compared to remotely for 55% of sessions. One possible explanation for the variance in outcomes for students with BiN can lie in the motivating operations. It is possible that receiving instruction in the home setting which is not typically used for instruction serves as a setting event to decrease the reinforcing effectiveness of the educational reinforcers available. Further, when the instructor is delivering instruction remotely, there is a greater delay in delivery of a reinforcer following a target response than when the instructor is in-person.

Finally, it is important to note that rate a faster rate of instructional delivery has been related to improved student learning (Ingham & Greer, 1992). Thus, the finding that in 46% of comparisons, participants receiving instruction in-person demonstrated an increase of at least 20% in number of learn units completed per minute is important as it is also related to better student outcomes. Of note, the rate of instructional presentation was not 20% faster during remote instruction across a single experiment.

Contributions to Existing Literature

This experiment adds to the existing literature on clinician-delivered ABA services provided remotely in which the client participates in a session with all antecedents, prompts, and reinforcers delivered by the clinician via synchronous videoconferencing. Three recent studies report on the effects of direct clinician-delivered ABA service delivery. Ferguson et al., (2020) reported effectiveness of clinician-delivered discrete trial teaching with instructive feedback in a dyad arrangement on tact instruction to six children diagnosed with ASD between 3-7 years of age over telehealth. All participants mastered target skills and demonstrated maintenance of responding 9 days following teaching. Pellegrino and Digennarro-Reed (2020) report on the efficacy of clinician-delivered training using total task chaining with least-to-most prompting to teach skills such as cooking and managing a budget to two adults with IDD over telehealth. Both participants met mastery criterion for the target skills within 15 sessions. Finally, Pollard et al., (2021) report of 17 cases of clients who transitioned from receiving in-person ABA services to receiving services remotely and maintained or improved correct responses across all programs.

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 Digennarro-Reed (2020) but extends those findings by evaluating the relative effectiveness of in-person 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 service delivery transitioned to remote. However, this study diverges from Pollard et al., (2021) in that it provides a detailed analysis of variables such as rate of LU per minute, objectives mastered, and rate of learning for each student while 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 while 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.

Limitations and Future Research

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. Firstly, all participants were under strong instructional control and required minimum prompts in order to remain seated and attend to the instructional sessions. This is important to factor when considering the generalization of these findings. For many students, 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, all participants in this study operated at a minimum on a listener/speaker level of verbal behavior (Greer & Ross, 2008). Thus, this research should be extended to include participants with pre-speaker levels of verbal behavior to test if the modality of instruction will relate to changes in their rate of learning, instruction completion, and maintenance. Further, the target operants that were taught in this study included intraverbal tacts and whole-word sight words which required, for these participants, relatively low levels of response-effort compared to other programs which require greater response effort in each learn unit. It might be the case that target operants which 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.

Despite these limitations, these findings are significant in that they demonstrate in a tightly controlled experiment where all else is held constant, the modality of instruction can affect rate of learning and rate of instructional delivery. Further, this difference is found even in students who demonstrate BiN who learn from incidental exposure to new word-object relations. If the participants with BiN learned faster 61% of the time when receiving instruction in-person compared to remote, it is reasonable to be concerned that children functioning on lower levels of verbal behavior might not be benefitting as much from remote instruction as they would need. Further research should investigate if instructional outcomes are related to level of verbal behavior development. Such an investigation on a wider scale would be important as it would improve external validity to help guide policy over remote instruction as well as helping to identify potential prerequisites which may allow students to benefit more from remote instruction. Further, differences in instructional outcomes can be tied to the funds used for special education. Comparing the funds allocated per student to receive educational services and the instructional outcomes for a school offering remote instruction compared to in-person instruction can be helpful in providing the relevant stakeholders a dollar value in costs of school closures in terms of immediate student outcomes.

Another limitation is that measures of caregiver/parent responding (e.g. rate of approvals, prompting, involvement in instruction) were not included in this investigation. As the caregiver is helping to facilitate instruction during remote sessions, it is reasonable to assume that variables of caregiver behavior will affect the student's learning. For example, participants with caregivers who deliver tangible reinforcers of strong reinforcing effectiveness may learn better in remote session compared to students whose caregivers rely on the teacher to deliver reinforcers over the teleconference modality which may be limited to vocal praise or videos. 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 help the students learn.

Conclusion

The findings of this experiment are significant in that they demonstrate that the modality of instruction was related to faster rate of learning, faster instructional delivery, and better maintenance when instruction was delivered in an in-person setting compared to a remote setting for half of the comparisons. The remaining comparisons demonstrate no difference or improved outcomes in remote instruction. This is an important addition to literature on remote instruction in that it provides an empirical investigation demonstrating that, when variables such as distribution of instruction, materials used for instructional presentation, target operants, and instructional methods are held constant, improved instructional outcomes occur in only half of the comparisons. These preliminary data are promising to support further investigation into using a remote instructional modality as a second line of treatment in environments that do not allow for in-person instructional delivery.

Further, including the levels of verbal behavior development of the participants in our analysis suggests that further research is needed to determine who can be recommended to benefit from remote instruction. In addition to preverbal foundational cusps such as conditioned reinforcement for observing faces or voices, further research can investigate additional cusps which may serve as prerequisites to best benefit from remote instruction. Specifically, research can extend these findings by comparing instructional outcomes across a greater number of students, students with a greater variety of prerequisite skills and verbal behavior developmental cusps, more complex instructional targets, and inclusion of an analysis of caregiver behaviors during remote sessions.

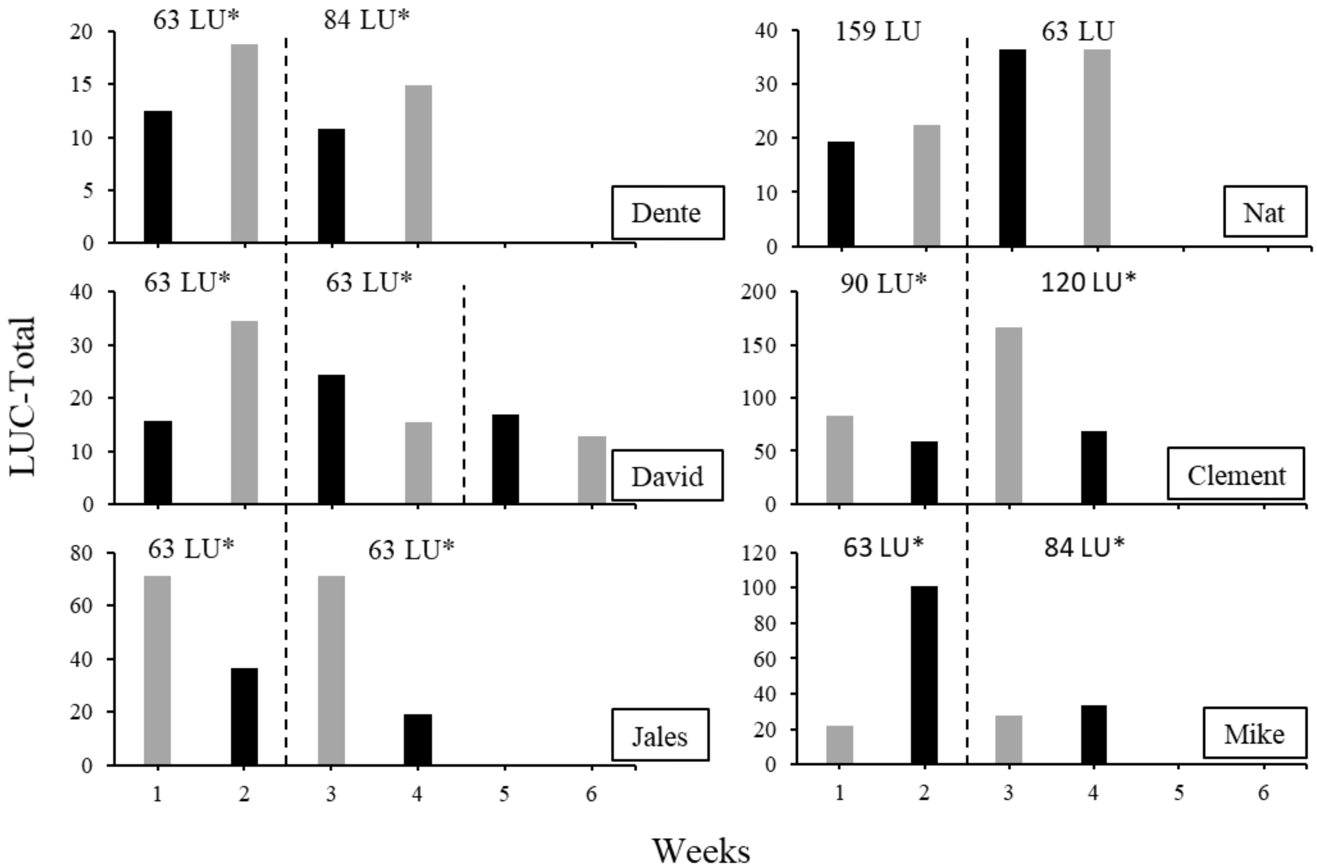
Table 1.

Operants Mastered Per Condition

	Dente	David	Jales	Nat	Clement	Mike
In Person 1	Sloth Lynx Condor Hedgehog Meerkat Yak	Sloth Lynx Condor Meerkat Yak	Condor	Soccer Tennis Hockey Golf Where Play Three	Writing Climbing	Scissor Pen Marker Eraser
Remote 1	Ham salmon Mint	Ham Mint	Ham Pad Thai	Tired Sad Hot Bored Help Jump Hungry My	Honey Camping Pushing	Cubby
In-Person 2	Tying Ribs Mink Tea Pug Shopping Skiing Breadfruit Mink	Shopping Mink Skating	Tying	Crawling Ax	Surprised	Shopping Cleaning Hungry Relaxed Meerkat
Remote 2	Shouting Holding Fig Yam Cashew Mole Crane	Bouncing holding Fig Mole Squid	Holding Pushing Shouting fig yam mole	Yam Hydrant	Coughing Skating Skiing	Surprised Condor Lynx Writing
In-Person 3	N/A	Badge Chalkboard Branch Hydrant Ginger Wok	N/A	N/A		N/A
Remote 3	N/A	Brush Fountain Nail Oyster radio wrench	N/A	N/A		N/A

Figure 1

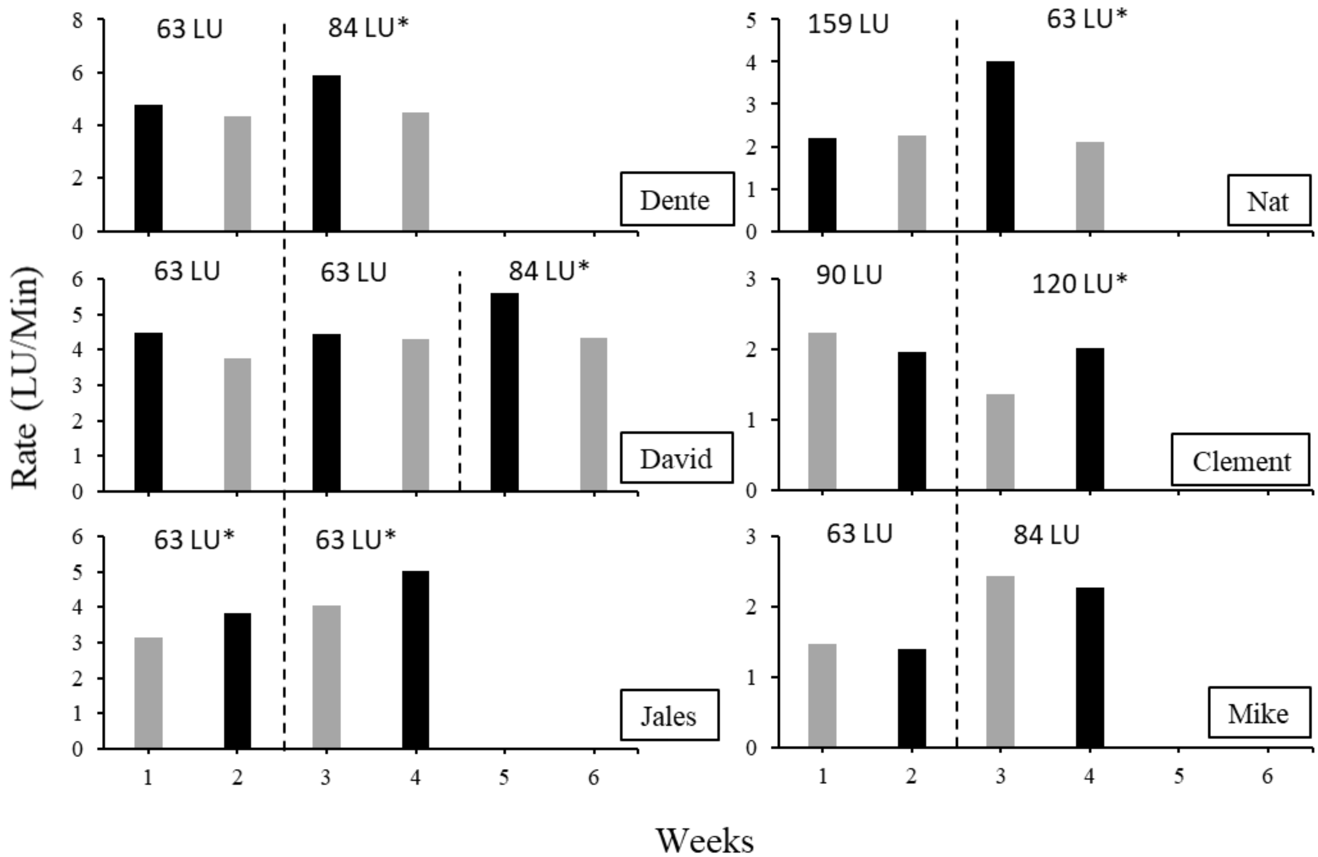
Average Number of Learn Units to Meet a Criterion per Comparison Condition



Note. Dark bars represent in-person instruction while light bars represent remote. A star on top of the condition indicates a difference of 20% or more. 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. David received 3 comparisons as he was the only participant who had different outcomes between comparison 1 and comparison 2.

Figure 2

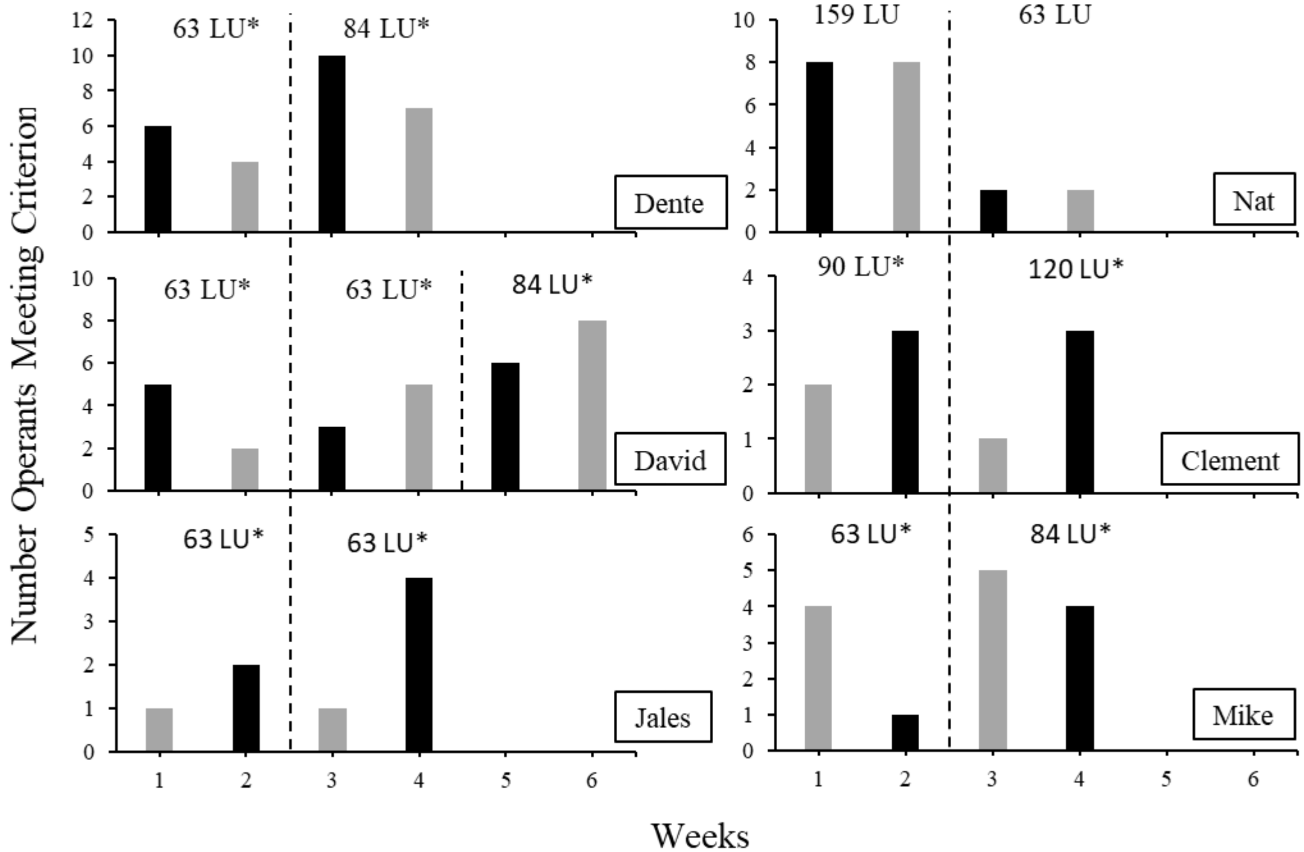
Average Number of Learn Units per Minute Across In-person and Remote Comparisons



Note. Dark bars represent in-person instruction while light bars represent remote. A star indicates a difference of 20% or more 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.

Figure 3

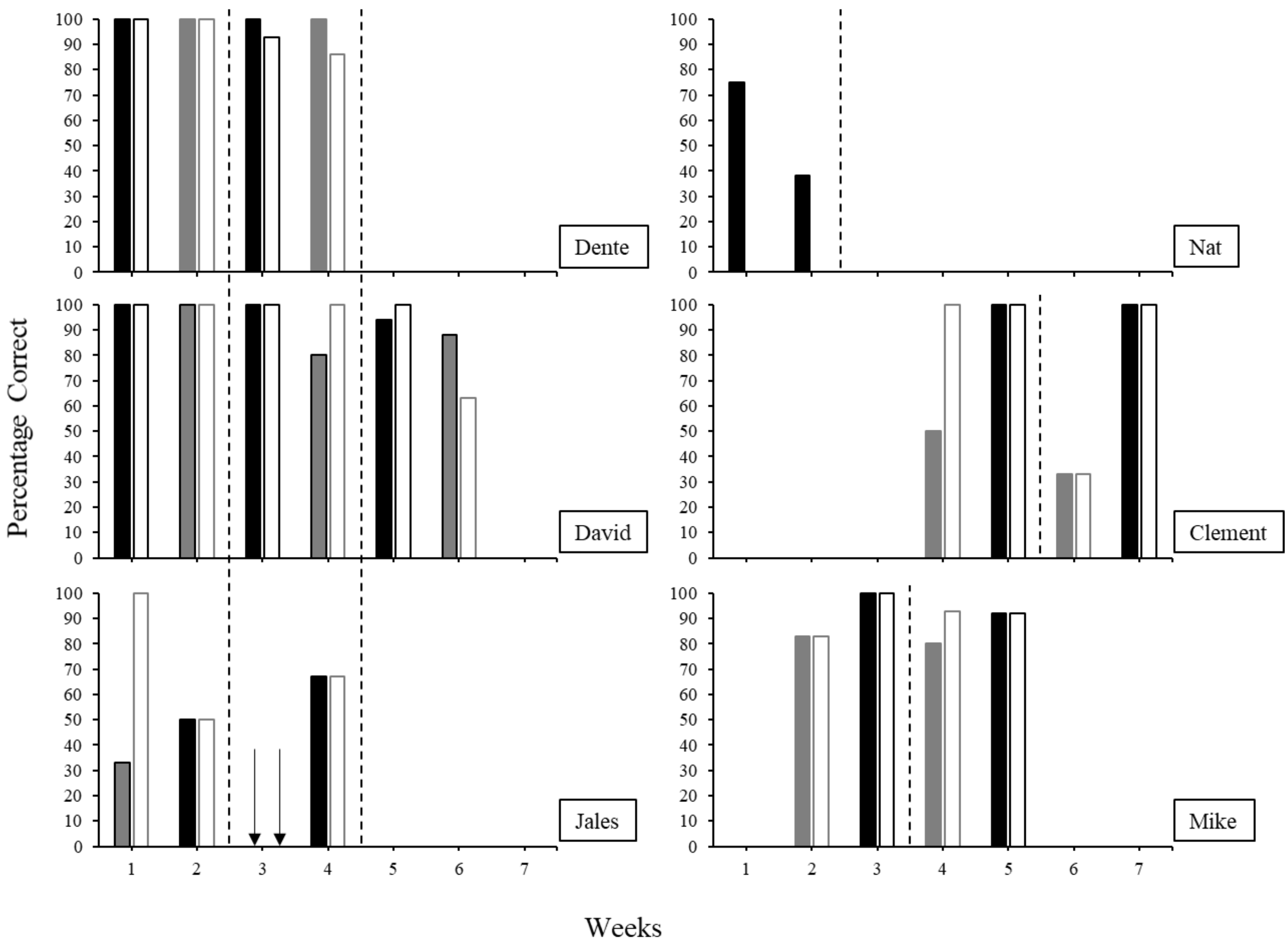
Number of Operants Meeting Criterion per Comparison Condition.



Note. Dark bars represent in-person instruction while light bars represent remote and a star indicates a difference of 20% or more. 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.

Figure 4:

Percent of Correct Responses for 14-Day and 21-Day Follow-Up Probes



Note. Dark bars represent follow-up from operants mastered during in-person instruction while light bars represent follow up of operants mastered. Each cluster represents data from 14 (dark bars) and 21-day (white bars) follow up probes. Clusters with black bars represent maintenance of targets mastered in-person while clusters with grey bars are maintenance of targets mastered remotely. The x-axis reflects the temporal order of weeks during intervention. Full follow-up data was available to be collected for Nat.

Chapter III

A Descriptive Analysis Comparing a Hybrid Educational Model to In-person

Abstract

I used a between groups analysis design to compare learning outcomes and dollar cost per learning outcome for 4 preceding years of in-person instruction in a CABAS® model compared to a hybrid model across in-person and remote instructional delivery. I further grouped students by level of verbal behavior development and compared learning outcomes across groups of students. Learning outcomes from the hybrid instructional model were collected from 108 students. In the hybrid model, 60% of instruction was delivered remotely while the remaining 40% was delivered in-person. The results indicate that although the fidelity of instruction during the hybrid model was equivalent to fidelity achieved during fully in-person instruction, learning outcomes were significantly improved ($p < .05$) in the in-person model for all learning measures except for number of learn units required per short-term objective mastered. Further, the dollar cost per each learning outcome was nearly double in the hybrid model compared to the fully in-person model. In the between-student analysis, results indicate that with one exception, no significant differences were among groups of students who did not demonstrate Naming. In contrast, significant differences were found across all learning measures between students with some degree of Naming and students who did not demonstrate any degree of Naming. The implications of these findings are discussed in terms of comparing these outcomes to other predicted and reported outcomes for remote instruction as well as in terms of the importance of the Naming capability as a prerequisite to benefit from remote instruction.

Introduction

In March of 2020 the COVID-19 pandemic led to governmental orders the closure of schools in all 50 states in America (Fronapfel & Demchak, 2020). Despite the closure of in-person school, many schools continued to provide instruction virtually (Basilaia & Kvavadze, 2020). The shift of entire educational systems to virtual learning is a new phenomenon and the potential effects on educational outcomes are yet unknown. One preliminary hierarchical linear model predicted that students were expected to lose 63-68% of reading gains in reading scores and 37-50% of gains in mathematics scores compared to gains expected during a typical year of full in-person instruction (Kuhfeld et al., 2020). The concern of effectiveness of virtual instruction is heightened for students with special needs. Virtual instruction for young children necessarily requires the presence of caregivers to provide support. Yet, Azoulay (2020) found that parents of students with special needs reported considerable difficulty in supporting virtual instruction for their children. In addition to these predictive models, there have been several studies which reviewed actual educational outcomes for students receiving remote instruction.

For example, one study evaluating the effects of school closures and shift to remote instruction in the Netherlands found that the average learning loss measured by standardized testing at the end of the 2019-2020 school year across 350,000 participants was equivalent to a fifth of a school year (Engzell et al., 2020). This number is nearly identical to the period for which schools were closed to in-person instruction These data are concerning as they suggest that, in a country with a high degree of technological preparedness, the time for which students received virtual instruction resulted on almost no learning.

On the other hand, a recent study by Pollard et al., (2021) of 17 individuals with disabilities found that nearly all participants transitioned to a remote service delivery method and received an equivalent number of instructional sessions remotely as they had been receiving in-person. Further, the authors reported that across all participants, the level of correct responding across acquisition, generalization, and maintenance targets was equal to or improved upon transition to telehealth. Peysin et al., (2021, Study 1 in this dissertation) compared learning across in-person and virtual modalities while holding all else constant for six preschoolers with disabilities. The results suggest that the participants did learn more efficiently and effectively during in-person instruction compared to remote instruction. Specifically, the results show that students learned with at least 20% fewer LU per criterion in-person for 54% of comparisons (compared to 31% of remote comparisons completed faster remotely) and completed at least 20% more learn units per minute for 46% of comparisons (compared to 8% of comparisons in which more 20% learn units were delivered remotely).

The studies mentioned above (Engzell et al., 2020, Pollard et al., 2021 and Peysin et al., 2021 Study 1 in this dissertation) each have different methodologies as well as findings. While Engzell et al., (2020) suggests that close to no learning occurred during remote instruction, both Pollard et al., (2021) and the Peysin et al., (2021, Study 1 in this dissertation) indicate that learning does occur remotely. Specifically, Pollard et al., (2021) found that students maintained levels of correct responding across acquisition, maintenance, and generalization targets when instruction switched to a remote model of service delivery. Peysin et al., (2021, Study 1 in this dissertation) found that a carefully controlled experimental analysis indicated advantages in terms of rate of learning and rate of instructional presentation for in-person instruction.

Comparing the telehealth literature from before the COVID-19 pandemic to data being published since the pandemic provides a further interesting distinction. While reviewing data published from before the pandemic, Uholz-Bowden et al., (2020) stated that the available evidence suggests that remote service delivery of ABA can produce similar results to direct service delivery. As more data are being published from widespread transition to remote instruction since the onset of school closures due to the COVID-19 pandemic, more distinctions are being made. For example, more recent data suggest that when instruction was delivered remotely, either learning didn't occur (Engzell et al., 2020), learning was less effective (Kuhfeld, 2021), or that some in-person instruction was slightly more effective than remote instruction, however students still learned remotely (Peysin et al., 2021, study 1 in this dissertation).

Thus, the current available literature has some disparate findings and a closer analysis is warranted. Each study cited above has some limitations that are worthy to address. Specifically, Engzell et al., (2020) reported population data on standardized test outcomes. Although these can be meaningful data, scores on standardized tests are not a direct measure of learning, but rather a secondary measure of outcomes. Peysin et al., (2021) had a small sample size but while carefully controlling almost all variables, presented with a limitation in that the types of objectives taught were necessarily limited in order to compare near-identical objectives.

Given these limitations, it would be worthwhile to collect data that serve as a direct measure of learning occurring, across a wide variety of objectives which should overall mirror objectives taught in the comparison, for a larger sample of participants. Controlling for these factors in an analysis can help determine which aspects of learning are different between in-person and remote instruction. Further, a detailed analysis across students with clearly identified characteristics can help identify which students benefit or lose most while receiving instruction remotely as well as which prerequisites or developmental repertoires help predict achievement in a remote educational model. Finally, such an analysis conducted in an educational model for which the core components are able to remain constant across remote and in-person instruction can help identify what system-wide components can be emphasized to improve.

CABAS®

One system that is worth analyzing is the Comprehensive Application of Behavior Analysis to Schooling (CABAS®) model. The CABAS® model incorporates tactics from across applied behavior analysis (ABA) literature and applies them to teaching, supervision, and parent involvement. Some central components used in the CABAS® model include the Learn Unit (Albers & Greer, 1991) as a fundamental unit of instruction, the decision-making algorithm (Keohane & Greer, 2005) to optimize instruction, and the Teacher Performance Rate and Accuracy (TPRA) measure (Ingham & Greer, 1992) to provide ongoing supervision and training of teachers and staff.

When delivered in-person, this model has been shown to be effective through repeated investigations. Some examples which have consistently shown effective outcomes with this treatment package include a sustained analysis of the program in the United States (Selinske, Greer, & Lodhi, 1991), a comparative analysis of the program after 20 years (Singer-Dudek, Speckman, & Nuzzolo, 2010), a functional analysis of the model (Greer, McCorkle, & Williams, 1989), and successful replications of the schooling model across programs in Italy (Lamm & Greer, 1991) and South Korea (Park, Du, & Choi, 2020).

The above research base demonstrates that the components which make up this model result in learning when delivered in-person. A close consideration of each of the component features of the CABAS® model suggests that they can be maintained when instruction is delivered remotely. Thus, if the components of the CABAS® model result in positive learning outcomes when present in-person, it would be worthwhile to investigate if replicating delivery of the CABAS® model over a remote platform will result in comparable outcomes. What follows is a brief overview of core CABAS® components which can be extended to remote instruction.

The Learn Unit

The fundamental component of the instructional delivery used in CABAS® is that of the learn unit (Albers & Greer, 1991). The learn unit is a potential three-term contingency which involves an interaction between a teacher and a student. In the learn unit the teacher presents an unambiguous and complete antecedent for the student, the student responds appropriately, and the teacher delivers an appropriate consequence in the form of reinforcement for a correct response or an error correction procedure following an incorrect response. The rate of learn unit presentation in CABAS® schools can be as frequent as four per minute. In the literature, the learn unit is described as a natural fracture of instruction and as such, it serves as a direct measure of instruction delivered by the teacher to the student as well as a means of measuring the student's learning (Greer and McDonough, 1999). Related literature pertaining to variations in learn units (Hranchuk et al., 2018; Greer et al., 2011a) reports that students who demonstrate the Naming capability learn faster when receiving instructional demonstration learn units even faster than direct learn units. Observation of learn unit presentations by trained observers allows for analysis and identification of sources of instructional difficulties. This observation and subsequent analysis is often conducted through the use of the Teacher Performance Rate Accuracy (TPRA) measure (Ingham & Greer, 1992).

The Teacher Performance and Accuracy (TPRA)

During the TPRA observation, the observer records data on the antecedents and consequences provided by the instructor, calibrates the accuracy of data collection on student responses, and records the time of the observation in order to calculate the rate of learn unit delivery (Hranchuk & Williams, 2021; Ingham & Greer, 1992). In this process, the observer can help identify and then correct potential errors including ambiguity in the antecedent, lack of appropriate manipulation of motivating conditions to ensure reinforcement is in place following correct responses, ambiguity in the target response, and so on. The format of the observation serves as an evaluative measure of teacher behavior for fidelity of instruction as well as a diagnostic tool to aid in troubleshooting aspects of instruction which are withholding progress.

As described above, the learn unit serves as a fundamental measure of instruction. Thus, if the learn unit is being delivered learning should occur (Greer and McDonough, 1999). If a student is not learning at a target rate, an observation of the teacher's instruction can be the source for a further analysis to determine the source of the error in instruction which is resulting in the student not learning. This observation is conducted in the form of a TPRA. The behavior analyst supervisor who conducts the TPRA ensures that intact learn units are being presented at rates which help achieve active student responding. Further, in the seminal research on the TPRA, an increase in errorless TPRAs was shown to result in increased student learning (Ingham and Greer, 1992; Ross et al., 2005).

Decision Making Algorithm for Data Analysis

In addition to being the unit of instruction, the learn unit also serves as a unit of analysis. Each learn unit delivered is measured and recorded by the teacher. Immediately upon completion of the session the data are graphed to aid in further analysis using a decision-tree protocol as described in Keohane and Greer (2005). The decision-tree protocol provides an algorithm for teachers to use in order to conduct a multilayer analysis. The first level of analysis consists of a set of rules for identifying trends in graphed data which demonstrate difficulties in instruction. If the trend in the data suggests that the student is not learning, the algorithm leads the teacher to a deeper analysis which takes into consideration all possible sources of instructional difficulty. Possible sources of instructional difficulty include faulty instructional materials, lack of fidelity in the delivery of antecedent or consequence portions of the learn unit, setting events, or missing prerequisite instructional history – including identification of missing verbal behavior developmental cusps and capabilities (Greer and Ross, 2008). In the seminal research (Keohane & Greer, 2005), use of the decision-making protocol across three teachers and six students resulted in significantly more learning objectives being mastered after the decision protocol was implemented when compared to before.

Curriculum

The CABAS® schools use the Early Learner Curriculum and Achievement Record (ELCAR) (Greer et al., 2020) to determine the scope and sequence of instructional objectives. The ELCAR is a comprehensive curriculum which aligns with state and national preschool standards and is designed to help arrange an instructional sequence so as to result in learning for all students. The ELCAR curriculum is built on the Preschool Inventory of Repertoires for Kindergarten curriculum which has been shown to positively impact outcomes for children with ASD (Reed et al., 2006), as well as improving the ability of children to transfer from special to mainstream schools (Waddington and Reed, 2009).

The curriculum is sequenced so that the teacher following the curriculum screens for fundamental prerequisites needed for successful instruction (e.g. preference assessment screening to identify potential reinforcers, instructional control prerequisites for accurate assessment of skills the student has in repertoire, observing response screening to identify potential missing reinforcers etc.) as well as sequences of instructional objectives across self-management, communication, academic literacy, and fine and gross motor skills needed for successful movement in and use of classroom materials. The curriculum specifies the antecedent (e.g., speaker, listener, tact/intraverbal) and type of consequence (i.e., tangible reinforcement in the form of preferred food, items, or activities, generalized reinforcement in the form of vocal praise, and natural reinforcement which specifies that the instructor does not deliver any reinforcement but rather engagement or completion of the task should serve to reinforce for the target behavior).

The series of educational objectives outlined in the ELCAR (Greer et al., 2020) are classified in this experiment as long-term objectives (LTOs) while the subcomponents of the ELCAR objective that are taught individually are classified as short-term objectives (STOs). Thus, an example of an LTO would be for a student to identify nine different colors as a speaker when presented with a stimulus in the target color and asked, "What color is this?" with 100% correct responding across all nine colors presented one time each a probe. An example of a STO would be for a student to identify three different colors as a speaker.

Teacher Training

Another important component of the CABAS® system of instruction is the competency-based teacher training modules which are taught using the personalized system of instruction as outlined in Keller (1968). Teacher training modules are divided in three domains: 1) mastery of the verbal behavior of the science, 2) expertise in applying ABA instruction including delivery of intact learn units, measurement of progress and graphic display of data, and 3) expertise in applying findings from behavior analytic research to help identify and solve instructional difficulties which may arise as part of instruction.

Verbal Behavior Developmental Theory (VBDT)

A central component of CABAS® includes the research base relating to verbal behavior development (Greer & Ross, 2008; Pohl et al., 2018). Identification of sources of stimulus control relevant for students on varying levels of verbal behavior (VB) development allows for optimization of instruction personalized to which stimuli the child has conditioned reinforcement for observing. For example, using the criteria outlined by Greer and Ross (2008) can help a teacher identify that a student does not have conditioned reinforcement for observing human voices. Once this deficit in stimulus control is identified, instruction can be sequenced in a way to establish the relevant sources of stimulus control (e.g., conditioning reinforcement for observing voices protocol as outlined by Greer et al., 2011) before attempting to teach target operants such as following vocal directions, which would likely result in slow or unsuccessful learning if the prerequisite stimulus control of observing human voices is not established. Further, the new stimulus control allowed through the establishment of a cusp or capability allows for learning of new classes of responses that were previously not able to be learned prior to the establishment of the cusp or capability.

Who Benefits?

Comparing system-wide outcomes from fully in-person instruction to a model that incorporates remote instruction can help provide valuable insights as to which aspects of virtual instruction are accounting for differences in outcomes. Data from such a study can help inform future research as to which aspects of a virtual instructional model can be improved to result in improved student outcomes. In addition, conducting group comparisons of students while accounting for differing characteristics can help extend the available literature relating to VB development.

Current literature relating to VB development has identified ways to measure the presence or absence of VB developmental cusps as well as identified interventions to establish these cusps if they are missing (Greer & Ross, 2008). However, the effects of the presence of VB cusps and outcomes from remote instruction is an area that has yet to be subject to an analysis. Findings from such an analysis can help guide further research on which VB cusps may serve as prerequisites for students to benefit from remote instruction. Further, an analysis of VB cusps and learning outcomes can help identify what type of interventions may need to be emphasized in order to produce desirable student outcomes (e.g., emphasizing training of caregivers is likely to be warranted for students who lack preverbal foundational cusps such as conditioned reinforcement for observing voices while instructor-delivered interventions are likely to be beneficial for students who have more listener and speaker cusps such as bidirectional Naming or BiN). Finally, identifying effects of the presence of certain VB cusps on instructional outcomes can be an initial step to eventually identifying instructional protocols to allow students to benefit from remote instruction.

A behavioral cusp, as identified by Rosales-Ruiz and Baer (1997) is defined as a behavior change the allows for an individual to come into contact with new contingencies that result in consequences beyond the immediate change itself. An example of a behavioral cusp is the VB cusp of conditioned reinforcement for attending to voices (Greer and Ross, 2008). Once this new source of stimulus control is established for a child, they are able to then contact contingencies from the spoken word of others. Now it may eventually be possible for vocal praise to become conditioned as a reinforcer, thereby increasing the ability of caregivers to shape habilitative behaviors by praising them when they occur. Or, eventually, a child for whom this cusp is established can come under control of instructions delivered vocally. In addition to contacting new contingencies, previous research suggests that establishment of VB developmental cusps allows for a child to learn much faster (e.g. Greer et al., 2005; Greer et al., 2011a; Greer et al., 2011b). It would be meaningful to see how these cusps affect educational outcomes. Specifically, it would be worthwhile to compare learning outcomes for students categorized based on the four stages outlined by Greer et al., (2017): preverbal, listener, speaker, and joining of listener and speaker (note that the fourth stage consists of two stages: unidirectional (UniN) and bidirectional Naming (BiN). For the sake of this analysis I will count the four stages as five.

Educationally, a child is identified to be missing VB cusps when they either require many instructional presentations to master a novel operant or fail to master novel operants despite the presence of sound instructional methods. This can result from a lack of environmental experiences that allow for establishment of the cusp or as a result of native disabilities (Greer et al., 2017). Although detailed procedures for identifying missing cusps as well as interventional protocols for establishing those cusps are found in the literature (e.g., Greer & Longano, 2010; Greer & Ross, 2008; Greer & Speckman, 2009) a detailed discussion on this topic is outside the emphasis of this experiment. However, it is important to emphasize that identification of these cusps for a student can allow the instructor to design instruction that is appropriate for the student's level of VB development as well as the appropriate prompt level. Table 1 provides a brief overview of the component cusps in each of the five categories as outlined by Greer et al., (2017) that I will use for classification of students to compare educational outcomes by level of verbal behavior development.

Table 1

Brief Overview of VC Developmental Cusps

Level of VB/ Group	VB Cusp	Brief Definition	Examples of Instructional Objectives Student Can Benefit from Once Cusp is Established
Preverbal Group 1	Conditioned reinforcement for observing 2D and 3D stimuli (Du et al., 2015)	Student will observe instructional 2D and 3D stimuli	Generalized matching of identical and non-identical stimuli, discrimination/point-to responses
Preverbal Group 1	Conditioned reinforcement for observing faces and voices (Maffei et al., 2014	Student will observe adult faces and voices	Listener instruction such as point to body parts, discrimination responses such as identify pictures of various stimuli.
Preverbal Group 1	Generalized imitation (Du & Greer, 2014).	Student will imitate motor movements modeled by others where the correspondence between movements serves as the reinforcer for the imitation response	Child learns novel responses from see-do correspondence without requiring consistent physical prompting
Listener Group 2	Auditory Matching (Du et al., 2017)	Student will select identical sounding words in the presence of a nonexemplar	Child learns to emit echoics with point-to-point correspondence. Presence of this cusp is also shown do induce UniN (Greer & Longano, 2010) and Listener Literacy (Du et al., 2017)
Listener Group 2	Listener Literacy (Greer et al., 2005)	Child comes under phonemic stimulus control and fluently responds to spoken directions without any visual aids	Child learns to follow spoken vocal directions without reliance on physical prompts. and begin to emit responses as a listener to word-object relations in listener instructions
Speaker Group 3	Echoic-to-tact/mand (Tsiouri & Greer, 2007).	Verbal operant where child emits speaker response with point-to-point correspondence to verbal antecedent leading to tact or mand function	Child learns novel tact and mand forms through echoing models presented by instructor.
Speaker Group 3	Independent mand (Ross et al., 2006)	Verbal operant where child requests for access to object/activity or for termination of stimulus	Child mediates environment as a speaker to obtain preferred objects or activities

		under deprivation or aversive antecedent conditions where the delivery of specified outcomes serves as reinforcer	
Speaker Group 3	Independent tact (Pistoljevic & Greer, 2006)	verbal operant where child emits speaker response under nonverbal antecedent conditions and response is reinforced by generalized reinforcement	Child emits novel word-object relations for social function.
Joining of the Listener and Speaker Outside and Inside			
Listener/Speaker with UniN Group 4	UniN (Greer et al., 2017)	Listener half of Naming where child learns to respond to word-object relations as a listener incidentally without direct instruction.	Child learns from instructional demonstrations for listener responses
Listener/Speaker with BiN Group 5	BiN (Greer et al., 2017)	Joining of listener and speaker repertoires where child is learns to respond to word-object relations as a speaker incidentally without direct instruction.	Child learns from instructional demonstrations for speaker responses

Current Study

In this experiment I look to conduct a between-year analysis to answer questions relating to educational outcomes for students receiving a model of education which is 60% virtual and 40% in-person compared to 5 prior years of wholly in-person instruction. I seek to build on existing literature demonstrating effectiveness of the CABAS® model of instruction including Singer-Dudek and colleagues (2010). Across this school model I will analyze educational outcomes to determine how transitioning the CABAS® school model to include remote provision affects rate of learning, attendance, and total learning opportunities. Further, I will divide students into groups to by their level of verbal behavior development and conduct a between-student analysis to see if educational outcomes differ based on the level of verbal behavior development of the students.

Specifically, In Experiment III I seek to answer the following research questions:

1) Will the number of supervisor observations to ensure instructional fidelity conducted in the hybrid educational model differ from the number of supervisor observations conducted during a fully in-person model?
2) What is the reliability of the measurement and fidelity of teaching observed in TPRA observations?
3) Will the total number of learn units delivered in a hybrid model differ from a model delivering 100% of instruction in-person?
4) Will the number of objectives mastered and LU per objective differ in a hybrid model with one delivering 100% of instruction in-person?
5) Will the cost of educational achievement measures differ in a hybrid model from a model delivering 100% of instruction in-person?
6) Will students of differing levels of VB development demonstrate a difference in attendance and instruction received in the hybrid educational model?
7) Will students of differing levels of VB development demonstrate a difference in rate of learning and objectives mastered in the hybrid educational model?
8) Will the cost of educational achievement measures differ for students of differing level of VB in a hybrid educational model?

Method

Participants

There were 108 full-time students enrolled for the 2020-2021 school year. Out of 108 students, 88 were boys and 20 were girls. All students were between 3 and 5 years of age. For the 2020-2021 school year the school operated on a hybrid model where 60% of each student's instruction was delivered remotely while the remaining 40% of instruction was delivered in-person. For the purpose of this study, the students were categorized by their level of VB development (Greer and Ross, 2008). There were five categories of VB development included in this analysis: 1) Pre-listener (fewer than 20 fluent listener responses), 2) Listener/Pre-Speaker (20 fluent listener response but fewer than 50 independent tacts/mands), 3) Listener/Speaker with NiN, (50 or more independent tacts/mands but did not demonstrate any degree of Naming), 4) Listener/Speaker with UniN (50 or more independent tacts/mands and demonstrated the listener half of Naming but not the speaker half), and 5) Listener/Speaker with BiN (50 or more independent tacts/mands and demonstrated BiN). The experimenter collected these classifications by interviewing the teachers in the school and asking them to classify their students using the aforementioned criteria for classification.

Twenty eights students were classified as pre-listener/pre-speaker. This means that these students were missing one or more preverbal foundational cusps and did not yet learn to follow 20 vocal one-step directions. An overview of the component cusps for this category are included in Table 1. These students required frequent physical or visual prompts in order to benefit from instruction being presented and generally did not benefit from vocal prompts.

Sixteen students were classified as listener/pre-speaker. This means that these students fluently followed at least 20 vocal directions but emitted fewer than 50 independent tacts or mands. Further, previous research indicates that these students learn novel listener responses faster than students without the listener literacy cusp (Greer et al., 2005). An overview of the component cusps for this category are included in Table 1. These students were missing one or more cusps that allowed them to mediate their environment as speakers. These students demonstrated some more independence and could benefit from vocal prompts during instruction but did not fluently use vocal speech for requests or social functions. These students did not demonstrate any degree of naming.

Twenty-five students were classified as listener/speaker with no degree of Naming. This means that the students emitted at least 50 independent tacts or mands but did not demonstrate any degree of Naming. An overview of the component cusps for this category are included in Table 1. These students demonstrated a greater level of independence than the two categories above but the absence of the Naming cusp meant that these students needed to be directly taught (i.e., occasioned to emit a response and then delivered reinforcement immediately following that response) any new repertoires and did not learn from instructional demonstrations or incidentally to respond as a speaker or listener.

Eighteen students were classified as listener/speaker with UniN. This means that the students emitted at least 50 independent tacts or mands and demonstrated the listener component of Naming but did not demonstrate the speaker component of Naming. An overview of the component cusps for this category are included in Table 1. These students demonstrated a greater level of independence than the categories above and were able to learn listener repertoires (e.g., pointing to or discriminating) incidentally or from instructional demonstrations. Further, the presence of the Naming capability is also related to an increased rate of learning (Greer et al., 2011a). Thus, if students demonstrate UniN they should learn novel listener responses at a faster rate than students who do not demonstrate this capability. The absence of the speaker component of the Naming cusp meant that these students needed to be directly taught (i.e., occasioned to emit a response and then delivered reinforcement immediately following that response) any new speaker repertoires and did not learn from instructional demonstrations or incidentally for speaker skills.

Finally, 13 students operated on a listener/speaker level of VB and demonstrated Bidirectional Naming (BiN). An overview of the component cusps for this category are included in Table 1. This means that these students were fully verbal and could learn incidentally or from instructional demonstrations for both listener and speaker responses. Further, the presence of the Naming capability is also related to an increased rate of learning (Greer et al., 2011a). Thus, if students demonstrate BiN, they should learn novel listener and speaker responses at a faster rate than students who do not demonstrate this capability. The criteria used to classify the students for their level of VB development drew on VBDT as outlined by Greer and Ross (2008).

We examined the Fall 2020 60/40 hybrid model data compared to 4 previous years of data when instruction was delivered entirely in-person. Data for number of students enrolled was collected by retrieving the school list that was updated after the end of the Fall term (i.e., in late December/early January) for each of the years included in the analysis. As the school had some students that attended only for half-day during Fall 2015-2018, every student that attended only half day as counted as half of a full-time student. There were 126.5 full-time students enrolled for Fall 2015. There were 124.5 full-time students enrolled for Fall 2016. There were 128.5 full-time students enrolled for Fall 2017, and there were 127 students enrolled for Fall 2019. Data from Fall 2018 were not included in this analysis as the school opened a new campus and transferred a large percentage of the student body to a new location. This move complicated the process of matching up data reported and student enrollment as the number of students enrolled was not consistent throughout Fall 2018.

It is important to mention that the teachers delivering remote instruction, for the most part, were the more senior teachers who had demonstrate mastery of instructional procedures during in-person instruction. Junior teachers typically would receive multiple errorless TPRAs across a variety of programs for in-person delivery of instruction prior to delivering remote instruction.

Setting and Materials

Data for this experiment were collected in a publicly funded private preschool located outside of a major metropolitan city in the Northeast U.S.A. The school had 13 different classrooms and primarily served students with developmental disabilities such as ASD. The materials used in this study consisted of a wide variety of materials used in educational settings (e.g. pens and date sheets for data collection, backup reinforcers, manipulative toys for mathematical instruction, and so on). Remote instruction was delivered over computer using the Zoom teleconferencing application.

CABAS® Fully In-Person Vs. Hybrid

In our analysis, for years 2015-2019 all instruction was delivered in-person. During fully in-person instruction, students are in the classroom for 5 hr per day for 5 days per week (Monday-Friday). Due to restrictions relating to number of students allowed in the building at one time because of COVID-19, the 2020 school year was a hybrid application of CABAS®. In the hybrid application of CABAS®, 60% of instruction was delivered remotely while 40% remained in-person. The duration for which each student participated in remote instruction varied by student and the availability of a caregiver to help facilitate instruction. The duration for which students engaged in instruction ranged from one 30-min session per week to four sessions per day.

The allocation of instruction, whether in-person or remote, was achieved by dividing the school into half where half the students attend in-person Monday-Thursday for one week while the remaining students receive instruction remotely those days. Every Friday the school campus was closed for deep clean disinfecting due to COVID-19 and resultantly all students received instruction remotely. Thus, in every two-week cycle, all students received four days of in-person instruction (one week of Monday-Thursday) and six days of remote instruction (one week of Monday-Thursday and both Fridays).

During both remote and in-person instruction all components of CABAS® were maintained as outlined above. However, when students were scheduled to receive instruction remotely, they were required to have an adult caregiver present to help facilitate the instruction. Caregiver responsibilities included manipulating materials, delivering reinforcers, or providing any gestural or physical prompts that the teacher was unable to deliver remotely.

Measurement

The data analyzed in this study were teacher-reported data across the first 14 weeks of the Fall semester across the five years 2015-2020 (excluding 2018[3]). All instruction from years 2015-2019 was conducted in-person while instruction during the year 2020 was delivered using a hybrid model incorporating both in-person and remote instruction as described above. Two analyses are reported in this experiment: 1) A comparative analysis for educational outcomes between years 2015-2019 (excluding 2018) ,which will be referred to as the between-years analysis, and 2) An analysis of data collected during the 2020 school year on the hybrid model

[3] The year 2018 was excluded from this analysis as a new campus of the school opened that year and many students switched between campuses thus making it difficult to properly include the data from 2018 in our analysis.

comparing outcomes for different categories of students classified by different levels of VB development. This second analysis will be referred to as a between-student analysis. What follows are descriptions of the categories of data included in the analysis.

TPRAs Conducted

In addition to the learning variables described above, two systems variables are included in our analysis. The first systems variable is included in the between-years analysis and is that of the cumulative number of TPRAs delivered. The TPRA (Ingham & Greer, 1992) is a central component of CABAS® outlined above and serves as a training and diagnostic tool as well as a fidelity measure to ensure instruction is being delivered as it is designed to be. The TPRA consists of an observation of instruction by a classroom teacher or supervisor who measures the fidelity and rate of instructional delivery in a session of instruction. In our comparison, data are reported on cumulative TPRAs (both with error and errorless) conducted each week.

Classification of TPRA. TPRAs were further classified as errorless TPRAs or TPRAs with error. A TPRA was classified as errorless if all antecedent, measurement of target behavior, and consequence components of the LU were delivered properly. The antecedent component was scored as correct if the antecedent was delivered only once the student was attending to the teacher, materials or presentation were clear and unambiguous, and the form of antecedent was delivered as specified in the curriculum (i.e., a verbal antecedent for an intraverbal target operant while a nonverbal antecedent for a pure tact operant). The antecedent was scored as incorrect if any of the above criteria were not met. Measurement of target behavior was scored as correct if the instructor marked a correct or incorrect response correctly. Measurement of target behavior was scored as incorrect if the instructor scored a correct response as incorrect or an incorrect response as correct. Finally, the consequence component of the LU was scored as correct if the instructor delivered the correction procedure with fidelity following an incorrect response or if the instructor delivered the curriculum-specified form of reinforcement following a correct response. The consequence was scored as incorrect if the instructor did not deliver the appropriate form of correction following an incorrect response or reinforcement following an incorrect response. The appropriate forms of reinforcement are specified in the ELCAR (Greer et al., 2020) and in an outline of appropriate instruction by Greer (2002).

Average Number of Weeks Attended

A second systems variable is included in the between-student analysis. Data on number of weeks attended was calculated by reviewing the data of each student and counting the number of weeks for which the student received any instruction. Thus, if the student attended an instructional session and received any LU instruction in a week it was scored as a week attended. If a student did not attend any instructional sessions that week, the week was scored as absent.

Learn Units

One learn unit (LU) consists of a potential interlocking complete three-term contingency including an antecedent from the instructor delivered after ensuring the student is attending, the target response of the student, and the reinforcement or correction procedure following a correct or incorrect response respectively. A measure of learn units completed is included in both the between-year and between-student comparison. For the between-year analysis the measure of learn units compared is the average number of learn units per student per day. This measure was obtained by dividing the total number of learn units completed schoolwide in a week and dividing it by the number of students enrolled and then dividing that number by the number of instructional days in the school calendar that week (LU/students enrolled/instructional days). The number of students enrolled was obtained by reviewing a school roster on file from the end of the fall semester for each respective school year. The number of instructional days was obtained from the school calendar from each respective school year. This measure is reported as *average number of learn units per student per day*.

In the between-student comparison, the LU measure was obtained by calculating the total number of LU by all students in one VB category and dividing that number by the number of students in the category. This measure is reported as the *average total number of LUs per student in a category*.

In addition to the total learn units received, the dollar cost per learn unit was calculated for the between-year analysis as well as the between-student comparison. The dollar cost was calculated by obtaining the total revenue the school received and selecting out the percentage of the revenue that corresponded to the percentage of weeks of the year that are included in our analysis. This number was then divided by the total number of LU completed in our analysis for that year to determine the dollar cost per LU in that year.

For the between-student analysis the total revenue was calculated as in the between-years analysis but then divided again by number of students in each group. This extra step was included to determine the allocation of revenue for those students. This number was then divided by the total number of LU completed by the number of students in that category to determine the dollar cost per LU for students in each category.

Short Term Objective (STO)

Number of STOs. The number of short-term objectives mastered is included in both the between-year and between-student comparison across groups of students with differing levels of VB development. If needed, when teaching a long-term objective (e.g., respond intraverbally to nine different colors) to a student, the teacher may subdivide the long-term objective into smaller components (e.g., identify three colors at a time). Mastery of one sub-component is considered one STO.

A measure of STOs mastered is included in both the between-year and between-student comparison. For the between-year analysis the measure of STOs mastered is the average number of STOs mastered per student per day. This measure was obtained by dividing the total number of STOs mastered schoolwide in a week and dividing it by number of students. This number was then divided by the number of instructional days in the school calendar that week (STOs/students enrolled/instructional days). This measure is reported as *average number of STOs mastered per student per day*.

In the between-student comparison, the STO measure was obtained by calculating the total STOs mastered by all students in one VB category and dividing that number by the number of students in the category. This measure is reported as the *average total number of STOs mastered per student in a category* throughout the weeks included in our analysis. In addition to the total STOs mastered, the dollar cost per STO was calculated for the between-year analysis as well as the between-student comparison. The dollar cost was calculated by obtaining the total revenue the school received and selecting out the percentage of the revenue that corresponded to the percentage of weeks of the year that were included in our analysis. This number was then divided by total number of STOs mastered in our analysis for that year to determine the dollar cost per STO mastered in that year. Thus, it is important to note that the dollar cost measure reported herein is obtained from an aggregated mean across the entire school as opposed to individual student week-by-week billing which may vary based on attendance and nature of services provided.

LU per STO. Number of LU per STO is included in both the between-year and between-student comparison. The number of LU per STO reflects the rate of learning in that it is a direct measure of how many learning opportunities were presented until a student mastered a target objective. The rate of learning was measured by calculating the total number of LU a student received and dividing that number by the number of STOs mastered. This measure is reported in both the between-years and between-student analysis. For the between-years analysis, this measure was obtained by dividing the total number of LU completed (on both a weekly basis as well as aggregated across all 13 weeks of the analysis) for all students enrolled for that year and dividing that number by the total number of STOs mastered. For the between-students analysis the number of LU completed by all students in a VB category was divided by the total number of STOs mastered by all students in that category to obtain the LU per STO measure.

Long Term Objective (LTO)

Number of LTOs. The number of LTOs mastered is included in both the between-year and between-student comparison across groups of students with differing levels of VB development. A LTO consists of one complete repertoire selected from the ELCAR (Greer et al., 2020) curriculum. An example of an LTO would be for a student to identify 10 common colors as a speaker under generalized or prosthetic reinforcement conditions with 100% accuracy one time across all 10 colors.

A measure of LTOs mastered is included in both the between-year and between-student comparison. For the between-year analysis, the measure of LTOs mastered is the average number of LTOs mastered per student per day. This measure was obtained by dividing the total number of LTOs mastered schoolwide in a week and dividing it by number of students. This number was then divided by the number of instructional days in the school calendar that week

(LTOs/students enrolled/instructional days). This measure is reported as *average number of LTOs mastered per student per day.*

In the between-student comparison, the STO measure was obtained by calculating the total STOs mastered by all students in one VB category and dividing that number by the number of students in the category. This measure is reported as the *average total number of LTOs mastered per student in a category* throughout the weeks included in our analysis. In addition to the total LTOs mastered, the dollar cost per LTO was calculated for the between-year analysis as well as the between-student comparison. The dollar cost was calculated by obtaining the total revenue the school received and selecting out the percentage of the revenue that corresponded to the percentage of weeks of the year that were included in our analysis. This number was then divided by total number of LTOs mastered in our analysis for that year to determine the dollar cost per STO mastered in that year.

Learn Units Per LTO. Number of LU per LTO is included in both the between-year and between-student comparison. The number of LU per LTO reflects the rate of learning in that it is a direct measure of how many learning opportunities were presented until a student mastered an LTO. On a larger scale this rate of learning is reported by calculating the total number of LU a student received and dividing that number by the number of LTOs mastered. This measure is reported in both the between-years and between-student analysis. For the between-years analysis, this measure was obtained by dividing the total number of LU completed (on both a weekly basis as well as aggregated across all 13 weeks of the analysis) for all students enrolled for that year and dividing that number by the total number of LTOs mastered. For the between-students analysis the number or LU completed by all students in a VB category was divided by the total number of LTOs mastered by all students in that category to obtain the LU per LTO measure.

Design and Analysis

A between-groups comparison design (Brown & Melamed, 1990) was used in this study. Data collected from students receiving instruction fully in-person during 13 weeks in Fall 2015-2019 (excluding 2018) were compared to students receiving instruction in a hybrid model in the Fall 2020 for the between-years analysis. For the between-student analysis, data were collected from students across 13 weeks of instruction in Fall 2020. These students were then placed into categories by their level of VB development and data on their educational outcomes were compared.

Statistics

For the between-years analysis, a one-tailed independent samples t-test with significance set at <.05 was used to determine significance in difference between data collected in Fall 2020 to the average data collected from Fall 2015-2019 (excluding 2018) for measures of total LU, STO, and LTO. A two-tailed independent samples t-test with significance set at <.05 was used to compare results between rate of learning for LU to STO and LU to LTO. For the between-student analysis, an ANOVA with Bonferroni post-hoc analysis was used to compare variance between the groups of students grouped by level of VB.

Results

Between-Year Analysis

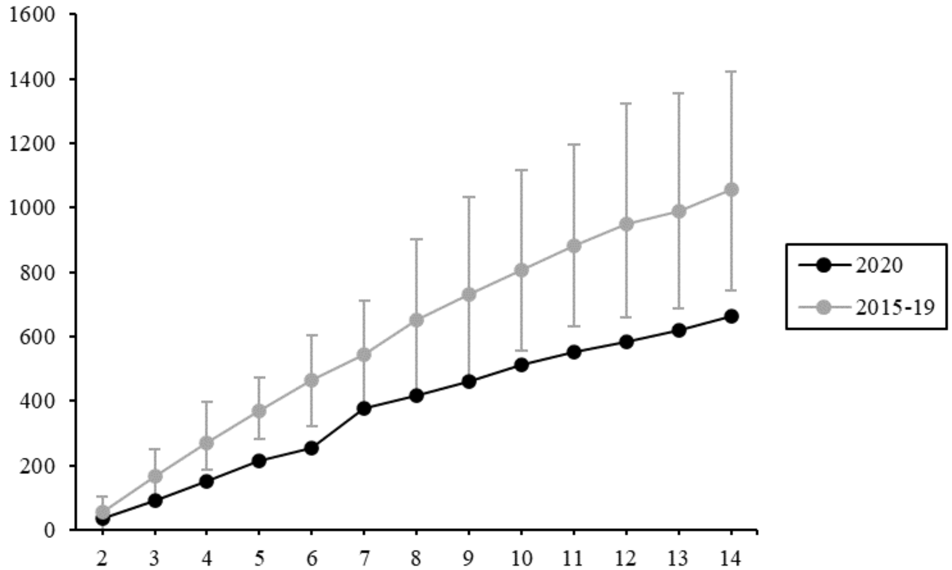
The first analysis compared data from Fall 2020 in a hybrid model of instruction compared to data collected during the same period in Fall 2015-2019 when the program was entirely in-person. Each year included 13 weeks of data spanning from the second week of each Fall semester through the 14^{th} week of the semester. What follows is a summary of the results as well as an interpretation of a visual analysis for each of our measures.

TPRA Observations Conducted

Figure 1 displays the cumulative TPRAs delivered across weeks 2-14 in each of the five years of data in our analysis. The grey data paths represent cumulative TPRAs delivered during in-person instruction during each Fall 2015-2019 and the black data path represents cumulative TPRAs during hybrid instruction delivered in 2020. For the four years of in-person instruction included in this analysis, there were a total of 958 TPRAs conducted in 2015, 1109 TPRAs conducted in 2016, 742 TPRAs conducted in 2017, 1421 TPRAs conducted in 2019. The average of total TPRAs conducted in across Fall 2015-2019 during in-person instruction was 1060 (range = 742 – 1421, SD 285). The total TPRAs conducted in Fall 2020 during the hybrid model with both in-person and remote instruction was 664. The total number of TPRAs delivered in 2020 was 1.46 standard deviations below the average of the previous 4 years. In terms of percentage, the number of TPRAs delivered in Fall 2020 was 37% fewer compared to the average across the preceding 4 years (664 in Fall 2020 compared to an average of 1060 in preceding 4 years).

Figure 1

A Graph of Cumulative TPRAs Delivered Weeks 2-14 for Fall 2015-2020



Note. The black data path represents TPRAs from the Fall 2020 hybrid model. The grey data path is for TPRAs from Fall 2015, Fall 2016, Fall 2017, and Fall 2019. Grey data points reflect the average while the errors bars represent the range of data from prior years

Fidelity of TPRA Observations

The percent of errorless TPRAs out of total TPRAs delivered was 86% (825 errorless out of 958 total TPRAs) in Fall 2015, 84% (936 errorless out of 1109 total TPRAs) in Fall 2016, 84% (622 errorless out of 742 total TPRAs) in Fall 2017, 88% in Fall 2019 (1246 errorless out of 1421 total TPRAs) for an average of 85% errorless (3629 errorless out of 4260 total TPRAs). This is compared to 87% correct TPRAs (580 errorless out of 664 total TPRAs) in Fall 2020. Thus, the average percent of correct TPRAs in Fall 2020 was comparable to the preceding 4 years in the comparison.

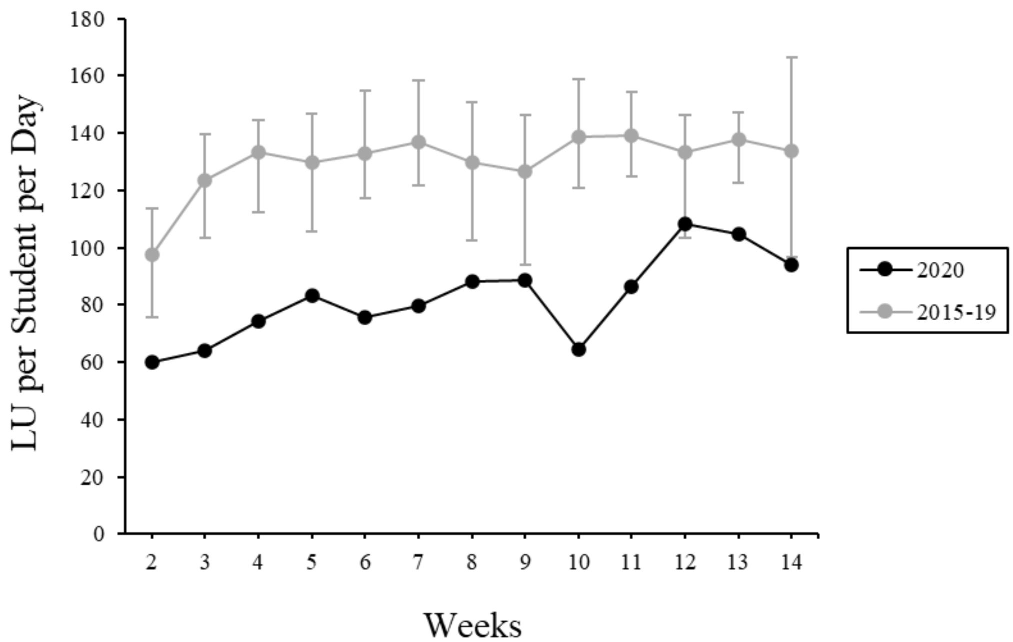
LU Per Day Per Student

Figure 2 displays the average number of learn units delivered per student per day across weeks 2-14 in the five years of data in our analysis. The grey data paths represent in-person instruction delivered in 2015-2019 and the black data path represents hybrid instruction delivered in 2020. For the four years of in-person instruction included in this analysis, there was an average of 105 (range = 76 - 125, SD = 14.5) LU delivered per student enrolled per instructional day in 2015, 136 (range = 97-167, SD = 16.9) LU per student per day in 2016, 124 (range = 104 – 149, SD = 11.6) LU per student per day in 2017, 135 (range = 97 – 158, SD = 16.8) LU per student per day in 2019. The average LU per student per day across all 4 years of fully in-person instruction was 125 (range = 75.6 – 166.5, SD = 19.5) LU per student per day. In Fall 2020, in a hybrid instructional model including both in-person and remote instruction there was an average of 79 (range = 61-110, SD = 14.9) LU delivered per student per day of instruction.

The average LU delivered per student per day in 2020 was 2.25 standard deviations below the average from the previous 4 years of data collected with fully in-person instruction. An analysis of the data using an independent samples *t*-test reveals that the average LU completed per student per day during in-person instruction across Fall of years 2015-2019 (M = 123, SD = 19.5) was significantly higher than the average LU completed per student per day in the hybrid model in Fall 2020 (M = 79, SD = 14.9), $t(12) = -9.4$, $p = .00001$. In terms of percentage, the average number of LU delivered per day per student in Fall 2020 was 37% fewer than the average of the preceding 4 years of in-person instruction (79 in Fall 2020 compared to average of 125 in preceding 4 years).

Figure 2

A Graph of LU Delivered per Student Per Day



Note. The black data path represents LU from the Fall 2020 hybrid model. The grey data path is for LU from Fall 2015, Fall 2016, Fall 2017, and Fall 2019. Grey data points reflect the average while the errors bars represent the range of data from prior years.

Dollar Cost Per LU

Table 2 displays the average dollar cost per LU delivered across years 2015-2020. The average dollar cost per LU delivered was $2.43 in Fall 2015, $1.92 in Fall 2016, $2.14 in Fall 2017, $2.39 in Fall 2019 for an average of $2.22. This is compared to the cost per LU delivered in 2020 of $4.85. Thus, the cost per LU delivered in this hybrid educational model was 2.18 times the cost of a LU delivered during fully in-person instruction.

Table 2

Dollar cost per learning measure per year.

Measure	2015	2016	2017	2019	2020
LU	$2.43	$1.92	$2.14	$2.39	$4.85
STO	$409.75	$351.82	$413.30	$431.76	$891.77
LTO	$3932.84	$3369.77	$3875.83	$3799.85	$7686.69

STOs

Number of STOs Met. The top panel in Figure 3 displays the average number of STOs mastered per student per day across weeks 2-14 in the five years of data in our analysis. The grey data paths represent in-person instruction delivered in 2015-2019 and the black data path represents hybrid instruction delivered in 2020. For the four years of in-person instruction included in this analysis, there was an average of 0.61 (range = 0.20 – 0.97, SD = 0.21) STOs mastered per student enrolled per instructional day in 2015, 0.73 (range = 0.37 – 2.08, SD = 0.28) STOs mastered per student per day in 2016, 0.63 (range = 0.33 – 0.84, SD = 0.13) STOs mastered per student per day in 2017, 0.74 (range = 0.45 – 0.92, SD = 0.14) STOs mastered per student per day in 2019. The average STOs mastered per student per day across all 4 years of fully in-person instruction was 0.68 (range = 0.20 – 1.08, SD = 0.17) STOs mastered per student per day. In Fall 2020, in a hybrid instructional model including both in-person and remote instruction there was an average of 0.42 (range = 0.12 - 0.66, SD = 0.15) STOs mastered per student per day of instruction.

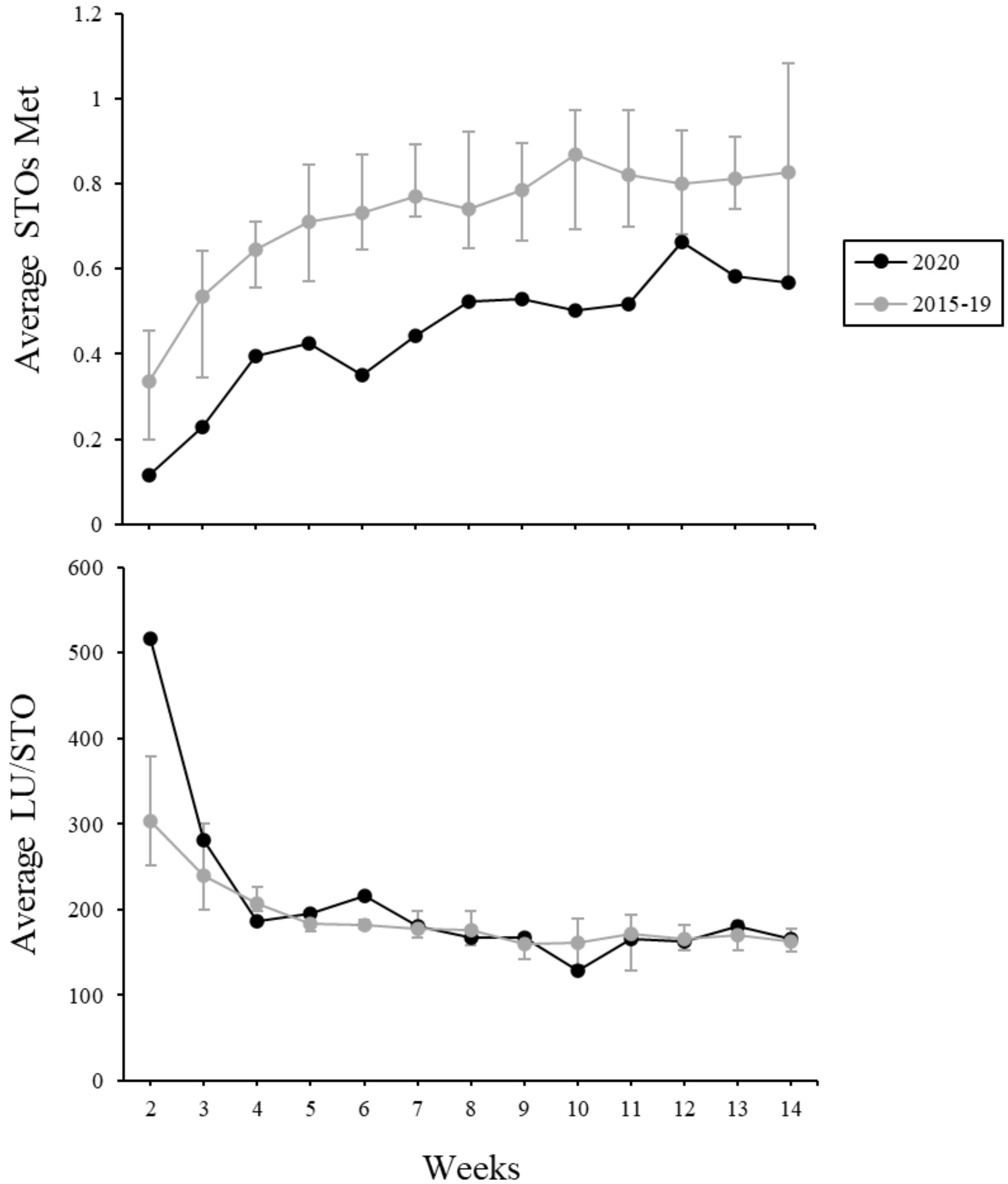
The average STOs mastered per student per day in 2020 was 1.47 standard deviations below the average from the previous 4 years of data collected with fully in-person instruction. An analysis of the data using an independent samples *t*-test reveals that the average STOs mastered per student per day during in-person instruction across Fall of years 2015-2019 (M = 0.68, SD = 0.17) was significantly more than the average STOs mastered per student per day in the hybrid model in Fall 2020 (M = 0.42, SD = 0.15), $t(12) = -3.3$, $p = .001$

LU Per STO. The bottom panel in Figure 3 displays the average number of LU delivered per STO mastered across weeks 2-14 in the five years of data in our analysis. The grey data paths represent in-person instruction delivered in 2015-2019 and the black data path represents hybrid instruction delivered in 2020. For the four years of in-person instruction included in this analysis, there was an average of 187 (range = 128 - 379, SD = 73) LU delivered per STO mastered in 2015, 187 (range = 154 – 264, SD = 29) LU delivered per STO mastered in 2016, 199 (range = 166 – 318, SD = 41) LU delivered per STO mastered 2017, 185 (range = 158 – 251, SD = 26) LU delivered per STO mastered in 2019. The average LU per STO mastered across all 4 years of fully in-person instruction was 189 (range = 128 – 379, SD = 45) LU delivered per STO mastered. In Fall 2020, in a hybrid instructional model including both in-person and remote instruction there was an average of 209 (range = 129 - 517, SD = 99) LU delivered per STO mastered.

The average number of LU delivered per STO mastered in 2020 was 0.44 standard deviations below the average from the previous 4 years of data collected with fully in-person instruction. An analysis of the data using an independent samples *t*-test reveals that the average LU per STO mastered during in-person instruction across Fall of years 2015-2019 (M = 189, SD = 45) was not significantly more than the average LU delivered per STO mastered in the hybrid model in Fall 2020 (M = 209, SD = 99), $t(12) = -0.7$, p = .52. In terms of percentage, the average number of STO mastered per day per student in Fall 2020 was 38% fewer than the average of the preceding 4 years of in-person instruction (average of .42 in Fall 2020 compared to average of .68 in preceding 4 years). For the LUSTO measure, the average number of LU required for a student to master an STO in Fall 2020 was 11% higher than preceding 4 years (average of 209 in Fall 2020 compared to average of 189 in preceding 4 years).

Figure 3

Number of STOs Mastered (top panel) and LU per STO Mastered (bottom panel)



Note. The black data path represents data from the Fall 2020 hybrid model. The grey data path is for data from Fall 2015, Fall 2016, Fall 2017, and Fall 2019. Grey data points reflect the average while the errors bars represent the range of data from prior years.

Dollar Cost Per STO. Table 1 displays the average dollar cost per STO mastered across years 2015-2020. The average dollar cost per STO mastered was $409.75 in Fall 2015, $351.82 in Fall 2016, $413.30 in Fall 2017, $431.76 in Fall 2019 for an average of $401.66 per STO mastered. This is compared to the cost per STO mastered in 2020 of $891.77. Thus, the cost per STO mastered in this hybrid educational model was 2.22 times the cost of a STO mastered during fully in-person instruction.

LTOs

Number of LTOs Met. The top panel in Figure 4 displays the average number of LTOs mastered per student per day across weeks 2-14 in the five years of data in our analysis. The grey data paths represent in-person instruction delivered in 2015-2019 and the black data path represents hybrid instruction delivered in 2020. For the four years of in-person instruction included in this analysis, there was an average of 0.07 (range = 0.02 – 0.15, SD = 0.04) LTOs mastered per student enrolled per instructional day in 2015, 0.08 (range = 0.02 – 0.12, SD = 0.03) LTOs mastered per student per day in 2016, 0.07 (range = 0.03 – 0.03, SD = 0.02) LTOs mastered per student per day in 2017, 0.07 (range = 0.03 – 0.11, SD = 0.02) LTOs mastered per student per day in 2019. The average LTOs mastered per student per day across all 4 years of fully in-person instruction was 0.073 (range = 0.02 – 0.15, SD = 0.028) LTOs mastered per student per day. In Fall 2020, in a hybrid instructional model including both in-person and remote instruction, there was an average of 0.047 (range = 0.01 - 0.09, SD = 0.02) LTOs mastered per student per day of instruction.

The average LTOs mastered per student per day in 2020 was 0.90 standard deviations below the average from the previous 4 years of data collected with fully in-person instruction. An analysis of the data using an independent samples *t*-test reveals that the average LTOs mastered per student per day during in-person instruction across Fall of years 2015-2019 (*M* = 0.073, *SD* = 0.028) was significantly more than the average LTOs mastered per student per day in the hybrid model in Fall 2020 (*M* = 0.047, *SD* = 0.02), $t(12) = 3.06$, $p = .003$.

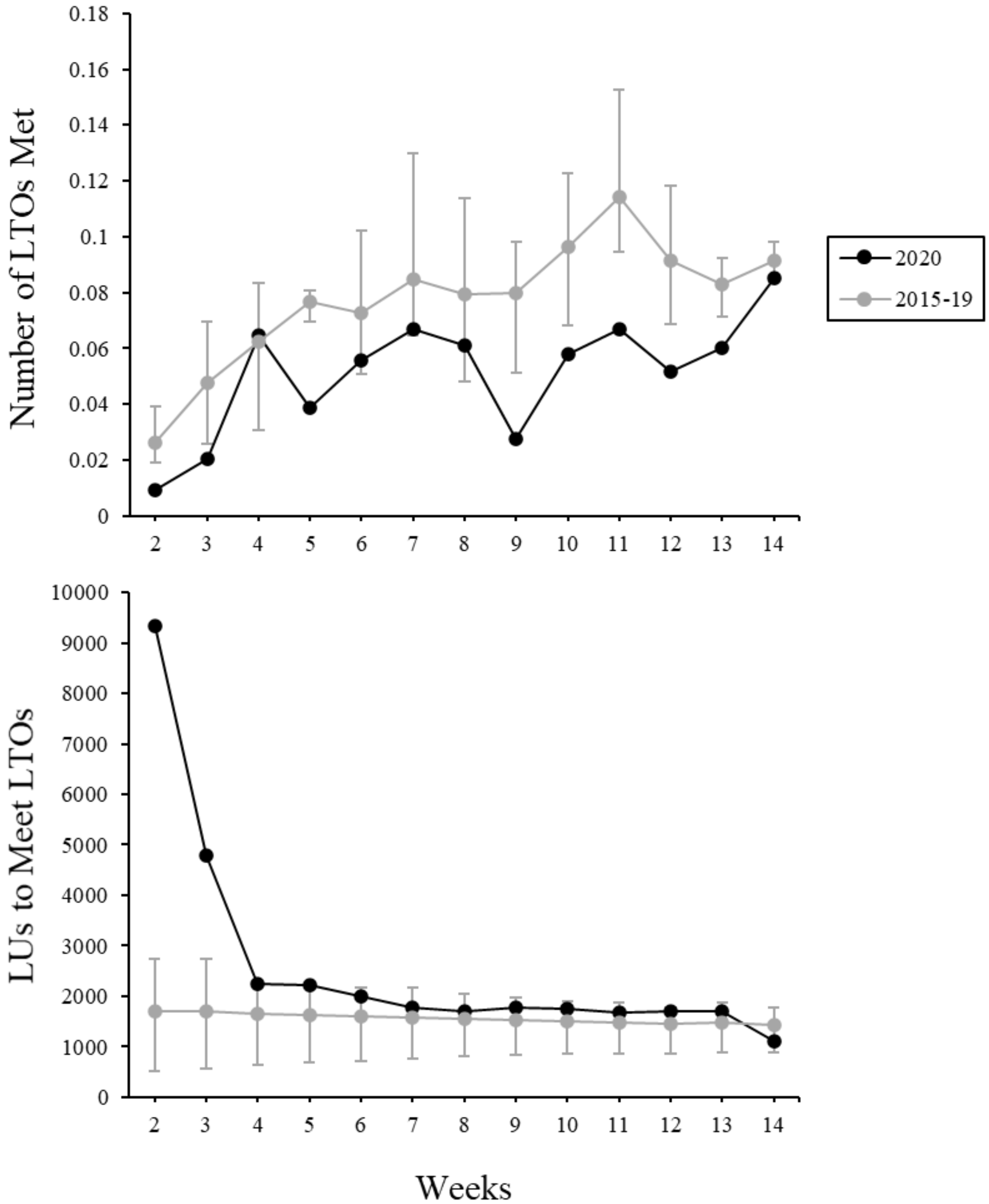
LU Per LTO. The bottom panel in Figure 4 displays the average number of LU delivered per LTO mastered across weeks 2-14 in the five years of data in our analysis. The grey data paths represent in-person instruction delivered in 2015-2019 and the black data path represents hybrid instruction delivered in 2020. For the four years of in-person instruction included in this analysis, there was an average of 757 (range = 505 - 895, SD = 126) LU delivered per LTO mastered in 2015, 2027 (range = 1771 – 2248, SD = 171) LU delivered per LTO mastered in 2016, 1501 (range = 1333 – 1593, SD = 84) LU delivered per LTO mastered 2017, 1961 (range = 1604 – 1750, SD = 405) LU delivered per LTO mastered in 2019. The average LU per LTO mastered across all 4 years of fully in-person instruction was 1560 (range = 505 – 2750, SD = 559) LU delivered per LTO mastered. In Fall 2020, in a hybrid instructional model including both in-person and remote instruction there was an average of 2597 (range = 1103 - 9324, SD = 2203) LU delivered per LTO mastered.

The average LU per LTO mastered in 2020 was 1.86 standard deviations above the average from the previous 4 years of data collected with fully in-person instruction. An analysis of the data using an independent samples *t*-test reveals that the average LU per LTO mastered per student per day during in-person instruction across Fall of years 2015-2019 (*M* = 2597, *SD* = 2203) was not significantly more than the average LU delivered per LTO mastered in the hybrid model in Fall 2020. In terms of percentage, the average number of LTO mastered per day per student in Fall 2020 was 36% fewer than the average of the preceding 4 years of in-person instruction (average of .047 in Fall 2020 compared to average of .0073 in preceding 4 years). For the LULTO measure, the average number of LU required for a student to master an LTO in Fall 2020 was 66% higher than preceding 4 years (average of 2597 in Fall 2020 compared to average of 1560 in preceding 4 years).

It is important to note that when removing the first two weeks from the analysis and only including the final 11 weeks, LULTO was only 14% higher during remote instruction compared to in-person instruction. Thus, this indicates that for LULTO the first few weeks were far less efficient but that by Week 4 the rate of learning for LULTO was within one standard deviation to the mean of the preceding four years of in-person instruction.

Figure 4

Number of LTOs Mastered (top panel) and LU per LTO Mastered (bottom panel)



Note. The black data path represents data from the Fall 2020 hybrid model. The grey data path is for data from Fall 2015, Fall 2016, Fall 2017, and Fall 2019. Grey data points reflect the average while the errors bars represent the range of data from prior years.

Dollar Cost Per LTO. Table 2 displays the average dollar cost per LTO mastered across years 2015-2020. The average dollar cost per LTO mastered was $3932.84 in Fall 2015, $3369.77 in Fall 2016, $3875.83 in Fall 2017, $3799.85 in Fall 2019 for an average of $3744.57 per LTO mastered. This is compared to the cost per LTO mastered in 2020 of $7686.69. Thus, the cost per LTO mastered in this hybrid educational model is 2.05 times the cost of a LTO mastered during fully in-person instruction.

Between-Student Analysis of Hybrid Model of Instruction

This analysis compares educational outcomes for students classified by levels of verbal behavior who attended the program during Fall 2020 for hybrid instruction. The hybrid model included 1 week of 4 days of in-person instruction plus one day remote and 1 week of fully remote instruction. Students were grouped in five groups: Pre-listener (Group 1, *N* = 28, listener/pre-speaker (Group 2, *N* = 18), listener/speaker who demonstrated NiN (Group 3, *N* = 30), listener/speaker who demonstrated UniN (Group 4, *N* = 18), and listener/speaker who demonstrated BiN (Group 5, *N* = 14). The means and standard deviations for each measure were calculated from data collected across all students within each group. What follows is a summary of the results for each measure.

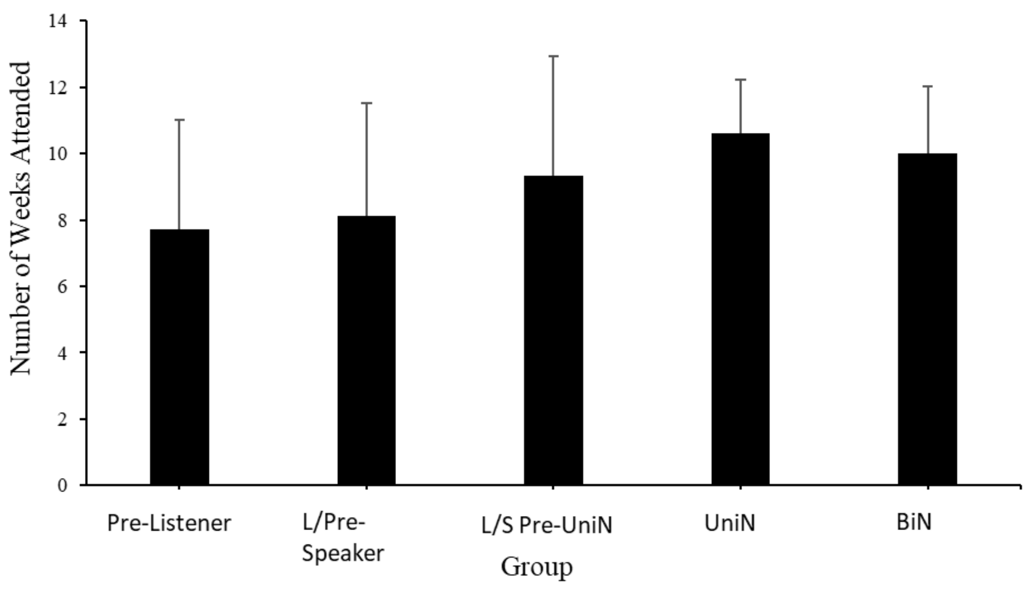
Number of Weeks Attended

Figure 5 displays the mean number of weeks attended per student in each group in total across both in-person and remote modalities. Students were scored as receiving instruction in a week if they received at least on instructional session in a week (Monday-Friday). If the student did not receive any instruction they were scored as not having attended that week of instruction. The students across in Groups 1, 2, 3, 4, and 5 received instruction across a mean of 7.7 weeks (range = 1-12, SD = 3.3), 8.1 weeks (range = 2–12, SD = 3.4), 9.3 weeks (range = 1–12, SD = 3.6), 10.6 weeks (range = 7–12, SD = 1.6), and 10 weeks (range = 6-12, SD = 2.0), respectively.

An analysis of the data using a one-way ANOVA with a Bonferroni post-hoc analysis reveals that the average number of weeks attended differed significantly per group of students, $F(4, 103) = 3.33$, $p = .013$. Post-hoc comparisons revealed a significant difference between Group 1 (pre-listener) and Group 4 (UniN) ($p = .022$) in average number of weeks attended and no other comparisons were different.

Figure 5

Number of Weeks Attended Per Group



Note. L/Pre-Speaker refers to students operating on the Listener/Pre-Speaker level of VB. The error bars represent the range within 1 SD away from the mean.

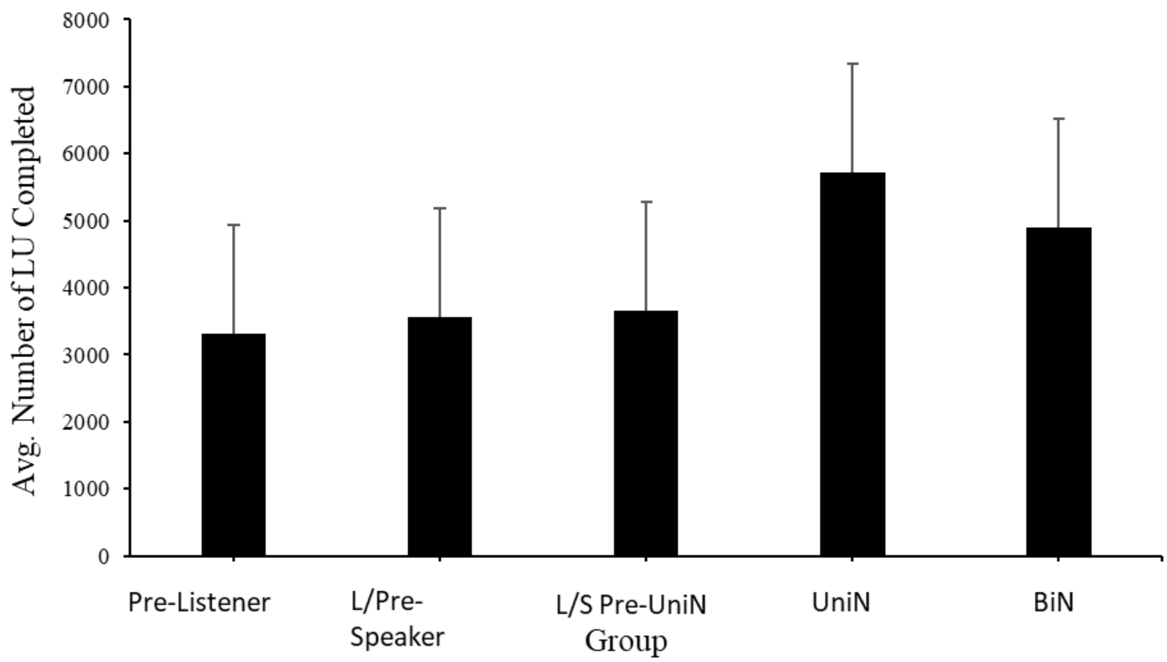
Number of LU Received

Figure 6 displays the mean number of LU received per student in each group across all 13 weeks included in this analysis. The students in Groups 1, 2, 3, 4, and 5, received a mean of 3304 LU (range = 30-8533, SD = 2164.0), 3567 LU (range = 686-9777, SD = 2767), 3664 LU (range = 42- 9422, SD = 2652), 5714 LU (range = 2266 – 11976, SD = 1961), and 4889 LU (range = 1356 – 7490, SD = 1623) respectively across all 13 weeks of school included in this analysis.

An analysis of the data using a one-way ANOVA with a Bonferroni post-hoc analysis reveals that the average number of LU received differed significantly per group of students, $F(4, 103) = 3.86$, $p = .006$. Post-hoc comparisons revealed a significant difference between Group 4 (UniN) and Groups 1 (pre-listener) and 3 (L/S pre-UniN), and 1 (pre-listener) ($p = .009$ and .04) in average number of LU received and no other comparisons were different.

Figure 6

Average Number of LU Completed per Group



Note. L/Pre-Speaker refers to students operating on the Listener/Pre-Speaker level of VB. The error bars represent the range within 1 SD away from the mean.

Table 3 displays the cost per LU delivered for each category of students in Fall 2020. The average dollar cost per LU received was $7.02 per LU for Group 1 (pre-listener) , $6.51 per LU for Group 2 (listener/pre-speaker), $6.33 per LU for Group 3 (L/S pre-UniN, $4.06 per LU for Group 4 (L/S with UniN), and $4.75 per LU for Group 5 (L/S with BiN).

Table 3

Dollar Cost Per Learning Measure Per Group of Students

Group	$LTO	$LU	$STO
Pre-Listener	$12257.89	$7.02	$1804.63
Listener/ Pre-Speaker	$9943.90	$6.51	$1481.01
L/S Pre-UniN	$9667.68	$6.33	$1152.44
L/S UniN	$8523.34	$4.06	$662.93
L/S BiN	$7922.78	$4.75	$565.91

Short Term Objectives

Number of STOs Mastered. The top panel of Figure 7 displays the mean number of STOs mastered per student in each group across all 13 weeks included in this analysis. The students in Groups 1 (pre-listener), 2 (listener/pre-speaker), 3 (L/S pre-UniN), 4 (L/S with UniN), and 5 (L/S with BiN) mastered a mean of 12.9 STOs (range = 0-43, SD = 11.4), 15.7 STOs (range = 1-37, SD = 11.1), 20.1 STOs (range = 0- 49, SD = 13.9), 35 STOs (range = 13-59, SD = 12.2), and 41 STOs (range = 5-72, SD = 22.0) respectively across all 13 weeks of school included in this analysis.

An analysis of the data using a one-way ANOVA with a Bonferroni post-hoc analysis reveals that the average number of STOs mastered differed significantly per group of students, $F(4, 103) = 14.49$ $p = .000$. Post-hoc comparisons revealed a significant difference between Groups 4 (UniN) and 5 (BiN) and all other groups (range of p up to .005) in number of STOs mastered. No significant difference was found between Groups 1-3 (all pre-UniN) or between Groups 4 (UniN) and 5 (BiN).

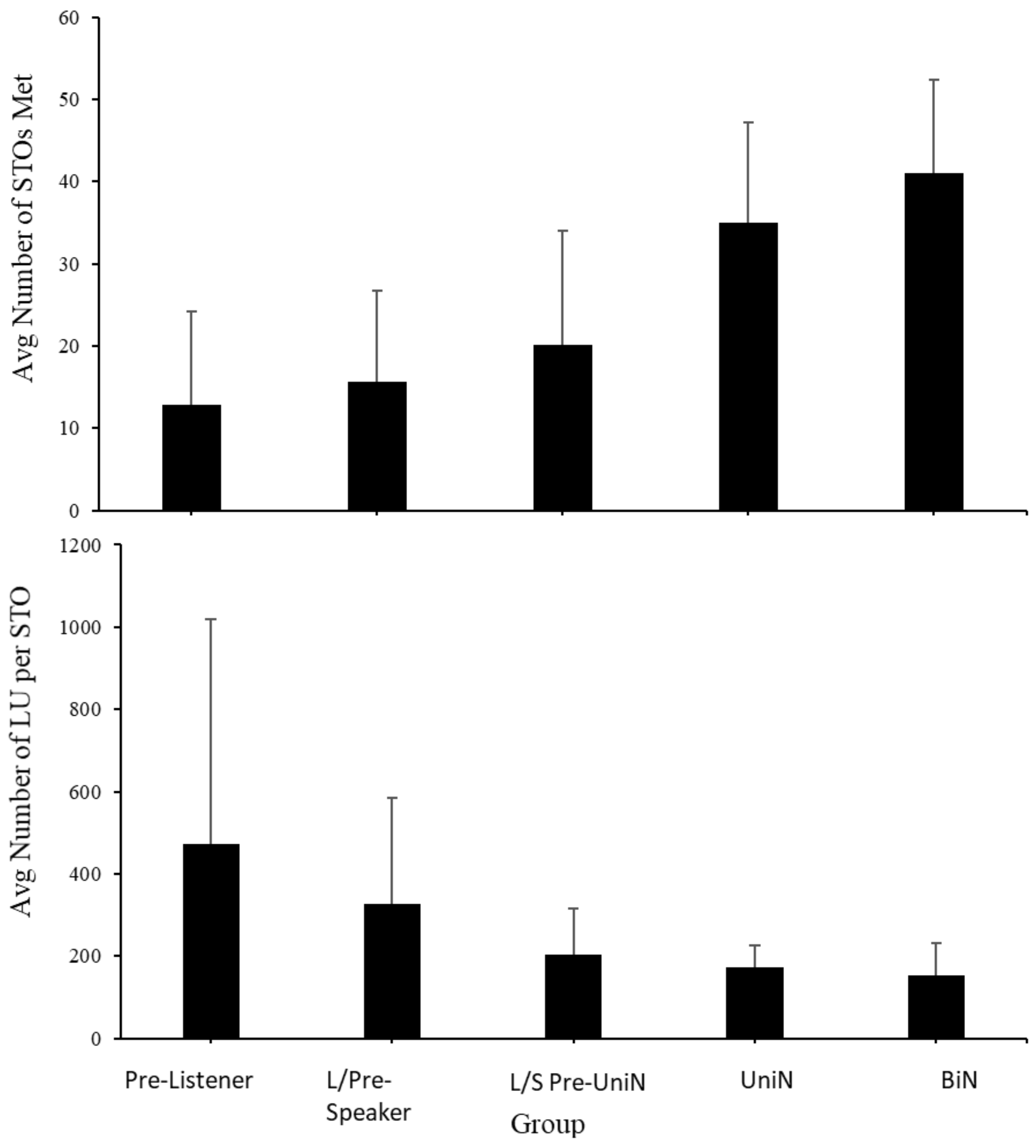
Table 2 displays the cost per STO mastered for each category of students in Fall 2020. The average dollar cost per STO mastered was $1804.63 per STO for the pre-listener group, $1481.01 per STO for the listener/pre-speaker group, $1152.44 per STO for the L/S pre-UniN group, $662.93 per STO for the L/S with UniN group, and $565.91 per STO for the L/S with BiN group.

Number of LU per STO Mastered. The rate of learning for each STO is measured by the number of LU required until criterion for mastery was met. The bottom panel of Figure 4 displays the number of LUs delivered per STO mastered across all students in the group with the SD calculated from the average LU to STO for each student. The students in Groups 1 (pre-listener), 2 (listener/pre-speaker), 3 (L/S pre-UniN), 4 (L/S with UniN), and 5 (L/S with BiN) required an average of 257 LU per STO mastered (range = 0-2558, SD = 257.0), 227.7 LU per STO mastered (range = 68-1152, SD = 255.8), 182.0 LU per STO mastered (range = 0-608, SD = 111.3), 163.3 LU per STO mastered (range = 67-271, SD = 51.4, and 119.2 LU per STO mastered (range = 51-288, SD = 76.8) respectively across all 13 weeks of school included in this analysis.

An analysis of the data using a one-way ANOVA with a Bonferroni post-hoc analysis reveals that the average number of LU per STO differed significantly per group of students, $F(4, 103) = 4.44$, $p = .002$. Post-hoc comparisons revealed a significant difference between Group 1 (pre-listener) and Groups 3 (L/S pre-UniN), 4 (L/S with UniN), and 5 (L/S with BiN) (range of $p = .01 - .02$) with no other differences between groups.

Figure 7

Number of STOs and LU per STOs per Group



Note. L/Pre-Speaker refers to students operating on the Listener/Pre-Speaker level of VB. The error bars represent the range within 1 SD away from the mean.

LTO

Number of LTOs Mastered. The top panel of Figure 8 displays the mean number of LTOs mastered per student in each group across all 13 weeks included in this analysis. The students in Groups 1 (pre-listener), 2 (listener/pre-speaker), 3 (L/S pre-UniN), 4 (L/S with UniN), and 5 (L/S with BiN) mastered an average of 1.9 LTOs (range = 0-20, SD = 4.5), 2.3 LTOs (range = 0-1, SD = 3.5), 2.4 LTOs (range = 0-15, SD = 3.4), 2.7 LTOs (range = 0-9, SD = 2.7), and 2.9 LTOs (range = 0-7, SD = 2.1) respectively across all 13 weeks of school included in this analysis.

An analysis of the data using a one-way ANOVA with a Bonferroni post-hoc analysis reveals that the average number of LTOs mastered did not differ significantly per group of students.

Table 2 displays the dollar cost per LTO mastered across each group of students. The average dollar cost per LTO mastered was $12,257.89 per LTO for the pre-listener group, $9,943.90 per LTO for the listener/pre-speaker group, $9,667,68 per LTO for the listener/speaker with pre-UniN group, $8,523.34 per LTO for the UniN group, and $7,922.78 per LTO for the BiN group.

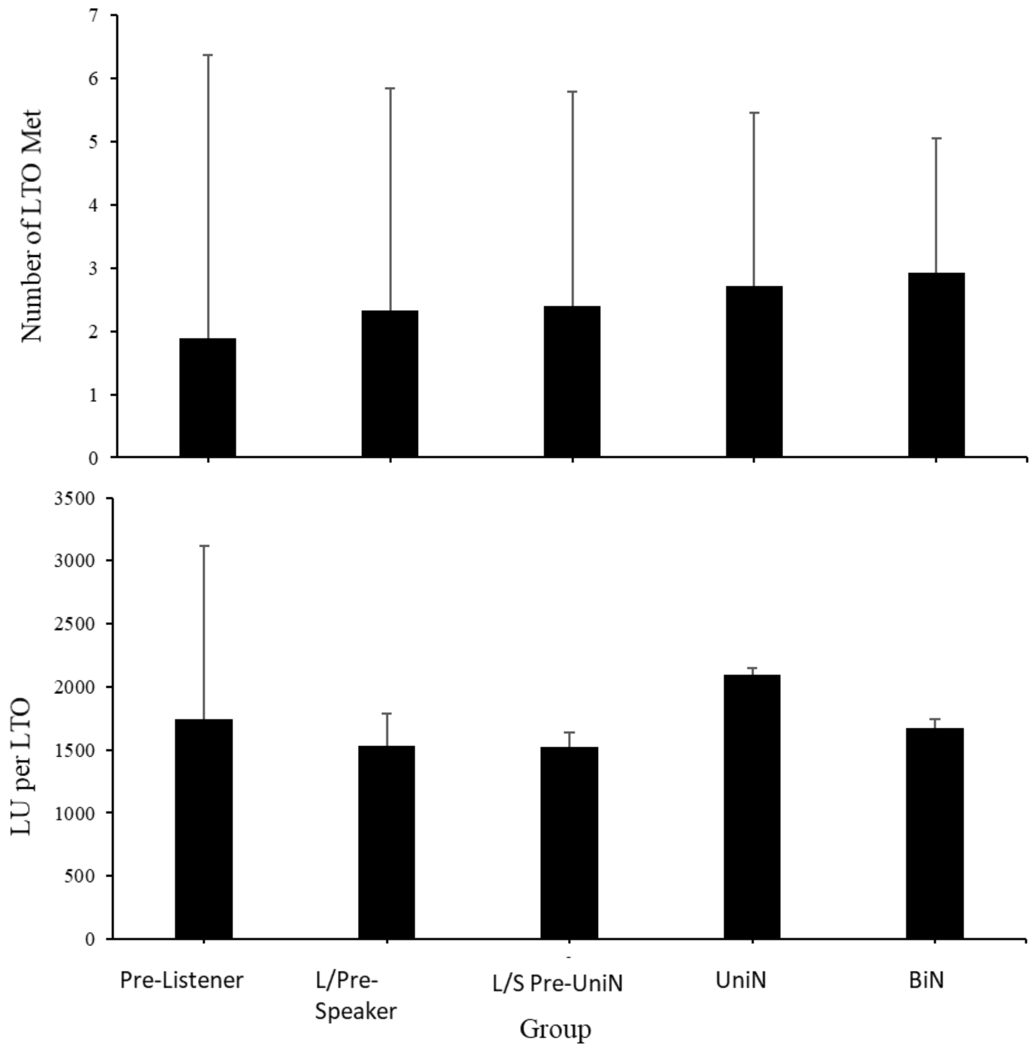
Number of LU per LTO Mastered. The rate of learning for each LTO is measured by the number of LU required until criterion for mastery was met for the LTO. The bottom panel of Figure 8 displays the number of LUs delivered per LTO mastered across all students in the group with the SD calculated from the average LU to LTO for each student. The students in Groups 1 (pre-listener), 2 (listener/pre-speaker), 3 (L/S pre-UniN), 4 (UniN), and 5 (BiN) required an average of 1745.6 LU per LTO mastered (range = 0-5377, SD = 547.5), 1528.6 LU per LTO mastered (range = 0-5490, SD = 255.8), 1526.7 LU per LTO mastered (range = 0-5075, SD =

111.3), 2099.0 LU per LTO mastered (range = 0-11976, SD = 51.4), and 1669.4 LU per LTO mastered (range = 0-5950, SD = 76.8) respectively across all 13 weeks of school included in this analysis.

An analysis of the data using a one-way ANOVA with a Bonferroni post-hoc analysis reveals that the average number of LU per LTO differed significantly per group of students, $F(4, 103) = 4.57$ $p = .002$. Post-hoc comparisons revealed a significant difference between Group 4 (UniN) and Groups 1 (pre-listener) and 3 (L/S pre-UniN) ($p = .001$ and .013 respectively) in average number of LU per LTO and no other comparisons were different.

Figure 8

Average Number of LTO and LU per LTO per Group



Note. L/Pre-Speaker refers to students operating on the Listener/Pre-Speaker level of VB. The error bars represent the range within 1 SD away from the mean.

Discussion

To evaluate the effects of including a remote learning component to the delivery of a CABAS® model of instruction, I compared educational outcomes achieved during the hybrid model with outcomes from the previous four years of in-person instruction. Further, to determine the relationship between level of VB development and educational outcomes in a hybrid learning model I grouped students by levels of VB and compared educational outcomes across those groups.

In the between-year analysis, the results indicate that in the hybrid model nearly all learning outcomes were decreased while the rate of learning was not significantly different. Specifically, the total number of TPRAs conducted, the average LU delivered per student per day, average STOs met, and average LTOs were all significantly fewer than those outcomes during the in-person comparison in the preceding four years. For the rate of learning, the average LU per STO and LU per LTO in the hybrid model were not significantly greater than LU per STO in the in-person model. Thus, although the total learning opportunities (LU) and learning objectives mastered (STO and LTO) were significantly higher during in-person instruction, no significant difference in rate of learning occurred during the hybrid model compared to in-person.

On the system-wide level, the average number of errorless TPRA observations conducted during the hybrid model was 87% compared to an average of 85% during the in-person model indicating that the average instructional fidelity during the hybrid model was comparable to the in-person model. Finally, the dollar cost per LU delivered, STO and LTO mastered in the hybrid model was nearly double the cost for each measure when compared to the in-person model.

In the between-student analysis I found that the differences among educational outcomes were most pronounced when comparing students with some degree of Naming to those who demonstrated no degree of Naming, or comparing outcomes for students with UniN (Group 4) or BiN (Group 5) to outcomes for students with pre-UniN (Groups 1, 2, and 3). Specifically, the results reveal that students with UniN attended significantly more weeks of instruction than students operating on a pre-listener level of VB and received significantly more LU than students on pre-listener and listener/speaker with pre-UniN levels of VB. Students with BiN and UniN mastered significantly more STOs than all students who demonstrated pre-UniN and learned at a significantly faster rate than students at the pre-listener level of VB development. Students with UniN mastered LTOs with significantly fewer LU when compared to students on the pre-listener and listener/speaker with pre-UniN levels of VB. Of note, there was no significant difference between level of VB and total number of LTOs mastered. Finally, the dollar cost was 50% more per LU delivered (average of $6.62 compared to $4.41), 140% more per STO mastered (average of $1,479 compared to $614), and 29% more per LTO mastered (average of $10,623 compared to $8223) for students with pre-UniN compared to students with UniN and BiN.

Addition to Existing Literature

Between-Years Comparison

The findings from the between-years analysis add to the existing literature relating to educational outcomes during remote instruction. In terms of educational outcomes during remote instruction, there have been few published studies reporting actual outcomes from in-person school closures due to COVID-19. One study by Engzell et al., (2020) found that the average learning loss measured by standardized testing at the end of the 2019-2020 school year across 350,00 participants in the Netherlands was equivalent to a fifth of a school year – nearly identical to the period of time for which schools were closed to in-person instruction. Compared to this report, measured by average performance on a standardized test, the findings from this experiment indicate that a hybrid delivery of the CABAS® model of instruction with 60% of instruction delivered remotely resulted in a decrease of 37% fewer LU delivered per student per day, 38% fewer STOs mastered per student per day, and 36% fewer LTOs mastered per student per day. This indicates that fewer overall learning opportunities are being presented which results in fewer objectives being mastered. It is important to note that the rate of learning, measured as number of LU delivered per STO were not significantly different (10% more LU per STO) remotely during remote instruction compared to in-person. For LULTO, the rate of learning was far slower in the first two weeks of our comparison but increasing drastically from Week 4 and on. Although there is no statistically significant difference when including Weeks 2-14 or 4-14, it is important to highlight that the rate of learning for LULTO, when including weeks 2-14, was 66% slower remotely. When including only weeks 4-14 the number drops drastically to being only 14% slower remotely. This indicates that the although students still learned as quickly during remote instruction despite the overall learning opportunities (and objectives) being fewer, from Week 4 and on the rate of learning remotely was very similar (within one standard deviation) for LULTO as well as LUSTO.

Compared to the findings of Engzell et al., (2020) these findings indicate greater success in remote instruction using the CABAS® model compared to the models in the Netherlands schools included in that study. Specifically, while Engzell et al., (2020) found that deficits in performance on a standardized test (nearly 20% lower scores) were almost identical to the percentage of the school year for which in-person instruction was closed (~20%), our results indicate that delivering 60%/40% of instruction remotely/in-person resulted in a decrease of 37% of learning. This indicates that delivery of instruction in a 60% remote and 40% in-person hybrid model was 63% as effective overall compared to fully in-person instruction.

Further, Kuhfeld et al., (2020) used a hierarchical linear model to predict that students were expected to lose an average of 63%-68% of gains in reading and 37%-50% of gains in mathematics due to school closures for COVID-19 in the 2019-2020 school year. Compared to the predictions of Kuhfeld et al., the CABAS® hybrid model found that students lost an average of 36%-38% educational outcomes in a hybrid model compared to a fully in-person model. This experiment did not select out math or reading skills specifically but included them in the overall educational outcomes measured. This preliminary analysis suggests that, on average, students receiving instruction in the hybrid CABAS® model outperformed predictions from Kuhfeld et al., (2020).

Between-Students Comparison

The findings from the between-students comparison revealed that, overall, significant differences in educational outcomes began to appear once students acquired some degree of the Naming capability. Specifically, the only significant difference between pre-listeners, listener/pre-speaker, and listener/speaker with pre-UniN (all who demonstrated no degree of Naming) in our measures was for LU per STO mastered in which listener/speakers with pre-UniN required significantly ($p = .02$) fewer LU per STO mastered. This is compared to significant differences between students with UniN or BiN and students from Groups 1-3 (all pre-UniN) in measures of weeks attended, average number of LU received per student, number of STOs mastered, number of LU per STO, and number of LU per LTO. In other words, there were significant differences between students who had demonstrated some degree of Naming (either UniN or BiN) and students who demonstrated NiN in every single measure except for number of LTOs met.

These findings demonstrate that students who demonstrate some degree of Naming (UniN or BiN) benefit from remote instruction in ways that students with pre-UniN do not. This builds on existing research demonstrating how the presence of Naming allows a student to learn in new ways. Specifically, Morgan et al., (2020) found that BiN is related to increased scores on tests of derived relational responding while students who did not demonstrate Naming scored lower. Greer et al., (2011a) found that students with Naming could learn from instructional demonstrations alone while students who did not have Naming required direct consequences and did not learn from instructional demonstrations. The findings of this experiment identify another form of instruction for which the Naming capability allows for improved learning.

Limitations and Future Research

One limitation of this experiment is that the educational outcomes were aggregated across all curricular domains and domain-specific outcomes (e.g. math, reading, or communication) were not reported. This precludes a more detailed comparison of outcomes from this model to the predictive model reported by Kuhfeld et al., (2020) which predicted an average loss of 63%-68% of gains in reading and 37%-50% of gains in mathematics due to school closures for COVID-19 in the 2019-2020 school year. Although the data reported in this experiment are compared and found to report improved overall educational outcomes than predicted by Kuhfeld et al., the lack of math or reading domain-specific data does not allow for a conclusive comparison. Future analysis would benefit from reporting domain-specific outcomes to allow for comparison to predictive models such as Kuhfeld et al., (2020).

A limitation in the between-year analysis for the TPRA measure relates to student enrollment. The number of students enrolled in the hybrid model in Fall 2020 was about 20 (~16%) fewer students than over the previous years. Thus, although there were fewer TPRAs delivered, that number may be influenced by number of students enrolled and receiving instruction. Despite this difference, it is important to note that the number of TPRAs delivered is a staff training and fidelity measure that should not be affected by the number of students.

It is important to note that the dollar cost measures reported are obtained by taking total school revenue and dividing it by outcomes reported. As billing for early intervention services differs from preschool special education services, and enrollment does vary slightly from week to week during the fall semester, the numbers we report reflect aggregated means as opposed to specific per-student payments.

A limitation in the between-student analysis is that the number of students in each group was not identical. Specifically, only 18 students were included in Group 4 (UniN) and 13 students in Group 5 (BiN). This is compared to 28 students in Group 1 (pre-listener), 16 students in Group 2 (listener/pre-speaker), and 25 students in Group 3 (listener/speaker with NiN). Thus, for example, the smaller sample of students in Group 5 places greater weight on the variation in data from one student which may be a result of outside variables. Further analysis can hold constant the number of students in each group as well as increase the sample included within each group to strengthen the external validity of these findings.

A final limitation in this experiment rests in number of weeks of instruction from which data were available for this analysis. The data indicate that far fewer learning opportunities and mastery of objectives were achieved in the first few weeks of this hybrid model, but the number of learning opportunities and objectives mastered increased as the students and teachers became more familiar with this model. To better compare the educational outcomes from a hybrid model of instruction to a fully in-person model, comparing data across a greater number of weeks (e.g. an entire school year) would be beneficial.

Conclusion

The findings from the between-year analysis indicate that, overall, educational outcomes from a hybrid model were decreased when compared to educational outcomes from an in-person model. However, the rate of learning was not significantly different in the remote model compared to the in-person model. Notably, the outcomes from the model in this experiment outperformed predictions for educational outcomes due to school closures from Kuhfeld et al., (2020) as well as performance on standardized tests from Engzell et al., (2020). This demonstrates that, although the overall learning outcomes were decreased in a hybrid extension of the CABAS® model, the outcomes were not as decreased as reported or predicted by other models. The findings from between-student analyses indicate that students with some degree of Naming benefit more from learning in a hybrid model compared to students who demonstrate no degree of Naming. In addition to other beneficial outcomes reported from establishment of Naming (e.g., increased rate of learning, learning from instructional demonstrations, learning incidentally), these findings support the importance of establishing the Naming capability for improved student outcomes.

Chapter IV

General Discussion

Remote Delivery of Services

The recent changes in how educational and ABA services are delivered due to COVID-19 opened up an opportunity to examine new modalities for how these services are to be delivered. Given the rising prevalence of Autism Spectrum Disorders (ASD) in the USA population and the documented effectiveness of ABA to help treat core deficits of this condition (Foxx, 2008), there is a need to increase the availability of ABA services now and even once the pandemic recedes. This need is underscored given that the supply of certified ABA service providers per capita is far less (Zhang & Cummings, 2020) than the benchmark needed to provide services to children with ASD in nearly every state in the USA (49 out of 50 states). This deficit is especially felt in many rural communities where children do not have access to in-person ABA services (Ferguson et al., 2019). To meet this need, remote delivery of ABA instructional services is one potential modality to explore in greater depth.

To address the gap in the literature relating to the evaluation of direct delivery of services through a remote modality, I conducted two experiments to explore the effectiveness of remote delivery of educational and ABA services. Given the documented concerns relating to delivering educational services remotely (e.g. see Partnership for Los Angeles Schools, 2020; Psacharopoulos et al., 2020; Engzell et al., 2020; Kuhfeld et al., 2020 among others) I wanted to know if remote instruction would produce inferior learning outcomes in an empirical investigation. The findings from Experiment I revealed that when the number of learning opportunities is held constant, the number of objectives mastered and rate of instruction (LU completed per minute), and efficiency of instruction (i.e., LU needed per STO) were equal or better during remote instruction in approximately half of the comparisons. The potential effectiveness of remote instruction was supported by the findings from Experiment II that found no significant difference in rate of learning (LU per STO and LU per LTO) across 13 weeks of remote instruction in a hybrid model compared to the same duration across 4 preceding years of in-person instruction.

However, along with these promising data supporting the efficiency of instruction, there are non-learning rate differences that are identified in the between-years component from Experiment II. Specifically, the number of instructional opportunities (i.e., LU), total TPRA observations, and number of weeks of instruction attended in the hybrid model were on average 37% lower when compared to the in-person model. Although this number represents a statistically significant (p = <.05) decrease, it is a better performance than the predictive model of Kuhfeld (2020) which projected 37-68% decreases across reading and math outcomes with remote instruction due to school closures relating to COVID-19. Based on the outcomes comparing in-person and remote instruction across both experiments it can be argued that, when instruction is successfully delivered to students in a remote model, the rate of learning could be comparable. Based on this analysis, one avenue of research to further improve remote instruction can be for identifying variables that could aid in increasing access for students to receive more instruction.

Prerequisites Needed for Remote Instruction

The data reported from both experiments comparing educational outcomes in remote and hybrid instruction were collected from across students with different VB developmental cusps. To provide remote instruction most effectively, it is necessary to identify which prerequisites allow for children to best benefit from remote instruction. Given the research supporting verbal behavior (VB) developmental cusps and capabilities as sources of stimulus control which allow for students to learn in new ways, it is worthwhile to identify which VB cusps allow for children to benefit from remote instruction.

As no prior research examined the effects of the presence of various VB developmental cusps and capabilities on outcomes during remote instruction, I wanted to fill this gap in the literature by conducting an analysis of student outcomes grouped by levels of VB development. To this end, I asked if educational outcomes for students would differ based on their level of VB development. Specifically, the research base has identified the joining of listener and speaker repertoires (i.e., Naming) as a milestone which allows for students to learn faster and in new ways they could not learn before (see Greer et al., 2017). As such, I wanted to know if students who demonstrated some degree of Naming would achieve different outcomes than students without this VB developmental capability.

To answer this question, I conducted a between-student analysis, in which I grouped students by their level of VB development. Based on the classifications described by Pohl et al., (2020) students were placed in the following five groups: pre-listeners, listeners/pre-speakers, listener/speaker with pre-UniN, UniN, and BiN. The first three groups did not demonstrate any degree of Naming while students from UniN and BiN groups demonstrated some degree of Naming. I found that once students acquired the listener half of Naming, learning outcomes were significantly different (p = <.05) on both rate of learning measures (LUSTO and LULTO) as well as nearly all learning outcomes (LU and STO) except for total LTO mastered. Notably, within the groups of students who did not acquire any degree of Naming, there were no significant differences on any measures except for LUSTO between the group of students operating on listener/speaker with pre-UniN level of VB and the students on pre-listener level of

VB. Thus, these findings extend the literature on Naming by demonstrating that students with Naming benefit significantly more from remote instruction than students without this capability.

Summary of Major Findings

The purpose of the experiments described in this dissertation was twofold. First, I wanted to provide a detailed analysis of how remote instruction compares to in-person instruction in terms of educational outcomes. Second, I wanted to investigate the effects of the presence of VB developmental cusps on educational achievement during remote instruction.

To address the first goal, I used a single-subject design with an emphasis on internal validity in Experiment I and a between-groups design with data collected from 108 students in Experiment II. The results from Experiment I indicated that remote instruction produced comparable outcomes in approximately half of the comparisons. The results from Experiment II indicated that in terms of rate of learning (LUSTO and LULTO) there was no significant ($p = <.05$) difference between in-person and remote instruction. However, in terms of overall learning measures (i.e., weeks of school attended, LU completed, STOs and LTOs mastered), students achieved significantly ($p = <.05$) less in the remote model compared to in-person with an average decrease of 37% across measures. Despite the significant decrease in overall student achievement in the hybrid model compared to the in-person model, the outcomes from the CABAS® extension to remote provision outperformed reported outcomes from a large-scale study in the Netherlands (Engzell et al., 2020) demonstrating nearly no learning occurring when in-person instruction was not available, as well as a predictive model from Kuhfeld et al., (2020) predicting decreases of 37-50% in reading and 63-68% decreases in reading due to in-person school closures during COVID-19.

To address the second goal relating to the identification of prerequisites to benefit from remote instruction, and specifically the predictive effects of the presence of various VB developmental cusps on student outcomes I conducted a between-student analysis. In the between-student analysis I grouped students by their level of VB development and analyzed their educational outcomes in terms of weeks of instruction and total LU received, total objectives mastered (STO and LTO), and efficiency of instruction measured as number of LU required to master an objective (STO and LTO). The findings revealed that significant (p = <.05) differences between groups consistently started to appear once students acquired some degree of Naming. Specifically, students with some degree of Naming achieved significantly better outcomes than students who were pre-UniN except for the total LTOs mastered. In contrast, the only measure in which there was a significant difference in outcomes between groups of students who were pre-UniN was in the LUSTO measure in which students of listener/speaker pre-UniN mastered objectives after significantly fewer LU than students in the pre-listener group.

Implications

Based on the results of the experiments reported herein, several claims can be made about the results and how they address gaps in research. First, the findings of Experiment I indicate that, in about half of comparisons, remote instruction results in equal or better outcomes than in-person instruction. These findings fill a gap in the literature as no other studies reported conduct a carefully controlled investigation of outcomes for in-person versus remote instruction for learning novel targets. The investigation differs from other published ABA literature on telehealth in that it emphasizes learning of novel outcomes across in-person and remote provision in a carefully controlled study. This is compared to reports of average levels of correct responding by Powell et al., (2021) or a systematic review of caregiver training via telehealth by

Unholz-Bowden (2020). A more detailed review of the literature is provided earlier in Chapters 1-4.

Second, the examination of educational outcomes tied to levels of VB development fills a gap in the literature relating to identifying prerequisites for children to benefit from remote instruction. The findings from the between-students analysis reveal that students with some degree of Naming achieved significantly ($p = <.05$) more learning outcomes while requiring fewer LU in every measure aside from total LTO. This contrasts with the three groups of students with pre-UniN (i.e., pre-listeners, listeners/pre-speakers, and listener/speakers pre-UniN) who did not demonstrate any significant differences in any measures aside for LUSTO in which listener/speakers with pre-UniN required significantly fewer LU per STO than students on pre-listener level of VB development.

Thus, the implication of these findings is that the presence of Naming can serve as a prerequisite which allows for a student to best benefit from receiving remote instruction. Thus, it is important to emphasize the establishment of Naming and related VB developmental cusps to help children best benefit from remote instruction. These outcomes extend and support the body of literature emphasizing the identification and establishment of VB developmental cusps as critical components of educational systems for children who do not develop this capability in the natural environment without special interventions.

The outcomes reported herein reflect outcomes which outperformed reports from actual outcomes of school closures nationwide in Netherlands (Engzell et al., 2020) as well as a predictive model by Kuhfeld and colleagues (2020). Some unique components of CABAS® that have been shown to be related to improved outcomes include the built-in teacher training component which a recent dissertation (Silsilah, 2019) showed that the number of teaching training components completed was significantly related to improved student outcomes. As the teaching training includes components across verbal behavior of the science, contingency shaped teaching behaviors, and verbally mediated repertoires, it is possible that the comprehensive training teacher receive in the CABAS® allows for greater success in transferring to remote provision of instruction. Another component that is important to highlight is the learn unit (Albers & Greer, 1991; Greer & McDonough, 1999) which is used as a fundamental component of instructional. Given the salient research the learn unit is built on, the delivery of learn unit across a remote modality could account for these outcomes.

Finally, data reported in the between-years analysis comparing outcomes from the remote provision of the CABAS® model to data from the preceding four years of in-person instruction provides an objective measure of student outcomes during remote instruction in the COVID-19 pandemic. As schools were closing due to COVID-19, there was much concern as to how student outcomes will fare given the unprecedented scale of school closures (e.g. see Kuhfeld et al., 2020; Engzell et al., 2020; Kaffenberger, 2021 among others). Although this analysis did not account for domain-specific outcomes, the overall results indicate that that the outcomes of the hybrid model were approximately 37% lower than a fully in-person model. Although a statistically significant decrease, these outcomes outperformed the predictive model reported by Kuhfeld et al., (2020) (37-67% decreases across reading and math) as well as the report by Engzell et al., (2020) from educational achievement in the Netherlands during school closures (almost no learning occurring during school closures). Specifically, when instruction is delivered with fidelity using a robust model built on principles of behavior analysis (as in the CABAS® model), outcomes achieved during remote instruction are approximately 73% as effective as in-person instruction. The implication of these findings, on one hand support the effectiveness of the CABAS® system of instruction in a novel modality, and on the other hand support the potential effectiveness of remote instructional delivery when in-person instruction is not available.

Limitations

As described in both experiments, there were several limitations to both studies. In Experiment I all participants operated on a minimum of a listener/speaker (pre-UniN) level of VB development. Thus, outcomes from the carefully controlled comparison between in-person and remote instruction are only discussed in terms of outcomes for students of those levels of VB. No implication from this study is available for students of pre-speaker or pre-listener levels of VB development. Further, the objectives taught in the carefully controlled comparison of Experiment I included whole-word sight words and tacts – objectives which required relatively lower response efforts for the students who had previously mastered responding to a large number of tacts or sight words. Thus, the outcomes of the analysis may or may not generalize to other students or even the same students when learning different types of responses.

Further studies extending the findings from Experiment I should conduct a carefully controlled experiment comparing in-person and remote instruction across students from pre-speaker and pre-listener levels of VB do determine if their educational outcomes will differ from what is reported in this manuscript. Further, the variety of instructional targets being compared should be extended to include a wider variety of targets including discrimination responses, conditioning reinforcement for novel toys or activities, as well as self-management objectives relating to repertoires needed to participate in remote instruction.

Limitations in Experiment II were related to the duration for which data were collected and reported. Data were only collected and reported for the first 14 weeks of the school year. The data reflect that, overall, as the weeks progressed, the trend of the data paths demonstrated that students achieved better outcomes across every single measure. More LUs were completed, and more objectives were mastered, while requiring progressively fewer LU per STO and LU per LTO as the year went on. If the trend apparent from weeks 2-14 continued, it is possible that as the school year continued with both teachers and students gaining more experience navigating the hybrid model, educational outcomes could have improved even further than what is reported above.

Further limitations in Experiment II relate to the size of the sample in each group of students included in the analysis. As some groups had approximately twice the number of students as other groups (e.g. 13 students with BiN compared to 25 listener/speakers pre-UniN or 28 pre-listeners), it is possible that the data would have been different when analyzed from outcomes of groups with larger samples.

A final limitation in Experiment II relates to the proportion of remote instruction to in-person instruction delivered during the hybrid model. The hybrid model reported herein was a model with 60% remote instruction and 40% in-person instruction overall. Thus, results from a hybrid model cannot necessarily be extended to a fully remote model. It is possible that effects of remote instruction are influenced by the 40% of in-person instruction that each student received. To better assess educational achievement from remote instruction, analysis of data collected in a fully remote provision would be needed. Alternatively, a parametric analysis of varying proportions of in-person to remote instruction can identify which proportions produce the best outcomes. Examples could include 50% remote to 50% in-person, or perhaps 80% remote to 20% in person. Data collected from such a model could identify if including a percentage of in-person instruction can help amplify the effectiveness of remote instruction as well as reporting educational outcomes from a fully remote provision of services.

Further studies can extend the findings of Experiment II by conducting a similar analysis across data collected from an entire school year in the hybrid model compared to preceding years of fully in-person instruction. This should be done while controlling for sample size across different groups.

Educational Significance

The school closures of COVID-19 were unprecedented in scale, duration, and intensity (Azevedo et al., 2020). Thus, a detailed analysis of how students learn remotely compared to in-person is important in that it provides a close analysis of the feasibility of remote instruction. The outcomes from Experiment I and the between-years analysis in Experiment II support the recommendation of remote instruction for students who do not have access to high-quality in-person services. The findings of the between-student analysis in Experiment II, where students were grouped by level of VB development, indicate that when some degree of Naming is present, students are likely to benefit most from remote instruction. These findings are important for school closures related due to a global pandemic as well as for students in geographically distant locations who do not have access to high-quality in-person educational or ABA services which serve their needs.

The findings from the between-groups comparison builds on the existing literature supporting the educational significance of identifying of the presence of absence of the Naming capability as well as interventions to establish this capability when it is missing. These findings extend the literature relating to how children with Naming can learn in new ways. In addition to increased rate of learning, the ability to learn incidentally (Greer & Longano, 2010) as well as the ability to learn faster from instructional demonstrations when compared to direct consequences (Hranchuk et al., 2019), these findings support that the presence of the Naming capability can help a child learn remotely in a way which they would not have without this capability.

Remote delivery of instructional and ABA services is, in some ways, a new frontier for using the science of behavior to produce socially significant beneficial outcomes for students with and without disabilities. As technology advances and new modalities for service delivery become available, it is important to identify ways in which new technologies can be used to help benefit those in need. There is much research demonstrating the power of using the science of behavior to help students learn. It is our responsibility to become familiar with the tools that exist and use them to help as many children as possible develop socially significant skills.

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.

American Telemedicine Association (2020, April). *Telehealth basics*. https://www.americantelemed.org/resource/why-telemedicine/

Azevedo, J. P., Hasan, A., Goldemberg, D., Iqbal, S. A., & Geven, K. (2020). *Simulating the potential impacts of COVID-19 school closures on schooling and learning outcomes: A set of global estimates*. The World Bank**.**

Azoulay, U. (2020, November 11). *Global Education Coalition* . UNESCO. https://en.unesco.org/covid19/educationresponse/globalcoalition

Barnett, M. L., Ray, K. N., Souza, J., & Mehrotra, A. (2018). Trends in telemedicine use in a large commercially insured population, 2005-2017. *Jama*, *320*(20), 2147-2149.

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.

Basilaia, G., & Kvavadze, D. (2020). Transition to online education in schools during a SARS-CoV-2 coronavirus (COVID-19) pandemic in Georgia. *Pedagogical Research*, *5*(4).

Benson, S. S., Dimian, A. F., Elmquist, M., Simacek, J., McComas, J. J., & Symons, F. J. (2018). Coaching parents to assess and treat self-injurious behaviour via telehealth. *Journal of Intellectual Disability Research*, *62*(12), 1114-1123.

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.

Brown, S. R., & Melamed, L. E. (1990). *Experimental design and analysis* (No. 74). Sage.
Center for Connected Health Policy (2020, March). *What is telehealth?* https://www.cchpca.org/about/about-telehealth

Cooper, J. O., Heron, T. E., & Heward, W. L. (2007). Applied behavior analysis. Pearson.

Du, L., Broto, J., & Greer, R. D. (2015). The effects of establishment of conditioned reinforcement for observing responses for 3D stimuli on generalized visual match-to-sample in children with autism spectrum disorders. *European Journal of Behavior Analysis*, *16*(1), 82-98.

Du, L., & Greer, R. D. (2014). Validation of adult generalized imitation topographies and the emergence of generalized imitation in young children with autism as a function of mirror training. *The Psychological Record*, *64*(2), 161-177.

Du, L., Speckman, J., Medina, M., & Cole-Hatchard, M. (2017). The effects of an auditory matching iPad app on three preschoolers' echoic and listener responses. *Behavior Analysis in Practice*, *10*(2), 118-130.

Engzell, P., Frey, A., & Verhagen, M. D. (2020). *Learning inequality during the COVID-19 pandemic*. (https://osf.io/download/5f995b4687b7df03233b06fe/). OSF preprint.

Ferguson, J., Craig, E. A., & Dounavi, K. (2019). Telehealth as a model for providing behavior analytic interventions to individuals with autism spectrum disorder: A systematic review. *Journal of Autism and Developmental Disorders*, **49**(2), 582–616.

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*.

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.

Foxx, R. M. (2008). Applied behavior analysis treatment of autism: The state of the art. *Child and adolescent psychiatric clinics of North America*, *17*(4), 821-834.

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*, 1.

Greer, R. D. (1991). The teacher as strategic scientist: A solution to our educational crisis?. *Behavior and Social Issues*, *1*(2), 25-41.

Greer, R. D. (1994). A systems analysis of the behaviors of schooling. *Journal of Behavioral Education*, *4*(3), 255-264.

Greer, R. D. (2002). *Designing teaching strategies: An applied behavior analysis systems approach*. Elsevier.

Greer, R. D., et al. (2020). *Early Learner Curriculum and Achievement Record (ELCAR): A CABAS® Developmental Inventory*.

Greer, R. D.., Chavez-Brown, M., Nirgudkar, A. S., Stolfi, L., & Rivera-Valdes, C. (2005). Acquisition of fluent listener responses and the educational advancement of young children with autism and severe language delays. *European Journal of Behavior Analysis*, *6*(2), 1-20.

Greer, R. D., Corwin, A., & Buttigieg, S. (2011). The effects of the verbal developmental capability of naming on how children can be taught. *Acta de Investigación P sicológica*, *1*(1), 23-54.

Greer, R. D., & Longano, J. (2010). A rose by naming: How we may learn how to do it. *The Analysis of Verbal Behavior*, *26*(1), 73-106.

Greer, R. D., McCorkle, N., & Williams, G. (1989). A sustained analysis of the behaviors of schooling. *Behavioral Interventions*, *4*(2), 113-141.

Greer, R. D., & McDonough, S. H. (1999). Is the learn unit a fundamental measure of pedagogy?. *The Behavior Analyst*, *22*(1), 5-16.

Greer, R. D., Pistoljevic, N., Cahill, C., & Du, L. (2011). Effects of conditioning voices as reinforcers for listener responses on rate of learning, awareness, and preferences for listening to stories in preschoolers with autism. *The Analysis of Verbal Behavior*, *27*(1), 103-124.

Greer, R. D., Pohl, P., Du, L., & Moschella, J. L. (2017). The separate development of children's listener and speaker behavior and the intercept as behavioral metamorphosis. *Journal of Behavioral and Brain Science*, *7*(13), 674-704.

Greer, R. D. & Ross, D. E. (2008). *Verbal behavior analysis: Inducing and expanding new verbal capabilities in children with language delays.* Boston: Pearson Education, Inc.

Greer, R. D., & Speckman, J. (2009). The integration of speaker and listener responses: A theory of verbal development. *The Psychological Record*, *59*(3), 449-488.

Greer, R. D., Stolfi, L., Chavez-Brown, M., & Rivera-Valdes, C. (2005). The emergence of the listener to speaker component of naming in children as a function of multiple exemplar instruction. *The Analysis of Verbal Behavior*, *21*(1), 123-134.

Hay-Hansson, A. W., & Eldevik, S. (2013). Training discrete trials teaching skills using videoconference. *Research in Autism Spectrum Disorders,7,* 1300–1309.

Heppen, J. B., Sorensen, N., Allensworth, E., Walters, K., Rickles, J., Taylor, S. S., & Michelman, V. (2017). The struggle to pass algebra: Online vs. face-to-face credit recovery for at-risk urban students. *Journal of Research on Educational Effectiveness, 10*(2), 272-296.

Higgins, W. J., Luczynski, K. C., Carroll, R. A., Fisher, W. W., & Mudford, 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.

Hranchuk, K., Greer, R. D., & Longano, J. (2019). Instructional demonstrations are more efficient than consequences alone for children with naming. *The Analysis of Verbal Behavior*, *35*(1), 1-20.

Ingersoll, B., Wainer, A. L., Berger, N. I., Pickard, K. E., & Bonter, N. (2016). Comparison of a self-directed and therapist-assisted telehealth parent-mediated intervention for children with ASD: A pilot RCT. *Journal of Autism and Developmental Disorders*, *46*(7), 2275-2284.

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.

Kaffenberger, M. (2021). Modelling the long-run learning impact of the Covid-19 learning shock: Actions to (more than) mitigate loss. *International Journal of Educational Development*, *81*, 102326.

Keller, F. S. (1968). Good-bye, teacher... *Journal of Applied Behavior Analysis*, *1*(1), 79.

Keohane, D. D., & Greer, R. D. (2005). Teachers' use of a verbally governed algorithm and student learning. *International Journal of Behavioral Consultation and Therapy*, *1*(3), 252.

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.

Kuhfeld, M., Soland, J., Tarasawa, B., Johnson, A., Ruzek, E., & Liu, J. (2020). Projecting the potential impact of COVID-19 school closures on academic achievement. *Educational Researcher*, *49*(8), 549-565.

Lamm, N., & Greer, R. D. (1991). A systematic replication and a comparative analysis of CABAS. *Journal of Behavioral Education*, *1*(4), 427-444.

Maffei, J., Singer-Dudek, J., &, Keohane, D., D. (2014). The effects of the establishment of adult faces and/or voices as conditioned reinforcers for children with ASD and related disorders. *Acta de Investigación Psicológica*, *4*(3), 1621-1641.

McGarrell, M., Healy, O., Leader, G., O'Connor, J., & Kenny, N. (2009). Six reports of children with autism spectrum disorder following intensive behavioral intervention using the Preschool Inventory of Repertoires for Kindergarten (PIRK®). *Research in Autism Spectrum Disorders*, *3*(3), 767-782.

Meadan, H., Snodgrass, M. R., Meyer, L. E., Fisher, K. W., Chung, M. Y., & Halle, J. W. (2016). Internet-based parent-implemented intervention for young children with autism: A pilot study. *Journal of Early Intervention*, *38*(1), 3-23.

Morgan, G. A., Greer, R. D., & Fienup, D. M. (2020). Descriptive Analyses of Relations among Bidirectional Naming, Arbitrary, and Nonarbitrary Relations. *The Psychological Record*, 1-21.

Nuzzolo-Gomez, R., & Greer, R. D. (2004). Emergence of untaught mands or tacts of novel adjective-object pairs as a function of instructional history. *The Analysis of Verbal Behavior*, *20*(1), 63-76.

Park, H. L., Du, L., & Choi, J. (2020). Education for Children with ASD in South Korea: A Case Study. *Behavior Analysis in Practice*, 1-11.

Partnership for Los Angeles Schools. (2020). *Listening to learn: What Los Angeles families say the need during distance learning [online]. Retrieved from https://partnershipla.org/wp-content/uploads/2020/09/Equity-Alert-USC-and-P4LA.pdf*

Pistoljevic, N., & Greer, R. D. (2006). The effects of daily intensive tact instruction on preschool students' emission of pure tacts and mands in non-instructional setting. *Journal of Early and Intensive Behavior Intervention*, *3*(1), 103

Pohl, P., Greer, R. D., Du, L., & Moschella, J. L. (2020). Verbal development, behavioral metamorphosis, and the evolution of language. *Perspectives on Behavior Science*, *43*(1), 215-232.

Pollard, J. S., LeBlanc, L. A., Griffin, C. A., & Baker, J. M. (2020). The effects of transition to technician-delivered telehealth ABA treatment during the COVID-19 crisis: A preliminary analysis. *Journal of Applied Behavior Analysis*. *53* (3), 1876-1888.

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.

Psacharopoulos, G., Collis, V., Patrinos, H. A., & Vegas, E. (2020). Lost Wages: The COVID-19 cost of school closures. *World Bank Policy Research Working Paper*, (9246).

Reed, P. Osborne, L. A., & Corness, M. (2006). Brief report: The effectiveness of different home-based behavioral approaches to early teaching intervention. *Journal of Autism and Developmental Disorders.*

Rodriguez, K. A. (2020). Maintaining treatment integrity in the face of crisis: A treatment selection model for transitioning direct ABA services to telehealth. *Behavior Analysis in Practice*, 1.

Rosales-Ruiz, J., & Baer, D. M. (1997). Behavioral cusps: A developmental and pragmatic concept for behavior analysis. *Journal of Applied Behavior Analysis*, *30*(3), 533-544.

Ross, D. E., Nuzzolo, R., Stolfi, L., & Natarelli, S. (2006). Effects of speaker immersion on independent speaker behavior of preschool children with verbal delays. *Journal of Early and Intensive Behavior Intervention*, *3*(1), 135.

Ross, D. E., Singer-Dudek, J., & Greer, R. D. (2005). The teacher performance rate and accuracy scale (TPRA): Training as evaluation. *Education and Training in Developmental Disabilities*, 411-423.

Selinske, J. E., Greer, R. D., & Lodhi, S. (1991). A functional analysis of the comprehensive application of behavior analysis to schooling. *Journal of Applied Behavior Analysis*, *24*(1), 107-117.

Silsilah, S. (2019). *What Makes Teachers Effective: Investigating the Relationship Between CABAS® Teacher Ranks and Teacher Effectiveness* (Doctoral dissertation, Columbia University).

Simacek, J., Dimian, A. F., & McComas, J. J. (2017). Communication intervention for young children with severe neurodevelopmental disabilities via telehealth. *Journal of Autism and Developmental Disorders,47,* 744–767.

Singer-Dudek, J., Speckman, J., & Nuzzolo, R. (2010). A comparative analysis of the CABAS® model of education at the Fred S. Keller School: A twenty-year review. *The Behavior Analyst Today*, *11*(4), 253.

South Carolina Medicaid. (2020, March 30). *COVID-19 Telehealth Policy Update to ABA Coverage.* https://www.scdhhs.gov/press-release/coronavirus-disease-2019-covid-19-temporary-telephonic-and-telehealth-services-updat-4.

Tomlinson, S. R. L., 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.

Tsiouri, I., & Greer, R. D. (2007). The role of different social reinforcement contingencies in inducing echoic tacts through motor imitation responding in children with severe languagedelays. *Journal of Early and Intensive Behavior Intervention*, *4*(4), 629.

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.

Wacker, D. P., Schieltz, K. M., Suess, A. N., Romani, P. W., Dalmau, Y. C. P., Kopelman, T. G., ... & Lindgren, S. D. (2016). Telehealth. In *Handbook of Evidence-based Practices in Intellectual and Developmental Disabilities* (pp. 585-613). Springer, Cham.

Waddington, E. M, & Reed, P. (2009). The impact of using the "Preschool Inventory of

Repertoires for Kindergarten (PIRK® on school outcomes on children with Autism Spectrum Disorders. *Research in Autism Disorders, 3,* 809-827.

Wolery, M., Gast, D. L., & Ledford, J. R. (2014). Comparison designs. In *Single case research methodology* (pp. 297-345). Routledge.

Wong, K. K., Bajwa, T., & Fienup, D. M. The Application of Mastery Criterion to Individual Operants and the Effects on Acquisition and Maintenance of Responses. *Journal of Behavioral Education*, 1-23.

Zhang, Y. X., & Cummings, J. R. (2020). Supply of certified applied behavior analysts in the United States: Implications for service delivery for children with autism. *Psychiatric Services*, *71*(4), 385-388.