ALTERNATIVES TO TRADITIONAL PROFESSIONAL DEVELOPMENT: A CLOSER LOOK AT IMPLEMENTATION

A DISSERTATION

SUBMITTED TO THE FACULTY

OF

THE GRADUATE SCHOOL OF APPLIED AND PROFESSIONAL PSYCHOLOGY

OF

RUTGERS,

THE STATE UNIVERSITY OF NEW JERSEY

BY

MADELINE WHITE

IN PARTIAL FULFILLMENT OF THE

REQUIREMENTS FOR THE DEGREE

OF

DOCTOR OF PSYCHOLOGY

NEW BRUNSWICK, NEW JERSEY MAY 2023

APPROVED: _____
Elisa Shernoff, Ph.D.

_____
Linda Reddy, Ph.D.

DEAN: _____
Arpana G. Inman, Ph.D.

ABSTRACT

Educators feel overwhelmed by managing challenging behaviors which often leads to burnout and turnover (Clotfelter et al., 2006; Kokkinos, 2007; Simon & Johnson, 2015). However, professional development (PD) in evidenced-based behavior management practices (EBMPs) is sparse with current PD models yielding varied outcomes (Darling-Hammond et al., 2009; Wei et al., 2010). Simulation training has been developed as teacher trainings to address limitations of traditional PD but, no studies to date have examined the implementation practices of simulation training (Martin et al., 2010; Proctor et al., 2011). This study examined the implementation practices of teachers' using IVT-T through descriptive patterns of use (fidelity) as well as associations between dosage, usability, and teachers' use of behavioral praise and behavioral corrective feedback. Subjects included 48 teachers from three urban K-8 schools. An exploratory analysis of fidelity revealed that on average teachers played IVT-T for less than 25 minutes per week and 92% of teachers played between 1-10 weeks. Adjusting for non-players, the weekly average minutes met the 40-minute threshold 70% of the weeks. Of the teachers that did not meet the threshold hold, 48% played between 20-39 minutes. Results of the negative binomial regression analyses revealed that dosage was not a significant predictor of behavioral praise or behavioral corrective feedback, when controlling for teachers' use of the behavioral strategy at Time 1. Finally, a correlational analysis revealed that the usability of IVT-T was moderately, positively correlated with dosage. Limitations and implications for future research and practice are discussed.

ACKNOWLEDGMENTS

As I reach the end of my graduate school journey, I would like to take a moment to express my deep appreciation, gratitude, and love for those who have supported me. I would like to thank my parents, Karen and Thomas White for instilling in me the perseverance and passion needed to complete my doctoral education. I would not be where I am today if it were not for your support and encouragement, thank you. I would also like to thank my sister, Abigail Menchel, for being my lifeline, source of humor, and the greatest role model to keep me going when things get tough.

To my fiancé, Chris Lehrer, I am so thankful and lucky to have had you by my side throughout graduate school. Your willingness to move to California for internship, read countless drafts, and help me prepare for interviews, presentations, etc., has meant so much to me. You have made even the hardest of hurdles feel achievable because, as you always say, we are able to conquer anything as a team. I deeply cherish you and your ability to bring humor and light into our lives every day. I am forever grateful to have you by my side in my life, now and forever.

Finally, I would like to thank those in the GSAPP Community who have provided me with mentorship, guidance, and friendship. In particular, Dr. Jeff Segal and Dr. Adam Lewka, thank you for spending so much of your time helping me find my way at GSAPP and making it that much easier to get where I am going. And, to Sydney Brinson, I am so grateful to have found you as my first-year roommate and to be able to share this journey together. I will never forget all of the time we spent getting coffee, staying late in the clinic, and living in Highland Park.

TABLE OF CONTENTS

ABSTRACT....ii
ACKNOWLEDGEMENTS....iii
LIST OF TABLES....iv
LIST OF FIGURES....v
INTRODUCTION....1
Teacher Stress Related to Behavior Management....2
Teacher Training in EBMPs....4
Effectiveness of Teacher Behavior Management Training/PD....5
The Value of Simulation Training....7
Integration of Simulation within Teacher PD....10
Implementation Science....12
Dosage....13
Fidelity....14
Usability....15
Design and Research Questions....17
METHOD....18
Schools....18
Teachers....19
IVT-T Simulation Training....19
Procedures for Larger Study....21
Measures....22

Analyses 25

Missing Values Analyses 27

RESULTS 28

How Frequently Did Teachers Use IVT-T? 28

Does IVT-T Dosage Predict EBMP Use? 30

Does Usability Impact IVT-T Dosage? 31

DISCUSSION 31

How Frequently Did Teachers Use IVT-T? 31

Does IVT-T Dosage Predict EBMP Use? 33

Does Usability Impact IVT-T Dosage? 36

Limitations 39

Implications for Research and Practice 41

Conclusions 43

REFERENCES 45

APPENDICES 60

LIST OF TABLES

Table 1 Teacher Characteristics.....60

Table 2 Descriptive Statistics for Independent and Depend Variables for Full Dataset 61

Table 3 Descriptive Statistics for Research Question 3.....61

Table 4 Intercorrelations for Study Variables.....62

Table 5 Number of Teachers Playing IVT-T Each Week.....63

Table 6 Regression Analysis for Dosage as a Predictor of CSAS Strategy Counts 64

Table 7 Correlation Analysis of Usability and IVT-T Mins.....65

LIST OF FIGURES

Figure 1 Average Time Teachers Played IVT-T per Week 66

Figure 2 Relationship Between Total Weeks Played and Change in CSAS Strategy Counts 67

Introduction

Studies document that teachers are overwhelmed by managing challenging behaviors, including student inattention, hyperactivity, and difficulty following teacher directions (Browers & Tomic, 2000; Clotfelter et al., 2006; Kokkinos, 2007; Simon & Johnson, 2015). Managing these challenging behaviors is time consuming for teachers and leads to diminished teaching efficacy and teacher burnout (Browers & Tomic, 2000; Clotfelter et al., 2006; Kokkinos, 2007; Simon & Johnson, 2015). While the use of evidence-based practices reduces challenging behaviors in the classroom and is indirectly associated with reductions in teacher stress, few educators report receiving sufficient training in these strategies (Mitchell et al., 2017; Poznanski et al., 2018; Reinke et al., 2011; Tschannen-Moran & Hoy, 2007).

Evidence-based behavior management practices (EBMPs) are classroom practices supported by research that promote positive behavioral change (e.g., praise, direct one-step commands, and selective attention; APA Task Force, 2006; Kern & Clemens, 2007; Parsonson, 2012). Current research suggests that some teacher professional development (PD) is either ineffective, does not train teachers in EBMPs, or is improperly implemented thereby reducing overall effectiveness of the training (Darling-Hammond et al., 2010; Durlak & DuPre, 2008; Franzen & Kamps, 2008; Martin et al., 2010; Mihalic, 2004; Noell et al., 2014; Proctor et al., 2009). PD includes efficacious treatment programs, workshops, conferences, observations, university courses and other training sessions (Darling-Hammond et al., 2009; Noell & Gansle, 2009; Wei et al., 2010). Despite the link between implementation and expected training outcomes in real-world settings, implementation practices are infrequently studied (Fixsen et al., 2005; Mihalic, 2004; Proctor et al., 2011).

Simulation training has recently been developed as an alternative to traditional PD in the field of education because it has been shown to increase skill acquisition and generalizability in other professional fields (Caro, 1973; Cook et al., 2011; Jentsch & Curtis, 2017). Scholars describe traditional PD as encompassing didactic instruction, facilitation of professional communities, and resource sharing (Elliot, 2017; Little & Housand, 2011). To date, only a handful of simulation trainings exist in the field of education and even fewer train teachers in behavior management strategies (Beisiegel et al., 2018; Dieker et al., 2014; Sitzmann, 2011; Zibit & Gibson, 2005). None of the simulation training studies examine the effect of implementation factors on training outcomes (Beisiegel et al., 2018; Dieker et al., 2014; Sitzmann, 2011; Zibit & Gibson, 2005). Interactive Virtual Training for Teachers (IVT-T) is a simulation training program that aims to support teacher skill development in EBMPs and demonstrate the effectiveness of simulation training as a PD model in the field of education (Shernoff et al., 2018). IVT-T is the focus of this study because it utilizes teaching characteristics associated with high-quality PD delivered in a simulated environment (Shernoff et al., 2018). This study is concerned with assessing the implementation of IVT-T to determine its viability in delivering EBMPs via simulation training.

Teacher Stress Related to Behavior Management

Teachers commonly cite managing challenging behaviors as time consuming, which contributes to high rates of turnover (Clotfelter et al., 2006; Kokkinos, 2007; Simon & Johnson, 2015). Clotfelter et al. (2006), found that stress related to time spent preparing students for instruction, addressing challenging behaviors, and limited school-wide support are the leading stressors associated with high turnover in schools (Simon & Johnson, 2015). Because teachers receive little training in behavior management skills, they spend much of their time managing

challenging behaviors which frequently results in less time for instruction. In turn, teachers' self-efficacy can be diminished due to challenges with managing behaviors.

Self-efficacy is defined as beliefs about the ability and capacity to accomplish tasks required (Bandura, 1977). Collie et al. (2012) studied a sample of 664 teachers across 17 schools to understand how school climate and social-emotional learning were related to teacher stress, teaching efficacy, and job satisfaction. Participants completed surveys for predictor variables (i.e., school climate and social-emotional learning) and outcomes variables (i.e., stress, teaching efficacy, and job satisfaction), which were then examined via factor analyses and structural equation models. Results indicated teaching efficacy was negatively associated with perceived stress related to challenging behaviors ($\beta = -.32$, $p < .001$); in other words, stress that educators experience when managing challenging behaviors is associated with feeling less confident in their teaching abilities. Teaching efficacy was positively associated with job satisfaction ($\beta = .33$, $p < .001$). Collie et al. also found that if teachers feel competent managing challenging behaviors, those behaviors minimally impacted their job satisfaction. These findings suggest higher teacher efficacy reduces stress when managing challenging behaviors. In addition, a recommended outcome from the study is for schools to provide sustained and appropriate PD in EBMPs to reinforce teacher confidence.

Teachers who are not trained in EBMPs may spend more time managing challenging behaviors, and thus may experience greater stress and burnout (Pas et al., 2010; Simon & Johnson, 2015). When teachers experience greater levels of stress, they are more likely to use ineffective and punitive strategies to address those challenging behaviors (Pas et al., 2010; Simon & Johnson, 2015). The cycle of stress is perpetuated by teachers using ineffective behavior management strategies that increase the occurrence of challenging behaviors. This

cycle results in increasing teachers' ineffectiveness and stress (Browers & Tomic, 2000; Pas et al., 2010). This cycle is predictive of teachers quitting and even changing careers (Browers & Tomic, 2000; Pas et al., 2010; Poznanski et al., 2018, Reinke et al., 2011; Simon and Johnson, 2015).

Reducing teacher stress through training in EBMPs can promote student engagement and is linked to higher academic achievement, higher intrinsic and extrinsic teacher satisfaction, and lower teacher turnover (Shinn et al., 1987). Intrinsic job satisfaction accounts for how well an individual teacher views their performance and internal growth, whereas extrinsic job satisfaction examines the teacher's comfort within the workplace environment (Bogler & Nir, 2012). High job satisfaction among teachers correlates with lower turnover rates, lower teacher absenteeism, and higher student achievement as well (Banerjee et al., 2017; Lawler, 1973; Locke, 1969). Given teachers receive limited training in EBMPs, further research is needed to examine how EBMPs can be successfully disseminated to teachers (Wei et al., 2010).

Teacher Training in EBMPs

Teachers request for training in behavior management is related to encountering challenging behaviors frequently and feeling unprepared to manage them (Darling-Hammond et al., 2009; Greenberg et al., 2014; Kokkinos, 2007; Melnick & Meister, 2008; Wei et al., 2010). Challenging behaviors are often observed among students with attention-deficit/hyperactivity disorder (ADHD). Congruent with teacher report, ADHD is the second leading psychiatric disorder among adolescents in the U.S. after anxiety disorders (American Psychiatric Association, 2013 [APA]; Kessler et al., 2012). Therefore, behavior management is often discussed in the context of ADHD (Reinke et al., 2011). Poznanski et al. (2018), found that preservice teachers on average had less than 50% accurate knowledge of ADHD characteristics

and a high number of misperceptions about appropriate treatments for ADHD. Conversely, student teachers were better able to identify the correct EBMPs to address anxiety in the classroom compared with ADHD (Pozanski et al., 2018). Teachers are more likely to begin their careers not knowing the appropriate EBMPs to implement when faced with challenging behaviors.

Greenberg et al. (2014) conducted a study examining the types of behavior management strategies taught in 122 teacher preparation programs. Results revealed that most programs do not teach EBMPs and less than two-thirds of programs require trainees to practice EBMPs (Greenberg et al., 2014). Reinke et al. (2011), found that only one third of educators were taught EBMPs in their graduate or undergraduate course work. Most participants cited PD as the setting they were taught EBMPs. In addition, approximately one-fourth of the teachers reported having minimal or no experience practicing EBMPs (Reinke et al., 2011).

After teachers are employed, many still do not have access to training in EBMPs. School districts infrequently select PD that trains teachers in behavior management strategies associated with meaningful behavioral change (Darling-Hammond et al., 2010). Trends in PD indicate that more than one-half of all teachers receive zero hours of training in EBMPs for the entire year despite many teachers requesting such training (Wei et al., 2010). Ameliorating this gap requires schools to provide ongoing training in EBMPs.

Effectiveness of Teacher Behavior Management Training/PD

Schools across the U.S. spend large sums of money on traditional PD and continue to rely on traditional PD to train teachers on ever evolving and increasingly complex skills (Chung & Kim, 2010; Guskey, 2000; National Staff Development Council, 2011). Darling-Hammond et al. (2009) reviewed nationally representative data of PD offered in U.S. schools and found it is most

effective when intensive and long-term (30-100 hours), with brief (5-14 hours), isolated trainings yielding nonsignificant outcomes. Yet, data from over 130,000 teachers across all states, indicated most teachers received less than 16 hours of PD relevant to the content they taught and around 1-2 days of PD in various topics per year (Darling-Hammond et al., 2009). Furthermore, less than one half of the PDs implemented in U.S. schools in 2010 trained teachers on behavior management strategies (Wei et al., 2010).

PD is also considered high-quality when it is continuous, collaborative, data-driven, interest-driven, and interactive (Elliot, 2017; Lutrick & Szabo, 2012). High-quality PD is associated with higher teaching quality and curriculum implementation and linked with higher student academic outcomes (Darling-Hammond et al., 2010; Desimone, 2009; Lawless & Pellegrino, 2007). High-quality PD is also correlated with greater job satisfaction and school improvement within the U.S. (Bogler & Nir, 2012; Desimone, 2009). However, one important finding from Martin et al. (2010), revealed that even high-quality PD may not be linked with optimal results if not implemented successfully. Martin et al. examined implementation variations of an instructional-technology PD used by 272 teachers and how it impacted student outcomes over two years. Higher rates of implementing PD as intended were associated with higher student achievement ($r = .309$, $p < .001$), with all individual factors of implementation positively correlated with higher student test scores. The way in which PD is implemented determines how successful it can be at attaining the intended outcomes.

Research specific to EBMPs reveals lower quality PD is not associated with significant changes in teacher knowledge and management of challenging behaviors (Joyce & Showers, 2002; Latouche & Gascoigne, 2019; Mitchell et al., 2017). However, multi-component training is linked with increased generalizability of behavior management skills (Joyce & Showers, 2002;

Latouche & Gascoigne, 2019; Reddy et al., 2020). Multi-component training refers to receiving coaching and immediate feedback on performance in combination with traditionally delivered instruction (Joyce & Showers, 2002; Latouche & Gascoigne, 2019). Coaching (i.e., is associated with significantly greater use of EBMPs by teachers (Fabiano et al., 2018). This finding suggests that high-quality PD (e.g., coaching and immediate feedback) can support teachers' behavior management skills. In addition, a large gap in the literature exists in the assessment of EBMPs implementation despite being connected with positive student outcomes (Reddy et al., 2020).

In sum, frequent challenging behaviors are associated with significant stress, especially when teachers receive little training in EBMPs, and often results in teachers using ineffective behavior management strategies. Training in EBMPs enhances teacher efficacy, job satisfaction, and reduces turnover but occurs infrequently with implementation practices irregularly studied. High-quality PD that can be implemented successfully is needed to address this gap and support educators in their roles. Therefore, this study will examine a new model of PD, simulation training, which infuses characteristics of high-quality PD to promote in-vivo skill development. In the next section, I will describe the benefit of simulation training as an alternative to traditional PD based on simulation's association with skill development and generalizability.

The Value of Simulation Training

Simulation training refers to practicing professional skills within a virtual setting to improve skill generalization to real-life settings through the re-creation of realistic conditions (Cook et al., 2011; Dieker et al., 2007, 2008; Lintern et al., 1990; Ozaki et al., 2013). While simulation training is relatively new in teacher PD, practicing skills critical to one's profession through simulation is not new in other professional fields (Caro, 1973; Cook et al., 2011; Jentsch

& Curtis, 2017). This section explores how simulation technology supports skill development across various professions and the impact of simulation on learning.

Simulation training has been an integral part of professional fields such as aviation, the military, and medicine for decades, training individuals in fundamental skills to proficiency in safe and realistic environments (Caro, 1973; Cook et al., 2011; Jentsch & Curtis, 2017). Simulation training allows professionals to develop skills necessary for high-stakes situations, such as surgery or flying an airplane, in a virtual setting practice before entering real-life scenarios (Cook et al., 2011; Jentsch & Curtis, 2017). Aviation has been implementing simulation training for over four decades (Caro, 1973). Progressing from stick and rudder control to full motion simulations, aviation has taken a leading role in the development of simulation training, dedicating many hours of its training curriculum to using this learning tool (Caro, 1973; Jentsch & Curtis, 2017). Aviation has progressed simulation training through engineering innovative technology and creating assessments of simulation as a method of learning (Alessi, 2000; Dorsey et al., 2009; Jentsch & Curtis, 2017). The aviation industry even mandates a certain amount of simulation training hours before becoming a licensed pilot (Federal Aviation Administration [FAA], 2019).

Cook et al. (2011) conducted a systematic review of simulation training among medical professions that revealed trainees displayed greater skill development and knowledge after completing simulation training. More than 35,000 health profession trainees, including medical students, physicians, nurses, emergency medical technicians, dentists, veterinarians, and chiropractors, participated in studies examining the outcomes of simulation training (Cook et al., 2011). Simulation was used to practice various procedures, including minimally invasive surgery, intubation, and obstetrics (Cook et al., 2011). Outcomes of the meta-analysis indicated

simulation training, overall, was associated with moderate to large positive effects when compared with no treatment, as were individual outcomes including amount of knowledge gained (ES = 1.20), time to execute skills (ES = 1.14), efficiency of process (ES = 1.09), and quality of the completed task (ES = 1.18; Cook et al., 2011). Cook et al. found a significant improvement of health professionals' behaviors including successful completion of key procedural tasks and procedural errors (ES = 0.79), and positive effects on patient care, defined as patient discomfort, complications, and survival rate (ES = 0.50).

More recently, simulation has been adapted into games and virtual reality to teach users through innovative game design. Games such as World of Warcraft (WoW; Blizzard Entertainment, 2004; Schrader & McCreery, 2008) have merged simulation-based skill development with a modern approach to maintaining the user's attention. Simulation games have been shown to be effective and are used more readily in the field of education due to the infusion of learning principles and theory (Schrader & McCreery, 2008; Sitzmann, 2011).

A meta-analysis of 55 studies, conducted by Sitzmann (2011), illustrated the instructional effectiveness of computer-based simulation games. Results from various fields including education, psychology, business, medicine, and science, indicated that simulation games were associated with improved self-efficacy (confidence in understanding and implementing learned information; $d = 0.52$), declarative knowledge (recall of facts and concepts and the relationship between each; $d = 0.28$), procedural knowledge (how to perform a task or action; $d = 0.37$), and retention (retaining information over time; $d = 0.22$; Sitzmann, 2011). A key finding of the Sitzmann meta-analysis is that engagement in the simulation game determined how effective the instruction was ($d = 0.49$). This finding highlights the importance of active engagement on learning outcomes (Sitzmann, 2011).

Dieker et al. (2014) recommended three critical components for an effective simulated environment in the context of teacher training: (1) personalized learning, (2) suspension of belief, and (3) cyclical procedures to ensure impact. The three components work together to allow the user to practice a realistic response in a particular situation and gain the benefit of analyzing what mistakes they made, how they can improve the interaction and, most importantly, attempt the same situation again with a different response (Dieker et al., 2014). The idea of virtual rehearsal separates simulation technologies from traditional skill development modalities because the user improves their skills in repeated, live situations to the point of proficiency that could not occur in the classroom (Dieker et al., 2014).

Given the success of simulation training in other professions, education and psychology are beginning to develop simulation programs, however, the uptake has been slower (Dieker et al., 2014; Jentsch et al., 2011). Despite the promise of simulation training established in other fields, more work is needed in education and psychology to understand how these programs work. Therefore, the goal of this study is to evaluate the implementation of a newly developed simulation training for teachers to identify factors that are associated with greater skill development.

Integration of Simulation within Teacher Professional Development

There is a growing body of literature to support simulation training in teacher preparation (Beisiegel et al., 2018; Dieker et al., 2007; Girard et al., 2012; Schrader et al., 2011; Schussler et al., 2017). Simulated environments are an advantageous format for teacher trainings in EBMPs for several reasons. Teachers can focus on improving specific targeted skills and practice their skills in a context-specific environment with multiple trials at ‘real-life’ scenarios without the risk of diminishing resources such as time, money, or relationships (Dieker et al., 2014).

Furthermore, simulation training may promote the integration of EBMPs that are associated with positive outcomes for skill acquisition, generalization, and increased student achievement (Schrader & McCreery, 2008; Sitzmann, 2011).

There are two EBMPs training simulators that are most relevant to the current study. First, Dieker et al. (2008) created a mixed reality virtual environment, TeachME™, designed for educators to support students in a myriad of areas including academic, behavioral, and social-emotional. Since it's conception, TeachME™, now named TeachLivE™/Mursion, has been studied in a variety of educational settings and has demonstrated improvements in EBMPs skill proficiency, acceptable usability perceptions, and varying degrees of generalization to the classroom (Dawson & Lignugaris-Kraft, 2017). In TeachLivE, educators navigate through various scenarios interacting with human-operated avatars within a realistic-looking classroom projected onto a screen (Dieker et al., 2008). Rosenberg et al. (2021), in a sample of 20 preservice teachers without prior training in EBMPs, found teachers were able to identify appropriate EBMPs to use in a classroom behavior plan after using TeachLivE. However, data were only collected on the teachers' ability to implement the EBMPs following TeachLivE™/Mursion training; data were not collected prior to the training. Therefore, changes in teachers' skill use from pre- to post-treatment was unable to be determined (Rosenberg et al., 2021).

In another study examining TeachLivE™/Mursion, Pas et al. (2016), looked at the effects of guided, simulated practice on teachers' use of proactive EBMPs. Among a small sample, researchers found statistically significant increases in proactive behavioral expectations ($p < .01$; $d = 0.92$) and approval ($p < .01$; $d = 1.06$) with sustained effects at the three-month follow-up. A statistically significant decrease in noncompliance ($p = .04$; $d = –0.36$), was also detected. While

TeachLivE^{TM}/Mursion is perhaps the most researched EBMPs simulation training models at the moment, there are only a handful of studies looking at the effectiveness of the model in the context of teaching EBMPs as well as the usability of the technology (Ersozlu et al., 2021).

SimSchool (Zibit & Gibson, 2005) is another teaching simulation wherein teachers are responsible for designing tasks within a virtual classroom that simulated students' responses based on their individual personalities and emotional state. Teachers then make decisions on how to respond to students and receive feedback to improve their decision making and understanding of the factors that lead to success for students (Christensen et al., 2011; Zibit & Gibson, 2005). Fifty-seven preservice teachers trained in EBMPs with SimSchool rated improvement in classroom management skills (M = 3.32) and behavior management skills (M = 3.25) on a 5-point scale (1 = Not at all, 5 = A great deal; Stavroulia et al., 2014). However, many participants (70.2%) found it difficult to play with the simulation (Stavroulia et al., 2014). A 2014 mixed-methods study conducted by Rayner and Fluck among 15 preservice teachers revealed that participants perceived SimSchool as helping them learn about effective classroom management strategies (M = 2.80) on a 5-point scale (1 = strongly agree, 5 = strongly disagree).

Simulation technology has emerged as a promising avenue for professional development for teachers. However, there are very few studies examining the implementation process in depth. The current study aims to address this gap in the literature.

Implementation Science

Implementation science examines how treatments and programs are implemented (Fixsen et al., 2005; Proctor et al., 2011). Implementation science principles inform how a simulation training model works in schools. These principles help separate the effectiveness of implementation from treatment outcomes and examine the degree to which various aspects of

implementation interfered with teachers interacting with simulation training (Larson et al., 2020; Proctor et al., 2011). Implementation science provides a framework to learn how much and why the technology was or was not used (Vince Garland et al., 2016; Proctor et al., 2009). Simulation training developers examine many variables including satisfaction with the innovation, compatibility in the given setting, whether it was delivered as intended, and sustained use of the innovation throughout the organization (Vince Garland et al., 2016; Proctor et al., 2009). There are three implementation science principles in the literature that are most relevant to the current study: dosage, fidelity, and usability.

Dosage

Dosage is defined as how much of a treatment or training was delivered (Proctor et al., 2011). Training dosage is important to consider when implementing EBMPs in naturalistic settings such as schools, due to effectiveness trials detailing the amount or dose of a treatment needed to produce effective outcomes (Dane & Schneider, 1998; Schoenwald & Hoagwood, 2001). EBMPs undergo many trials to evaluate the specific treatment characteristics, including dosage, attributed to outcomes and to determine if the outcomes are positive before being considered evidence based (Schoenwald & Hoagwood, 2001).

Durlak and DuPre (2008) found that effective implementation is associated with better outcomes in their five meta-analyses consisting of 483 studies and 59 additional studies looking at the influence of implementation dimensions (fidelity, dosage, quality, responsiveness, program differentiation, adaptation, monitoring control/comparison conditions, and program reach) on treatment outcomes. However, of the 29 studies that measured the influence of dosage on outcomes, 17 studies revealed a significant positive relationship with at least one half of all program outcomes (Durlak & DuPre, 2008).

August et al. (2003) evaluated the effectiveness of Early Risers "Skills for Success," an EBMPs program focused on prevention for children displaying early aggression. The program focused on reducing aggression and other internalizing behaviors and improving academic abilities and social skills. In a two-year trial, participants were randomized to one of three groups: full treatment model, partial strength treatment model, and no-intervention control. Of particular interest, the authors assessed how dosage, calculated as the number of hours present for the treatment, impacted the stated outcomes (August et al., 2003). Results showed that among the 218 children involved in the treatment condition, there was a significant positive relationship between treatment hours and teacher-rated social skills and internalizing problems (August et al., 2003). This held true when outcomes were compared with the control group (August et al., 2003).

While dosage is associated with development of EBMPs and is a dimension used to determine treatment efficacy in clinical trials, studies do not frequently examine the effect of dosage on the use of EBMPs in naturalistic settings (Durlak & DuPre, 2008; Novins et al., 2013; Powell et al., 2020). Therefore, one aim of this study is to understand the impact of dosage on the use of EBMPs following simulation training.

Fidelity

Fidelity examines the degree to which a program was implemented as intended, such that all prescribed components were executed accordingly (Proctor et al., 2011). An example to illustrate the differences between dosage and fidelity is as follows: a person prescribed a daily pill for five days would demonstrate high fidelity if they took the pill each day. If the individual took all the pills on last day, they would have taken the correct dosage but had a low rate of fidelity. In the context of real-world implementation of EBMPs, rates of fidelity historically have

been low, even though higher rates of fidelity are associated with better child outcomes (Novins et al., 2013; Real & Poole, 2005; Robbins et al., 2011; Schoenwald et al., 2008; Schoenwald et al., 2009). Despite newer studies identifying the link between fidelity and positive outcomes, little research on simulation training examines the impact of fidelity on outcomes.

Research on TeachLivE has contributed to our knowledge of the effectiveness of simulation training at teaching EBMPs, although none have examined the influence of fidelity on treatment outcomes (Dawson & Lignugaris-Kraft, 2017; Straub et al., 2014; Vince Garland et al., 2012). Prior research on TeachLivE assessed the average treatment fidelity of each training session but did not focus on answering questions related to impacts of fidelity on overall implementation and effectiveness (Dawson & Lignugaris-Kraft, 2017; Straub et al., 2014; Vince Garland et al., 2012, 2016).

Fidelity is rarely evaluated during the implementation of EBMPs training despite having significant influence over treatment outcomes (Carroll et al., 2007; Dane & Schneider, 1998; Durlak & DuPre, 2008; Mihalic, 2004). Fidelity also has not been examined in the few EBMPs simulation trainings that exist, which underscores a gap in understanding the impact of fidelity on treatment effectiveness when implementing simulation training. IVT-T utilizes a self-guided format that relies on teachers to use it independently, which poses a risk to the implementation fidelity and is therefore an important aspect to be evaluated.

Usability

Usability of IVT-T will be evaluated due to its relevance when developing technology and innovations. Usability is defined as the ease with which users interact with and conduct tasks on a system (Badiee & Kaufman, 2015; Bowman et al., 2002). During the formative stages of development, the usability of new systems is assessed to inform ongoing improvements

(Bowman et al., 2002; Schneiderman, 1992). Usability has not been assessed as a dimension of EBMPs implementation because the treatments are not technologically based (Badiee & Kaufman, 2015; Bowman et al., 2002). However, the limited literature that does exist on EBMPs simulation training highlights the importance of assessing the usability of a system to ensure strong implementation.

Badiee and Kaufman (2015) evaluated the usability of SimSchool and found that ease of use was rated the lowest and the user interface and general navigation were the least liked features. However, their study did not examine the impact of usability on overall use of SimSchool or teacher outcomes (Badiee & Kaufman, 2015).

There is a similar dearth of research on the impact of usability on TeachLivE™'s effectiveness. Based on a review of the literature conducted in 2021, TeachLivE™ has been studied across 102 publications with approximately 55 of those studies using an experimental or mixed methods design (Ersozlu et al., 2021). While some studies examine the implementation feasibility and fidelity of the model, to our current knowledge, none have evaluated TeachLivE™'s ease of use or the impact of its usability on teachers' desire to continue using the system. Vince Garland et al. (2016) describes participants' attitudes towards TeachLivE™ labelled as social validity. Authors reviewed focus group responses of five participants' experience within the simulated environment, noting the realism of the setting and ease of suspending disbelief. However, quantitative data were not collected, and usability was not explicitly studied nor were the impacts of social validity on treatment use or study outcomes (Vince Garland et al., 2016).

More work is needed to understand the implementation practices that facilitate and impede transfer of EBMPs, which is the goal of the current study. Simulation training holds

promise for improving skill development and transfer but requires further examination to determine the implications unique to the implementation of a novel training program. Thus, the current study will assess if training dosage, fidelity, and usability of IVT-T is associated with generalizing EBMPs to the classroom. This information helps fill a gap in the literature on implementing simulation training and informs future integration of technology-based programs into real-world settings.

Design and Research Questions

EMBPs are essential, frequently requested professional development skills that can be taught within a simulated learning environment to help teachers feel more competent and reduce burnout (Dieker et al., 2014; Wei et al., 2010). While EBMPs simulation training is growing within the education field, little is known about how implementation procedures and facets of the technology influence teachers' learning and interaction with simulators (Christensen et al., 2011; Vince Garland et al., 2016; Shernoff et al., 2018; Stavroulia et al., 2014). This study drew on extant data from a federally funded grant study which employed a between group, two condition by two time point, quasi-experimental design. The study examined the association between the amount of time teachers spend using IVT-T and use of EBMPs in the classroom and whether weekly exposure to IVT-T was associated with level of skill use. Additionally, this study explored if usability was related to the amount of time spent playing IVT-T among a sample of teachers.

Research Question 1*:* How frequently did teachers use IVT-T (total minutes per week and weekly use)? Descriptive information will be analyzed to better understand teachers' patterns of use of IVT-T to support future implementation.

Research Question 2: Does the total amount of time teachers play IVT-T (dosage) predict changes in use of EBMPs at time 2? It was predicted that higher IVT-T dosage would predict larger positive change in use of EBMPs.

Research Question 3: Is there an association between perceived ease of use (usability) and the amount of time teachers spend playing IVT-T? It was predicted that there would be a positive association between usability and time spent using IVT-T.

Method

Schools

The larger study was conducted in the final year of a development and innovation grant. The initial two years of the IVT-T grant focused on developing and refining the simulator with the final two years devoted to evaluating whether the virtual training improved student engagement and achievement (Shernoff et al., 2018). The larger study aimed to examine the promise of IVT-T using a between group, quasi-experimental design that assigned schools to an intervention (full model online course + IVT-T + Professional Learning Communities) or comparison (online course + standard district induction and mentoring) condition. In the larger grant-funded study, data were collected at two time points, Time 1 (Fall 2018) and Time 2 (Spring 2019), from teachers, school personnel, and students at risk for challenging behaviors. Data were collected at each school through observation, teacher-report, and measures of student achievement.

School districts were recruited from a list of schools meeting the study's inclusion criteria (70% or more students eligible for free or reduced lunch). Recruitment strategies included emailing district administrators to introduce the project and meeting with interested schools to secure school approval. Six schools were recruited and assigned to either the intervention or

comparison condition, resulting in three schools per condition in the larger funded study. In the current dissertation study, teachers in three schools in the intervention condition who had access to IVT-T were included in these secondary analyses.

In participating schools, 90.7% of students were African American or Black, 8.1% Hispanic or Latino, 0.5% White, 0.1% Asian and 0.6% Other. Seventy two percent of students were eligible for free or reduced lunch (Range = 70–78%) and median household income was $49,509. The combined four-year graduation rate in the district is 78%, and 42% of students attend college.

Teachers

This study used data from teachers (N = 48) with access to IVT-T. General and special educators who taught kindergarten through seventh grade and between the ages of 22 and 65 were eligible to participate. Teachers were not excluded based on ethnicity/race, gender, or religion. Teachers volunteered for the program on a first-come-first-served basis.

Sample characteristics are described in Table 1. In brief, 75% of consented teachers were female, 25% held Bachelor's Degrees, and 45% held Master's Degrees (20% missing). Mean age was 44 years (SD = 12.1 years, 14–64 years old), mean years of teaching experience was 14.53 (SD = 9.39, 1–36 years). Forty three percent of teachers self-identified as African American or Black, 8% as Hispanic or Latino, 16% as White, 2% as Asian, 2% as American Indian or Alaskan Native, and 8% as Other.

IVT-T Simulation Training

IVT-T simulates real-life classroom scenarios providing teachers with a supported environment to practice EBMPs. By leveraging simulation technology, IVT-T expands the role and reach of technology in PD. Practicing EBMPs in a low-stakes simulated environment

increases the likelihood of skill generalization and only requires a computer to utilize the program which increases teacher access to EBMPs (Dieker et al., 2014).

The aim of IVT-T is to support teachers in developing their knowledge and increase their use of EMBPs by practicing their skills with virtual students engaging in challenging behaviors. A functional behavioral framework and common practices approach were adopted to guide the interactions within IVT-T storylines. Creators applied common effective practices to the antecedent-behavior-consequences model of behavior to illustrate the environments contribution to maintaining challenging (Shernoff et al., 2021).

Teachers interact with three components of IVT-T, characters, classrooms, and storylines. Teachers interact with three different characters in first or sixth grade that engage in off task and aggressive behaviors. Teachers interact with these characters in developmentally realistic first or sixth grade classrooms. Characters behave and react in accordance with storylines created to illustrate challenging interactions (Shernoff et al., 2018). Storylines were written to incorporate EBMPs that inform the prevention and management of challenging behaviors (e.g., praise, ignore, redirect, proximity, instructions, punishment; Kazdin 2005; Simonsen et al., 2008). Teachers are presented with response options that allow them to diminish challenging behaviors and re-gain student attention, compliance, and engagement (Shernoff et al., 2017, 2018). Depending on the response option selected by the user, characters' challenging behaviors will escalate or de-escalate and become more engaged. Therefore, if the EBMP response option was selected, the character's challenging behaviors would diminish and appear more engaged and on task. However, if the non-EBMP response option was selected, the character's challenging behaviors would increase and become more intense. These interactions are reflective of typical behavioral progressions due to environmental influence seen in real classrooms (Kazdin, 2005).

In addition, IVT-T incorporates characteristics of PD that are correlated with positive outcomes by providing feedback and reflection, continuous in-vivo practice, and interactive learning (Elliott, 2017; Latouche & Gascoigne, 2019; Lutrick & Szabo, 2012; Shernoff et al., 2017). The initial interaction with virtual students was designed for teachers to practice responding to challenging behaviors. The second phase provided teachers the ability to reflect on the response options they chose during the practice phase and identify ways to improve on their selected responses in subsequent training sessions. In the third phase, teachers reviewed quantitative and qualitative feedback information about their performance, the effectiveness of their responses, and how to improve in the future. Teachers receive frequent feedback and opportunities for reflection at the end of each practice session (Shernoff et al., 2017). Feedback to teachers details the type of strategies they selected, effectiveness of their choices, time spent practicing, and time it took to make decisions in the practice session in the form of a system-generated skill report card (Shernoff et al., 2017). The simulation training is available through a website domain that allows for extensive opportunities for practice to develop teachers' skills (Shernoff et al., 2017).

Procedures for Larger Study

Teachers were consented at the beginning of the school year between September and October. Teachers completed demographic surveys and baseline observations were conducted during instructional periods to assess their current use of evidence-based behavior management and instructional practices. Observational data were collected using the Classroom Strategies Assessment System (CSAS; Reddy et al., 2013), an observational system that quantifies the behavioral management skills used by a teacher during a period of instruction. Undergraduate

and graduate students who were trained in and met 80 percent rating reliability in the CSAS, conducted two 30-minute classroom observations during academic instruction.

Following baseline observations, teachers were instructed to play IVT-T between 40 to 60 minutes each week, logging in at least once per week, for the duration of a school year. Teachers had access to IVT-T Prototype 4.0 which had three characters available, Michael, Jordan, and Ava. Michael was a first-grade student with inattention and hyperactivity. Jordan was a sixth grader that struggled with defiance, verbal aggression, and physical aggression. Ava was a sixth grader that was inattentive and had difficulty following directions. Teachers engaging with the Michael storyline were focused on supporting him to complete independent reading and locate materials. Teachers engaging with the Jordan storyline were focused on supporting him in completing assigned work after arriving late and being off task (Shernoff et al., 2021). Each storyline was designed to maximize teachers' opportunities to manage challenging behaviors, reflect on their performance, and receive feedback on the effectiveness of their choices.

Measures were administered at the beginning of the school year on a rolling basis between September and October (Time 1), and again at the end of the school year between May and June (Time 2). Usage data were collected electronically and automatically through the IVT-T program. Teachers were observed again and completed measures including the System Usability Scale (SUS) – the measure of relevance to this study.

Measures

Fidelity

Fidelity to the program is conceptualized as frequency of playing IVT-T and is operationalized as the number of weeks during which teachers used IVT-T for 40 or more minutes throughout the year. There were a total of 23 weeks teachers had access to IVT-T

following baseline observations and before data collection. Frequency counts will be grouped by week with a maximum frequency count being 23. Time spent playing IVT-T each week will also be examined as a facet of both fidelity and dosage. Login and logout time stamps collected by system analytics will support determining teacher fidelity to IVT-T.

Dosage

In the current study, dosage was operationalized as the total amount of minutes teachers engaged with IVT-T throughout the year. Tracking logs captured teachers use of the program on an ongoing basis. Login and logout time stamps measured the frequency with which they used IVT-T.

Usability

Responses to a measure of usability were collected after teachers interacted with IVT-T with the aim of understanding the relationship between teachers' perceptions of ease of use and how much time they played IVT-T. Usability will be measured using the SUS.

The System Usability Scale has been used extensively to measure perceived ease of use for a wide range of software, hardware, cell phones, and websites. The format of the questionnaire is a 10 item Likert scale measuring a respondent's subjective assessment of a system's usability (Brooke, 1996). Items include "I found the system unnecessarily complex" and "I felt very confident using the system." Responses are on a 5-point scale (1 = *strongly disagree* to 5 = *strongly agree*) and produce a single score indicative of global usability. Scores range from 0 to 100 with higher totals representing greater usability of a system. The SUS has strong reliability (Cronbach α ranging from 0.83 to 0.97) and significant correlations of concurrent validity, ranging from 0.22 to 0.96 (Lewis, 2018). In addition to being a widely used

measure of usability, the SUS provides an accurate sample mean with small sample sizes (Lewis, 2018; Sauro, 2010).

Behavior Management Strategies

The CSAS (Reddy et al., 2013) is an observational measure designed to assess teachers' use of evidence-based instructional strategies and behavioral management skills, identify and develop skill goals, and monitor progress towards the goals (Reddy et al., 2013). There are three parts of the CSAS (1) Strategy Counts (frequency of strategies used), (2) Strategy Rating Scales (quality with which strategy was used), and (3) The Classroom Checklist (presence of classroom resources and procedures; Reddy & Dudek, 2014). Observers count the frequency with which a teacher demonstrates eight instructional and behavioral management strategies. Next, observers rate the quality with which the teacher implemented instructional and behavioral management strategies on a 7-point Likert-type scale (1 = *not used*, 4 = *sometimes used*, 7 = *always used*). Ratings are based on information gathered from the Strategy Counts, targeted notes on nine dimensions completed during the observation, observation notes of the lesson content, academic engagement, and learning for the lesson, and knowledge of effective strategies learned during CSAS training (Kettler et al., 2019; Reddy et al., 2013).

The Strategy Rating Scales are composed of ratings based on the discrepancy between the observed frequency of a strategy and the recommended frequency of that strategy by the observer's approximation. The Strategy Rating Scales contributes information on the strategies teachers should be using more often, less often, or equally often. The Classroom Checklist is the final component used to collect data with the CSAS and is completed by the observer endorsing whether 10 common classroom features are present or absent in a teacher's classroom (Reddy et

al., 2013). The CSAS demonstrates strong internal consistency (Cronbach αs of .93 and .92), as well as good interrater reliability (r = .72 to .94).

Data from the behavioral management strategy counts were used to quantify the frequency with which teachers use EBMPs at Time 1 and 2. Strategies that were analyzed include behavioral praise statements and behavioral corrective feedback. Behavioral praise and behavioral corrective feedback are represented by a single sum score that indicates the total number of times the behavioral management strategy was used. The strategies were analyzed independently due to predicting different directional outcomes (i.e., an increase in behavioral praise, decrease in behavioral corrective feedback).

Demographic Survey

Teachers completed a basic demographic survey at Time 1. Teachers provided information related to gender, age, race/ethnicity, educational background, teaching experience, and prior training in EBMPs.

Analyses

Research Question 1

How frequently did teachers use IVT-T (total minutes per week and weekly use)?

Research question 1 was analyzed descriptively, which included examining weekly use of IVT-T, total minutes per week, and whether teachers met the 40-mintue threshold. Frequency graphs were examined to understand teachers' patterns of use of IVT-T, relationship to outcomes, and correlation with the dosage variable. Descriptive statistics that were examined included average IVT-T minutes played weekly, percentage of teachers that played weekly, number of weeks played, and change in behavioral praise and behavioral corrective feedback compared to total

weeks played. A new variable was computed to visualize the change in CSAS strategy counts from Time 1 to Time 2 as it related to weekly use.

Research Question 2

Does the total amount of time teachers play IVT-T (dosage) predict changes in the frequency with which teachers used behavioral praise and behavioral corrective feedback at Time 2? It was predicted that higher dosage would predict larger positive change in the use of behavioral praise and behavioral corrective feedback at time 2. In addition to descriptive analyses, behavioral praise and behavioral corrective feedback strategy counts data were analyzed with negative binomial regression models for over-dispersed count data (Beaujean & Morgan, 2016). After comparing the fit of various generalized linear models to these data, negative binomial regressions were used for several reasons. Negative binomial models support the analysis of count data that often do not adhere to assumptions of distribution normality. Poisson modelling was not selected, even though it conducts a similar analysis of count data, because it requires outcomes variables to adhere to a Poisson distribution, meaning the mean and variance of the outcome variable are equal. Both outcome variables' (i.e., behavioral praise, behavioral corrective feedback) variances were much larger than their means, demonstrating overdispersion and violating an assumption required to conduct Poisson regressions. In addition, zero-inflated negative binomial (ZINB) regressions were considered but not selected due to behavioral corrective feedback not having zero values, which is an assumption of the ZINB model. Further, the outcome variable, behavioral praise, displayed negligible differences in the goodness-of-fit between the negative binomial regression model and ZINB model. More specifically, the Akaike's information criterion (AIC) was examined to determine the goodness-of-fit of each model for behavioral praise, which indicated an insignificant difference between

the negative binomial regression model and ZINB (Beaujean & Morgan, 2016). Therefore, the negative binomial regression model was selected to maintain uniformity when reporting outcomes across variables and to promote ease of interpretation.

Two negative binomial regressions examined the association between dosage and behavioral praise, and dosage and behavioral corrective feedback. In the first negative binomial regression, the predictor variable was dosage, measured by total IVT-T minutes, and the outcome variable was total behavioral praise count at Time 2. The total behavioral praise count at Time 1 was entered as a covariate. In the second regression, the predictor variable was dosage, and the outcome variable was total behavioral corrective feedback count at Time 2. The total behavioral corrective feedback count at Time 1 was entered as a covariate. For research question 2, a general linear model prospective power analysis yielded power was sufficient (80%) to detect medium and large effect sizes at the $\alpha = .05$ level (Green & MacLeod, 2016). Power was sufficient to detect a large effect size at the $\alpha = .01$ level.

Research Question 3

Is there an association between perceived ease of use (usability) and the amount of time teachers spend playing IVT-T? The prospective power analysis for Spearman's rank order correlation indicated the current sample size was sufficient. A scatterplot analysis revealed the data did not appear perfectly linear indicating Spearman's rank order correlation may better capture the relationship between usability and dosage (Zar, 2005). The independent variable was usability, measured by the SUS, and the dependent variable was time using IVT-T, measured as total IVT-T minutes collected through system analytics.

Missing Values Analysis

Prior to running analyses, data were screened for invalid scores, outliers, and missing data. No invalid scores were found, and twelve scores were identified as outliers. After analyzing the scores, two of the twelve identified outliers were three or more standard deviations from the mean and omitted due to artificially inflating the results (Aguinis et al., 2013; Osborne & Overbay, 2004). A missing values analysis indicated 18% to 29% missing data on nine variables of interest. Little's MCAR test (Little, 1988) was conducted to analyze patterns of missing data and could not rule out that data were missing completely at random (MCAR; $\chi^2(24) = 13.25$, $p = .96$). Therefore, general ignorable procedures were followed for cases with missing data.

Results

Descriptive statistics are presented in Tables 2 and 3. Analysis of histograms, scatterplots, and frequency graphs revealed that behavioral praise and behavioral corrective feedback appeared to be distributed consistent with expectations for count data (White & Bennetts, 1996). Total IVT-T minutes and IVT-T weeks were also positively skewed.

Pearson product-moment correlations evaluated the relationship between independent and dependent variables (see Table 4). Total IVT-T minutes and IVT-T weeks were significantly correlated ($r = .95$, $p < .01$), indicating a large, positive linear relationship. This means that teachers who spent more time playing IVT-T also tended to play IVT-T for more weeks. Behavioral praise scores at Time 2 were significantly correlated with total IVT-T mins ($r = .40$, $p < .05$), and IVT-T weeks ($r = .53$, $p < .01$). These were moderate and large, positive linear relationships suggesting teachers that used more behavioral praise at Time 2 also spent more time and weeks playing IVT-T. Behavioral corrective feedback scores at Time 2 and behavioral corrective feedback scores at Time 1 were significantly correlated ($r = .47$, $p < .01$), indicating a moderate, positive linear relationship.

How Frequently Did Teachers Use IVT-T?

Table 5 and Figures 1 and 2 present descriptive statistics of teachers' weekly use of IVT-T, including total weeks played, total time played each week, and teachers' ability to reach 40-minutes of play. Table 5 presents a frequency table of the total number of weeks teachers played IVT-T. Descriptive analyses indicate a bimodal distribution in which 92% played IVT-T between 1-10 weeks and 8% teachers played more than 10 weeks (Range = 1 to 21). On average, teachers played IVT-T 20% of the total weeks they had access.

Figure 1 displays the average amount of minutes teachers played IVT-T each week, as well as the adjusted weekly average, which was the average minutes played by teachers that used IVT-T for 1 or more minutes. Teachers that logged 0 minutes were not included in the adjusted weekly average score. All teachers played on average less than 25 minutes per week except during week 4 when teachers averaged 40.1 minutes (*SD* = 8.3 mins, Range = 3.3 - 40.1). When adjusting for non-players, the weekly average minutes met the 40-minute threshold 70% of the weeks (*M* = 51.5, *SD* = 16.7, Range = 19.6-82.7).

Teachers' use of IVT-T from week to week was variable. The largest percentage of teachers (20-33%) played IVT-T during weeks 2, 5, 6, 8, 9, 10, 11, 20, 22, and 23, with the exception of week 4 when 77% of teachers played IVT-T. Furthermore, approximately 64% of teachers that played for more than one week, played IVT-T in consecutive weeks, ranging from 2 consecutive weeks to 20. When not meeting the 40-minute threshold, teachers played between 20-39 minutes 48% of the time, 15-19 minutes 17% of the time, 10-14 minutes 21% of the time, and 1-9 minutes 14% of the time. The lowest adjusted average weekly minutes occurred during week 5 (*M* = 19.6, *SD* = 15.4, Range = 1-45). Teachers played IVT-T for the highest number of adjusted average minutes during weeks 21 (*M* = 81.9, *SD* = 45.4, Range = 46-153) and 23 (*M* =

82.7, *SD* = 56.9, Range = 32-183). The total number of teachers who played IVT-T were highest during week 4 (*N* = 37) followed by week 2 (*N* = 16).

When comparing weekly use of IVT-T to the change in CSAS strategy counts from Time 1 to Time 2, there appeared to be a moderate amount of heterogeneity (see Figure 2). Specifically, of teachers that played between 1-5 weeks (*N* = 37) and were observed twice (*N* = 24), 37.5% provided less behavioral praise at Time 2, 16.7% provided the same amount, and 45.8% made an increase of 1 count or more. Similarly, for behavioral corrective feedback, 29.2% provided more corrective feedback at Time 2, and 66.7% provided less. Across all teachers, while it appeared that greater weekly use of IVT-T was related to positive outcomes, the overall heterogeneity in the data was significant. In addition, dosage and fidelity were significantly correlated ($r = .95$, $p < .001$), indicating weekly use may be too similar to total IVT-T minutes to be able to draw independent conclusions about the function of weekly use of IVT-T on teachers' use of EBMPs.

Does IVT-T Dosage Predict EBMP Use?

Behavioral Praise

A negative binomial regression examined whether IVT-T dosage predicted higher rates of behavioral praise at Time 2, when controlling for teachers' use of behavioral praise at Time 1. The negative binomial regression revealed total IVT-T minutes was not a statistically significant predictor of behavioral praise at Time 2 when controlling for behavioral praise at Time 1 ($b = 0.001$, $p = .06$).

Behavioral Corrective Feedback

A negative binomial regression also examined whether the amount of time teachers used IVT-T predicted rates of behavioral corrective feedback at Time 2, when controlling for

teachers' use of behavioral corrective feedback at Time 1. The negative binomial regression revealed total IVT-T minutes was not a significant predictor of behavioral corrective feedback Time 2, when accounting for behavioral corrective feedback Time 1. However, behavioral corrective feedback Time 1 was a significant predictor of behavioral corrective feedback at Time 2 ($b = 0.0138$, $p < .005$).

Does Usability Impact IVT-T Dosage?

A Spearman's rank-order correlation assessed the relationship between IVT-T usability and the time teachers spent playing IVT-T (dosage). Usability was measured using the SUS and time spent playing IVT-T was measured by IVT mins. The correlation between usability and dosage was significant $r_s(39) = .403$, $p = .011$.

Discussion

The purpose of this study was to examine the frequency with which teachers used IVT-T, the association between dosage and usability, and teachers' use of EBMPs in the classroom. Specifically, this study examined the characteristics of teachers' weekly use of IVT-T and examined the relationship between the dosage of IVT-T and teachers' use of EBMPs in the classroom. The relationship between system usability and time spent using the simulator was also examined. This research helps to inform the way simulation training in schools can be scaled up and disseminated in schools.

How Frequently Did Teachers Use IVT-T?

I found that, on average, teachers played IVT-T at a lower dose than recommended. Specifically, teachers played IVT-T for less than 25 minutes per week for an average of 20% of the total weeks studied. However, a subset of teachers met the weekly threshold of 40 minutes played each week. These findings suggest that most teachers used IVT-T for a small portion (3-5

weeks) of all weeks studied and for less than the recommended dose. These findings will help inform benchmarks for recommended IVT-T dosage. In the current study, although teachers did not meet the weekly threshold of 40 minutes, almost one half played between 20-39 minutes. This finding suggests that the 40-minute recommendation may exceed the amount of time teachers had to use IVT-T. Teachers may be able to better fit 20-39 minutes per week of training into their schedule than the recommended 40 minutes. This is particularly relevant given teachers were asked to use the system outside of the instructional day and/or on their own time.

These findings are also consistent with prior research demonstrating that teachers use technology-based EBMP programs at a lower dose than prescribed (Mixon et al., 2019; Owens et al., 2018; Owens et al., 2019). Mixon et al. (2019) examined elementary school teachers' use of an online intervention and found that slightly more than one half (51.5%) of the teachers used the intervention for at least 8 weeks and the other one half used the intervention for approximately 2.5 weeks. Despite the low usage rate, Mixon et al. found statistically significant improvements in student behavior (Mixon et al., 2019). These findings suggest that teachers are able to achieve intended outcomes despite not meeting recommended dosage benchmarks set by researchers.

Prior studies of IVT-T similarly found lower training dosages while teachers still reduced their use of punitive strategies and between 62% and 71% of teachers demonstrated an improvement in their perceived ability to manage challenging behaviors (Shernoff et al., 2021). Taken together, the conclusions of prior research suggest that lower dosage – both total weeks and weekly minutes – is not uncommon when implementing teacher training programs.

Another finding revealed that 64% of teachers used IVT-T during consecutive weeks, with an average of 5 consecutive weeks, and most teachers played IVT-T in weeks 2 and 4 followed by week 6 and week 9. This suggests that teachers were able to dedicate consecutive

weeks to playing IVT-T but that 23 weeks of training is not feasible. While Mixon et al. (2019) did not look at consecutive use, they found differences in the way in which teachers adopted the online intervention. Teachers, for example, were categorized as long-term (i.e., began use of the intervention in under 5 weeks, sustained use for 8 or more weeks) and short-term adopters (i.e., began use of the intervention within 10 weeks, sustained use for less than 8 weeks). Mixon also found certain demographic characteristics were associated with being a long-term user, including having a Master's degree, previous experience with the intervention, and older age. Although the current study was not set up to examine or distinguish between short-term and long-term users, the Mixon findings suggest that short-term adopters can still experience positive outcomes with brief interventions. In addition, although Mixon et al. did not examine consecutive use, the nature of their study was such that consecutive use was an inherent feature of teachers' participation. Consecutive use of the intervention was not accounted for in the present study but is worth further investigation to understand the implications on teacher skill development. Future work should be focused on examining brief, consecutive EBMPs with demographic characteristics as predictors of use.

Developing new forms of professional development to support teachers in acquiring EBMP skills requires a realistic understanding about how teachers engage with the training to maximize treatment outcomes. Although the present study did not yield a consensus on how often teachers should use IVT-T, dosage and fidelity have been shown to be important components of program implementation and success and should therefore be incorporated and evaluated when implementing IVT-T in the future.

Does IVT-T Dosage Predict EBMPs Use?

Results indicated that dosage was not a significant predictor of behavioral praise when controlling for behavioral praise at Time 1. This means that teachers who used IVT-T for more time did not necessarily use more behavioral praise at the second observation. This result is somewhat unexpected given that prior research on PD dosage has indicated a positive relationship between dosage and EBMPs (Holmes et al., 2021; Klaft & Codding, 2022; Pianta et al., 2014; Reinke et al., 2013; Reyes et al., 2012; Sorlie et al., 2015; Wei et al., 2010; Zhai et al., 2010). For example, in a non-randomized experiment, Sorlie et al. (2015), studied the effects of a behavior management intervention implemented in 17 primary schools on teachers' use of positive teaching strategies as compared with 20 control schools. Sorlie et al., found training dosage to be a statistically significant moderator of teachers using behavioral praise after the intervention (2015). The 2015 study examined other moderation variables, including implementation quality, which was positively related to greater use of behavioral praise. The relationship between dosage and behavioral praise, however, has yet to be studied extensively and studies of simulation training models have drawn even fewer scholarly studies (Floress et al., 2017; Trahan et al., 2019).

Similarly, in the present study, dosage was not a significant predictor of behavioral corrective feedback when controlling for behavioral corrective feedback at Time 1. This means that teachers who used IVT-T for more time did not necessarily use less behavioral corrective feedback at the second observation. This result is also incongruent with some research on PD dosage which, though only sparsely examining behavioral corrective feedback, has indicated a relationship between dosage and behavioral corrective feedback (Allday et al., 2012; Hirsch et al., 2019; Reinke et al., 2013). Reinke et al. (2013), investigated the association between the dosage of teacher coaching and teachers' use of EBMPs including behavioral corrective

feedback. Results indicated that teachers who received greater levels of coaching dosage had lower rates of behavioral corrective feedback when compared with teachers who received lower doses of coaching (Reinke et al., 2013). Considering incongruence with some prior research and gaps in the literature, the results of the present study related to dosage require further investigation and explanation.

Despite the lack of relationship between dosage and behavioral praise and corrective feedback in this study, mean behavioral praise counts increased at time 2, and behavioral corrective feedback counts also decreased at Time 2. Behavioral praise went from 4.2 mean counts at Time 1 to 8.8 mean counts at Time 2 and behavioral corrective feedback increased from 29.3 to 20.2. This result is consistent with Reddy et al.'s 2013 study which highlighted that a teacher's goal may not be to reduce or eliminate behavioral corrective feedback entirely but rather to balance the ratio of corrective feedback to praise. Previous standards found in the literature recommend five behavioral praise statements for every behavioral corrective feedback, although empirical evidence is limited to support an exact ratio (Flora, 2000; Reddy et al., 2013; Sabey et al., 2018). Recent studies have, however, found evidence suggesting that increasing the ratio is associated with positive behavioral outcomes for students (Cook et al., 2017; Friman et al., 1997; Pisacreta et al., 2011). Future research should focus examining the benefit of teachers using behavioral praise and corrective feedback in sets when managing challenging behaviors. Therefore, it is possible that IVT-T was associated with changes in teachers' use of EBMPs but looking at strategies in isolation from one another may not present the full picture.

One possible explanation for the lack of findings related to dosage predicting EMBPs use is due to the way in which dosage was measured in the current study in comparison to other research. Other studies examining PD implementation have operationalized dosage in ways that

differ from the current study. In particular, quality of training and teachers' engagement with training are two of many constructs that have been used to measure the dosage of training participants received (Holmes et al., 2022; Humphrey et al., 2018). Holmes et al. (2022), conceptualized engagement as participation, openness to, and comprehension of the intervention. Furthermore, Humphrey et al. (2018) theorized that reaching high levels of dosage may reduce the quality of the intervention possibly due to a "race to the finish line" approach. Thus, additional research that defines dosage in terms of engagement or quality of training could shed more light on the relationship between dosage of IVT-T and EBMPs use.

IVT-T is distinct from other PDs in that IVT-T is a self-directed simulated training used to develop teachers' skills in-vivo. Therefore, while identifying dosage benchmarks for IVT-T users will be necessary, it is likely not the only metric needed to provide a complete understanding of successful implementation. In the context of this study, minutes spent using IVT-T may not have fully captured the degree to which teachers were engaged and benefitting from the training. Future research should focus on broader dimensions of implementation in addition to possible interaction effects between implementation dimensions. The current study used a dosage measure that accounted for minutes using IVT-T but did not capture engagement, quality of training sessions, or the individual teachers' understanding of the skills. Therefore, while results of this study suggest that teachers' use of EBMPs increases over time, additional research would provide a more complete account of how dosage operates to impact teacher skills.

Does Usability Impact IVT-T Dosage?

Perceived ease of use of IVT-T was moderately associated with time spent using the system. This finding suggests that teachers who found IVT-T to be easier to interact with spent

more time playing IVT-T, which is consistent with prior research examining usability and the efficacy of technology (Cheung & Vogel, 2013; Ludwick & Doucette, 2009; Shernoff et al., 2021; Shernoff et al., 2022; Venkatesh & Bala, 2008). Cheung and Vogel (2013) studied factors that influenced teachers' engagement with internet-based collaborative learning technologies and found that ease of use predicted time spent using a system and was a stronger predictor than participants' perceptions of the technology's benefit to them (Cheung & Vogel, 2013). Shernoff et al. (2021), found similar results in their feasibility study of IVT-T; teachers' total time playing IVT-T increased as usability of the system increased.

In a study of a behavior management simulation training, researchers examined usability and identified specific factors of the technology that contributed to reduced participant dosage (Badiee & Kaufman, 2015). In this study, qualitative responses from users indicated that the technology's interface, general navigation, and ease of use were the least liked factors of the simulation training (Badiee & Kaufman, 2015). These findings reinforce the notion that the usability of educational technology is related to the amount of time teachers engage with such technology, which is congruent with this study's finding.

McLaughlin et al. (2010) detailed the nuance of designing effective simulated learning environments. Player immersion, or amount of time spent playing, has long been considered the determining factor of a technology's effectiveness (Mclaughlin et al., 2010). However, McLaughlin et al. posit that visual sophistication, or greater usability, is needed to increase player immersion (2010). This theory has been supported by Ludwick and Doucette's 2009 literature review examining medical professionals' experience with health information systems. The review of 86 studies determined usability was a significant factor that impacted medical professionals' use of health information systems and the successful adoption of the technology

(Ludwick & Doucette, 2009). It is a well-established idea in the educational and information technology literature that usability is directly related to whether a system will be adopted or not and is often seen as needing to be addressed before implementing such system (Nielsen & Molich, 1990; Wesson, 2002).

The relationship between usability and time spent playing IVT-T provides important information about the adoption of this technology by teachers and is critical to consider during the development of future PD technology. According to the technology acceptance model theory (Venkatesh et al., 2003), a direct effect exists between time engaged with technology and perceived ease of use. In 2003, Venkatesh and colleagues studied the determinants of intention to use and the usage of new technology across four organizations using eight evidence-based user acceptance models. Results from their study indicated that usability was predictive of acceptance of the new technology ($B = .22$, $p < .01$) and that acceptance of the new technology was predictive of usage of the technology (Venkatesh et al., 2003). A new user acceptance model was formulated and tested based on these findings and included usability – described as effort expectancy – as a determinant of intention to use and usage of new technology (2003). Outcomes of the new model were estimated using the original dataset and cross-validated by two additional organizations. Significant findings included usability being an indirect determinant of usage behaviors with the effect being greater for women and older participants. However, an additional finding indicated that usability only significantly impacted intention during limited exposure to the technology; as experience over time increased the effect decreased (Venkatesh et al., 2003). These findings support the notion that usability is at the crux of technology adoption and usage. The findings also highlight the need to address issues related to usability of new PD technology prior to dissemination, as it becomes a barrier to entry for users.

Defining usability as teachers' perceived ease of use is only one of many ways to examine the usability of a technology. Many components of a technology contribute to its overall usability, with some frameworks focusing on upwards of 20 dimensions (Jahnke et al., 2020; ISO 25010, 2011; Nielsen, 1990: Nokelainen, 2004, 2005). Other usability traits include error prevention and accessibility (technological), feasibility of intervention for promoting learning and motivation to learn (pedagogical), diversity, equity, and inclusion and collaborative learning (socio-culture; Jahnke et al., 2020; Lu et al., 2022). Jahnke et al. (2020), argued for using an interconnected framework of usability across the three dimensions when evaluating usability in order to receive the greatest benefit when identifying potential design flaws. In the current study, usability was assessed using a global measure (SUS), yet research points to the need for more comprehensive measurement. Thus, future studies should measure usability more comprehensively getting at pedagogical and socio-cultural traits in additional technological traits.

The relationship between specific factors that may influence teachers' perceptions of IVT-T's usability, such as demographic information, were not explored in our study. Shernoff et al. (2021) found that technical problems with IVT-T may have contributed to reduced dosage in a qualitative examination of usability. Additionally, demographic factors including age and prior experience with virtual reality gaming, were predictive of perceptions of usability (Shernoff et al., 2021). Given that usability is related to time spent using a new technology, future studies of IVT-T should examine the specific aspects of usability such as (list: age, exposure, culture, etc.,) (Badiee & Kaufman, 2015; Cheung & Vogel, 2013; Lu et al., 2022; Venkatesh et al., 2003). This information will help identify the technological aspects of IVT-T that may need further

refinement to support teachers playing IVT-T for longer (Badiee & Kaufman, 2015; Cheung & Vogel, 2013; Venkatesh et al., 2003).

Limitations

Results of the current study should be interpreted with caution due to several limitations. First, data were limited to what was collected by the grant-funded feasibility study, which consisted of a small sample size and nested data. The small sample size did not afford adequate power to detect small or medium effects among signficiant results. Moreover, the statistical analyses used did not account for the similarities shared and experienced by teachers working within the same district across three schools. In this study, teachers worked together within the same school, which means that they had opportunities to influence each other. Negative binomial regression analyses do not account for the shared characteristics of teachers within the same school or district, which violates the independence assumption (Beaujean & Morgan, 2016). Therefore, this work should be replicated using multilevel modeling with a larger sample size to account for nesting (Beaujean & Morgan, 2016; Cameron & Trivedi, 1998; Cohen & Cohen, 1992).

The method of measurement was another limitation of this study. Behavioral praise and behavioral corrective feedback counts were measured independent of one another and not as a set. However, measuring the ratio to which teachers used behavioral praise to behavioral corrective feedback may provide more accurate information about teachers' use of EBMPs (Reddy et al., 2013). For instance, measuring the ratio of behavioral praise to behavioral corrective feedback could show when teachers increase behavioral praise to balance the amount of praise and corrective feedback they give. Furthermore, teachers are not advised to reduce the amount of behavioral corrective feedback used to zero but rather balance corrective feedback

with praise (Reddy et al., 2013). This study did not examine the ratio of behavioral praise to behavioral corrective feedback. Future research with IVT-T can incorporate a ratio of behavioral praise to behavioral corrective feedback to evaluate teachers' use of EBMPs and to reflect teachers use of strategies more accurately in sets rather than in isolation.

Furthermore, usability of IVT-T was measured using a global rating of usability. Rating scales may be biased due teachers reporting their subjective experience that may be influenced by factors unrelated to the usability of IVT-T, such as how they wish to be perceived. In addition, the global SUS ratings may be too broad and may not provide enough nuanced information about the usability of IVT-T, possessing strong construct validity (Smith, 2005). Future research should examine multiple dimensions of usability including pedagogical and socio-cultural, and use several methods of measurement, such as task-based user testing along with questionnaires to increase the validity of the constructs being measured (Flay et al., 2005).

Finally, it is important to consider that this study was conducted in one high-poverty urban school district, which is not representative of the entire spectrum of national school districts or teachers. This urban district had a high percentage of African American students (94%), and nearly all (97%) of students qualified for free and reduced lunch. Teachers also volunteered on a first come, first served basis to participate in the study with the promise of compensation. Therefore, the motivation and willingness to engage with IVT-T may not be representative of the entire teaching population. Future studies in this area should include a more diverse sample in terms of locale, socioeconomic statuses, and student ethnicity to make findings more generalizable.

Future Research and Implications for Practice

Findings from the current study have implications for the development and implementation of EBMP simulation training as an alternative professional development model for teachers. Infrequent and inconsistent usage of IVT-T highlights barriers to implementing simulation training. This finding suggests the need to identify and promote implementation best practices when developing simulation trainings and education technology. Some barriers to implementation are well-known, such as the usability of the system and time available to teachers; these should be addressed accordingly in future. Specifically, ongoing usability evaluations are necessary to make refinements to the system even after initial assessments as issues of usability appear to persist throughout the development and use of simulation technology. Additionally, school administrators can facilitate the implementation of simulation training by setting aside time for training, communal space for teachers to engage with the system together, and in-person consultation for teachers to better access the concepts being taught. Other barriers to using simulation training may exist, including the age of the player, their experience with video games, and their beliefs about behavioral interventions. These barriers require future research to gain a more complete picture of the impact of EBMP simulation training.

The way in which fidelity and usability are defined and measured is important to understanding teachers' adherence to the training. In the case of fidelity, weekly use may not be the best measure of fidelity to a simulated training as it appears to be highly linked with dosage and provides little additional information about teachers' interaction with the simulator. While fidelity can be measured by the amount of a program delivered, it can also be measured as adherence to the protocol and quality of the program delivered (Proctor et al., 2010). The number of weeks spent using a simulator does not convey information about the user's concentration on

the technology, comprehension of the concepts being taught, or understanding of how to execute the EBMPs being taught. Future research of simulation training should measure additional aspects of fidelity such as focused engagement, concept comprehension checks, and understanding of how to execute the skills.

In addition, the usability of a system influences the amount of time teachers spend interacting with the system. However, many components contribute to the usability of a system and measuring multiple dimensions and traits is necessary to make appropriate adjustments to improve the system's usability. Education technology often superimposes concepts from information technology onto educational tools, such as technological usability, but may not account for distinct concepts of education (Lu et al., 2022). Additional study is needed to include more pedagogical and socio-cultural usability traits as a way to expand the field's understanding of these dimensions of usability within educational technologies.

Conclusions

Simulation training in EBMPs have been developed into a professional development tool to address limitations found in traditional PD. Due to the convenient, low-risk, and cost-effective nature, simulation training is a promising PD tool for teachers to practice managing challenging classroom behaviors, and thereby indirectly reducing associated stress and burnout. However, little is known about how the implementation practices of simulation training in school settings influence teachers' learning and interaction with simulators. Despite the findings from this study being limited by a small sample size and quasi-experimental design, results provide important information that suggest barriers to successfully implementing IVT-T as prescribed exist in the school setting. Furthermore, findings revealed the importance of how implementation practices are operationally defined and measured. Implementation science has long-documented that the

success of school-based programs relies on both the effectiveness of the program and the implementation, which aligns with the results of this study. By researchers expanding their scope of simulation training to include the examination of implementation factors, it will allow teachers to access highly requested skills more readily. School administrators are uniquely positioned to support the integration of simulation training in schools through dedicating time and resources to support teachers engaging with simulators.

References

Aguinis, H., Gottfredson, R. K., & Joo, H. (2013). Best-practice recommendations for defining, identifying, and handling outliers. *Organizational Research Methods*, *16*(2), 270-301.

Alessi, S. (2000). Designing educational support in system-dynamics-based interactive learning environments. *Simulation & Gaming*, *31*(2), 178-196.

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

Annetta, L. A., Cheng, M. T., & Holmes, S. (2010). Assessing twenty-first century skills through a teacher created video game for high school biology students. *Research in Science & Technological Education*, *28*(2), 101-114.

APA Presidential Task Force on Evidence-Based Practice. (2006). Evidence-based practice in psychology. *The American Psychologist*, *61*(4), 271-285.

August, G. J., Lee, S. S., Bloomquist, M. L., Realmuto, G. M., & Hektner, J. M. (2003). Dissemination of an evidence-based prevention innovation for aggressive children living in culturally diverse, urban neighborhoods: The Early Risers effectiveness study. *Prevention Science*, *4*(4), 271-286.

Badiee, F., & Kaufman, D. (2015). Design evaluation of a simulation for teacher education. *Sage Open*, *5*(2), 2158244015592454.

Banerjee, N., Stearns, E., Moller, S., & Mickelson, R. A. (2017). Teacher job satisfaction and student achievement: The roles of teacher professional community and teacher collaboration in schools. *American Journal of Education*, *123*(2), 000-000.

Bates, M. S., Phalen, L., & Moran, C. (2016). Online professional development: A primer. *Phi Delta Kappan*, *97*(5), 70-73.

Beaujean, A. A., & Morgan, G. B. (2016). Tutorial on using regression models with count outcomes using R. *Practical Assessment, Research, and Evaluation*, *21*(1), 2.

Beisiegel, M., Mitchell, R., & Hill, H. C. (2018). The design of video-based professional development: An exploratory experiment intended to identify effective features. *Journal of Teacher Education*, *69*(1), 69-89.

Blizzard Entertainment (2004). World of Warcraft (PC version) [Video game].

Bogler, R., & Nir, A. E. (2012). The importance of teachers' perceived organizational support to job satisfaction. *Journal of educational administration*.

Bowman, D. A., Gabbard, J. L., & Hix, D. (2002). A survey of usability evaluation in virtual environments: classification and comparison of methods. Presence: Teleoperators & Virtual Environments, 11(4), 404–424.

Brooke, J. (1996). SUS-A quick and dirty usability scale. *Usability evaluation in industry*, *189*(194), 4-7.

Browers, A., & Tomic, W. (2000). A longitudinal study of teacher burnout and perceived self-efficacy in classroom management. *Teaching and Teacher Education*, **16**, 239-253.

Caro, P. W. (1973). Aircraft simulators and pilot training. *Human Factors*, *15*(6), 502-509.

Carroll, C., Patterson, M., Wood, S., Booth, A., Rick, J., & Balain, S. (2007). A conceptual framework for implementation fidelity. *Implementation science*, *2*(1), 1-9.

Cauraugh, J. H., Martin, M., & Martin, K. K. (1999). Modeling surgical expertise for motor skill acquisition. *The American journal of surgery*, *177*(4), 331-336.

Cheung, R., & Vogel, D. (2013). Predicting user acceptance of collaborative technologies: An extension of the technology acceptance model for e-learning. *Computers & education*, *63*, 160-175.

Chin, J. P., Diehl, V. A., & Norman, K. L. (1988, May). Development of an instrument measuring user satisfaction of the human-computer interface. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 213-218).

Chin, J. P., Norman, K. L., & Shneiderman, B. (1987). *Subjective user evaluation of CF PASCAL programming tools*. University of Maryland. Computer Science.

Chung, H., & Kim, H. (2010). Implementing professional standards in teacher preparation programs in the United States: Preservice teachers' understanding of teaching standards. *KEDI Journal of Educational Policy, 7*(2).

Clotfelter, C., Ladd, H. F., Vigdor, J., & Wheeler, J. (2006). High-poverty schools and the distribution of teachers and principals. *NCL Rev., 85*, 1345.

Collie, R. J., Shapka, J. D., & Perry, N. E. (2012). School climate and social–emotional learning: Predicting teacher stress, job satisfaction, and teaching efficacy. *Journal of educational psychology*, *104*(4), 1189.

Cook, D. A., Hatala, R., Brydges, R., Zendejas, B., Szostek, J. H., Wang, A. T., ... & Hamstra, S. J. (2011). Technology-enhanced simulation for health professions education: a systematic review and meta-analysis. *Jama*, *306*(9), 978-988.

Cook, M., Dickerson, D. L., Annetta, L. A., & Minogue, J. (2011). In-service teachers' perceptions of online learning environments. *Quarterly Review of Distance Education*, *12*(2).

Dane, A. V., & Schneider, B. H. (1998). Program integrity in primary and early secondary prevention: are implementation effects out of control? *Clinical psychology review, 18*(1), 23-45.

Darling-Hammond, L., Wei, R. C., & Andree, A. (2010). *How high-achieving countries develop great teachers*. Stanford Center for Opportunity Policy in Education.

Dawson, M. R., & Lignugaris/Kraft, B. (2017). Meaningful practice: Generalizing foundation teaching skills from TLE TeachLivE™ to the classroom. *Teacher Education and Special Education*, *40*(1), 26-50.

Dede, C. (2005). Planning for neomillennial learning styles. *Educause Quarterly*, *28*(1), 7-12.

Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational researcher*, *38*(3), 181-199.

Dieker, L. A., Hynes, M., Hughes, C., & Smith, E. (2008). Implications of mixed reality simulation technologies on special education and teacher preparation. Focus on Exceptional Children, 40(6), 1-20.

Dieker, L. A., Rodriguez, J. A., Lignugaris/Kraft, B., Hynes, M. C., & Hughes, C. E. (2014). The potential of simulated environments in teacher education: Current and future possibilities. *Teacher Education and Special Education*, *37*(1), 21-33.

Dieker, L., Hynes, M., Stapleton, C., & Hughes, C. (2007). Virtual classrooms: STAR simulator.

Dorsey, D., Campbell, G., & Russell, S. (2009). Adopting the instructional science paradigm to encompass training in virtual environments. *Theoretical Issues in Ergonomics Science*, *10*(3), 197-215.

Durlak, J. A., & DuPre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American journal of community psychology*, *41*(3), 327-350.

Egert, F., Fukkink, R. G., & Eckhardt, A. G. (2018). Impact of in-service professional development programs for early childhood teachers on quality ratings and child outcomes: A meta-analysis. *Review of Educational Research*, *88*(3), 401-433.

Elliott, J. C. (2017). The evolution from traditional to online professional development: A review. *Journal of Digital Learning in Teacher Education*, *33*(3), 114-125.

Ersozlu, Z., Ledger, S., Ersozlu, A., Mayne, F., & Wildy, H. (2021). Mixed-Reality Learning Environments in Teacher Education: An Analysis of TeachLivE™ Research. *Sage Open*, *11*(3), 21582440211032155.

Fabry, D. L., & Higgs, J. R. (1997). Barriers to the effective use of technology in education: Current status. *Journal of educational computing research*, *17*(4), 385-395.

Federal Aviation Administration [FAA] website, Order 8900.1, link https://www.faa.gov/pilots/training/part_142/train_prog_curriclums/

Fixsen, D. L., Naoom, S. F., Blase, K. A., Friedman, R. M., Wallace, F., Burns, B., ... & Shern, D. (2005). Implementation research: A synthesis of the literature.

Girard, C., Ecalle, J., & Magnan, A. (2012). Serious games as new educational tools: how effective are they? A meta-analysis of recent studies. Journal of Computer Assisted Learning, 29(3), 207-219.

Green, P., & MacLeod, C. J. (2016). SIMR: an R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution*, *7*(4), 493-498.

Greenberg, J., Putman, H., & Walsh, K. (2014). Training Our Future Teachers: Classroom Management. Revised. *National Council on Teacher Quality*.

Greenberg, M. T., & Kusché, C. A. (2006). Building social and emotional competence: The PATHS curriculum.

Gregory, S., & James, R. (2011). VirtualPREX: Open and Distance Learning for pre-service teachers. In *Expanding Horizons-New Approaches to Open and Distance Learning. Presented at the 24th ICDE World Conference on Open & Distance Learning, Bali*.

Guskey, T. R. (2000). *Evaluating professional development*. Corwin press.

Halim, M. S. A. A., & Hashim, H. (2019). Integrating web 2.0 technology in ESL classroom: A review on the benefits and barriers. *Journal of Counseling and Educational Technology*, *2*, 1-8.

Hall, G. E. (2010). Technology's Achilles heel: Achieving high-quality implementation. *Journal of Research on Technology in education*, *42*(3), 231-253.

Harper, B. D., & Norman, K. L. (1993, February). Improving user satisfaction: The questionnaire for user interaction satisfaction version 5.5. In *Proceedings of the 1st Annual Mid-Atlantic Human Factors Conference* (pp. 224-228). sn.

Heppell, J., Beauchamp, G., & Chollet, A. (1995). Ten-year experience with a basic technical skills and perioperative management workshop for first-year residents. *Canadian Journal of Surgery*, *38*(1), 27.

Humphrey, N., Barlow, A., & Lendrum, A. (2018). Quality matters: Implementation moderates student outcomes in the PATHS curriculum. *Prevention Science*, *19*(2), 197-208.

Jentsch, F., & Curtis, M. (2017). *Simulation in aviation training*. Routledge.

Joyce, B. R., & Showers, B. (2002). Student achievement through staff development.

Kelley, M. J., & Wenzel, T. (2019). How TeachLivE™ transformed our teaching practices in reading education and pre-service. *SRATE Journal*, *28*(1), 9-22.

Kessler, R. C., Avenevoli, S., Costello, E. J., Georgiades, K., Green, J. G., Gruber, M. J., ... & Sampson, N. A. (2012). Prevalence, persistence, and sociodemographic correlates of

DSM-IV disorders in the National Comorbidity Survey Replication Adolescent Supplement. *Archives of general psychiatry*, *69*(4), 372-380.

Kettler, R. J., Reddy, L. A., Glover, T. A., & Kurz, A. (2019). Bridging the Gap: Classroom Strategies Assessment System–Observer Form. *Assessment for Effective Intervention*, *44*(2), 120-122.

Klaft, J. M., & Codding, R. S. (2022). Promoting teachers' implementation adherence and quality of the good behavior game using behavioral skills training. *Journal of Educational and Psychological Consultation*, *32*(2), 156-184.

Kokkinos, C. M. (2007). Job stressors, personality and burnout in primary school teachers. *British journal of educational psychology*, *77*(1), 229-243.

Larson, K. E., Hirsch, S. E., McGraw, J. P., & Bradshaw, C. P. (2020). Preparing preservice teachers to manage behavior problems in the classroom: The feasibility and acceptability of using a mixed-reality simulator. *Journal of Special Education Technology*, *35*(2), 63-75.

Latouche, A. P., & Gascoigne, M. (2019). In-service training for increasing teachers' ADHD knowledge and self-efficacy. *Journal of attention disorders*, *23*(3), 270-281.

Lawler III, E. E. (1973). Motivation in work organizations.

Lawless, K. A., & Pellegrino, J. W. (2007). Professional development in integrating technology into teaching and learning: Knowns, unknowns, and ways to pursue better questions and answers. *Review of educational research*, *77*(4), 575-614.

Lewis, J. R. (2018). The system usability scale: past, present, and future. *International Journal of Human–Computer Interaction*, *34*(7), 577-590.

Lintern, G., Roscoe, S. N., Koonce, J. M., & Segal, L. D. (1990). Transfer of landing skills in beginning flight training. *Human Factors*, *32*(3), 319-327.

Little, R. J. (1988). Missing-data adjustments in large surveys. *Journal of Business & Economic Statistics*, *6*(3), 287-296.

Locke, E. A. (1969). What is job satisfaction? *Organizational behavior and human performance*, *4*(4), 309-336.

Ludwick, D. A., & Doucette, J. (2009). Adopting electronic medical records in primary care: lessons learned from health information systems implementation experience in seven countries. *International journal of medical informatics*, *78*(1), 22-31.

Lutrick, E., & Szabo, S. (2012). Instructional leaders' beliefs about effective professional development. *Delta Kappa Gamma Bulletin*, *78*(3), 6–12.

Melnick, S. A., Meister, D. G. (2008). A comparison of beginning and experienced teachers' concerns. Educational Research Quarterly, 31(3), 40-56.

Mihalic, S. (2004). The importance of implementation fidelity. *Emotional and Behavioral Disorders in Youth*, *4*(4), 83-105.

Mitchell, B. S., Hirn, R. G., & Lewis, T. J. (2017). Enhancing effective classroom management in schools: Structures for changing teacher behavior. *Teacher Education and Special Education*, *40*(2), 140-153.

Mixon, C. S., Owens, J. S., Hustus, C., Serrano, V. J., & Holdaway, A. S. (2019). Evaluating the impact of online professional development on teachers' use of a targeted behavioral classroom intervention. *School Mental Health*, *11*(1), 115-128.

National Staff Development Council. (2011). *Standards for professional learning.* Retrieved from http://www.learningforward.org

Novins, D. K., Green, A. E., Legha, R. K., & Aarons, G. A. (2013). Dissemination and implementation of evidence-based practices for child and adolescent mental health: A systematic review. *Journal of the American Academy of Child & Adolescent Psychiatry*, *52*(10), 1009-1025.

Osborne, J. W., & Overbay, A. (2004). The power of outliers (and why researchers should always check for them). *Practical Assessment, Research, and Evaluation*, *9*(1), 6.

Owens, J. S., Holdaway, A. S., Smith, J., Evans, S. W., Himawan, L. K., Coles, E. K., ... & Dawson, A. E. (2018). Rates of common classroom behavior management strategies and their associations with challenging student behavior in elementary school. *Journal of Emotional and Behavioral Disorders*, *26*(3), 156-169.

Owens, J. S., McLennan, J. D., Hustus, C. L., Haines-Saah, R., Mitchell, S., Mixon, C. S., & Troutman, A. (2019). Leveraging technology to facilitate teachers' use of a targeted classroom intervention: Evaluation of the Daily Report Card. Online (DRC. O) System. *School Mental Health*, *11*(4), 665-677.

Ozaki, M., Downes, M. S., Baba, A., Yanagita, K., & Kashiwamura, H. (2013). Development of a virtual reality force feedback-enabled dental drill training simulation. *International Journal of Biomedical Soft Computing and Human Sciences: the official journal of the Biomedical Fuzzy Systems Association*, *18*(2), 49-55.

Parsons, S. A., Hutchison, A. C., Hall, L. A., Parsons, A. W., Ives, S. T., & Leggett, A. B. (2019). US teachers' perceptions of online professional development. *Teaching and Teacher Education: An International Journal of Research and Studies*, *82*(1), 33-42.

Pas, E. T., Bradshaw, C. P., Hershfeldt, P. A., & Leaf, P. J. (2010). A multilevel exploration of the influence of teacher efficacy and burnout on response to student problem behavior and school-based service use. *School Psychology Quarterly*, *25*(1), 13.

Pianta, R. C., DeCoster, J., Cabell, S., Burchinal, M., Hamre, B. K., Downer, J., ... & Howes, C. (2014). Dose–response relations between preschool teachers' exposure to components of professional development and increases in quality of their interactions with children. *Early Childhood Research Quarterly*, *29*(4), 499-508.

Polanczyk, G. V., Salum, G. A., Sugaya, L. S., Caye, A., & Rohde, L. A. (2015). Annual research review: A meta-analysis of the worldwide prevalence of mental disorders in children and adolescents. *Journal of Child Psychology and Psychiatry*, *56*(3), 345-365.

Powell, B. J., Patel, S. V., Haley, A. D., Haines, E. R., Knocke, K. E., Chandler, S., ... & Aarons, G. A. (2020). Determinants of implementing evidence-based trauma-focused interventions for children and youth: A systematic review. *Administration and Policy in Mental Health and Mental Health Services Research*, *47*(5), 705-719.

Poznanski, B., Hart, K. C., & Cramer, E. (2018). Are teachers ready? Preservice teacher knowledge of classroom management and ADHD. *School Mental Health*, *10*(3), 301-313.

Proctor, E. K., Landsverk, J., Aarons, G., Chambers, D., Glisson, C., & Mittman, B. (2009). Implementation research in mental health services: An emerging science with conceptual, methodological, and training challenges. Administration and Policy in Mental Health and Mental Health Services Research, 36(1), 24–34.

Proctor, E., Silmere, H., Raghavan, R., Hovmand, P., Aarons, G., Bunger, A., ... & Hensley, M. (2011). Outcomes for implementation research: conceptual distinctions, measurement

challenges, and research agenda. *Administration and policy in mental health and mental health services research*, *38*(2), 65-76.

Real, K., & Poole, M. S. (2005). Innovation implementation: Conceptualization and measurement in organizational research. In *Research in organizational change and development*. Emerald Group Publishing Limited.

Reddy, L. A., & Dudek, C. M. (2014). Teacher progress monitoring of instructional and behavioral management practices: An evidence-based approach to improving classroom practices. *International Journal of School & Educational Psychology*, *2*(2), 71-84.

Reddy, L.A., Fabiano, G. A., Dudek, C. M. (2013). Concurrent validity of the Classroom Strategies Scale for elementary school – Observer form. *Journal of Psychoeducational Assessment*, *31*, 258-270.

Reinke, W. M., Stormont, M., Herman, K. C., Puri, R., & Goel, N. (2011). Supporting children's mental health in schools: Teacher perceptions of needs, roles, and barriers. *School Psychology Quarterly*, *26*(1), 1.

Reyes, M. R., Brackett, M. A., Rivers, S. E., White, M., & Salovey, P. (2012). Classroom emotional climate, student engagement, and academic achievement. *Journal of educational psychology*, *104*(3), 700.

Reznick, R. K., & MacRae, H. (2006). Teaching surgical skills—changes in the wind. *New England Journal of Medicine*, *355*(25), 2664-2669.

Rigby, K., & Griffiths, C. (2010). Applying the method of shared concern in Australian schools: An evaluative study. Retrieved from http://www.deewr.gov.au/Schooling/NationalSafeSchools/Documents/covertBullyReports/MethodOFSharedConcern.pdf:

Robbins, M. S., Feaster, D. J., Horigian, V. E., Puccinelli, M. J., Henderson, C., & Szapocznik, J. (2011). Therapist adherence in brief strategic family therapy for adolescent drug abusers. *Journal of consulting and clinical psychology*, *79*(1), 43.

Sauro, J. (2010). A practical guide to measuring usability. *Measuring Usability LLC, Denver*.

Schneiderman, B. (1992). Developing the User Interface. Reading, MA: Addison-Wesley

Schoenwald, S. K., & Hoagwood, K. (2001). Effectiveness, transportability, and dissemination of interventions: What matters when? *Psychiatric services*, *52*(9), 1190-1197.

Schoenwald, S. K., Carter, R. E., Chapman, J. E., & Sheidow, A. J. (2008). Therapist adherence and organizational effects on change in youth behavior problems one year after multisystemic therapy. *Administration and policy in mental health and mental health services research*, *35*(5), 379-394.

Schoenwald, S. K., Sheidow, A. J., & Chapman, J. E. (2009). Clinical supervision in treatment transport: effects on adherence and outcomes. *Journal of Consulting and Clinical Psychology*, *77*(3), 410.

Schrader, P. G., & McCreery, M. (2008). The acquisition of skill and expertise in massively multiplayer online games. *Educational Technology Research and Development*, *56*(5), 557-574.

Schrader, P. G., Archambault, L. M., & Oh-Young, C. (2011). Training by gaming: Preparing teachers of today for tomorrow's learning environments. Journal of Technology and Teacher Education, 19(3), 261-286.

Schussler, D., Frank, J., Lee, T., & Mahfouz, J. (2017). Using virtual role-play to enhance teacher candidates' skills in responding to bullying. Journal of Technology and Teacher Education, 25(1), 91-120.

Seymour, N. E., Gallagher, A. G., Roman, S. A., O'brien, M. K., Bansal, V. K., Andersen, D. K., & Satava, R. M. (2002). Virtual reality training improves operating room performance: results of a randomized, double-blinded study. *Annals of surgery*, *236*(4), 458.

Sharples, M., Taylor, J., & Vavoula, G. (2010). A theory of learning for the mobile age. In *Medienbildung in neuen Kulturräumen* (pp. 87-99). VS Verlag für Sozialwissenschaften.

Shernoff, E. S., Frazier, S. L., Lisetti, C., Delmarre, A., Bibi, Z., & Gabbard, J. (2021). Supporting the Implementation of Evidence-Based Behavior Management Practices through Simulation: A Mixed Method Study. *Journal of Educational and Psychological Consultation, 31*(4), 463-497.

Shernoff, E. S., Lekwa, A. L., Delmarre, A., Gabbard, J., Stokes-Tyler, D., Lisetti, C., & Frazier, S. L. (2022). Bridging simulation technology with positive behavioral supports to promote student engagement and behavior. *Journal of School Psychology*, *95*, 121-138.

Shernoff, E.S., Frazier, S.L., Lisetti, C., Buche, C., Lunn, S., Brown, C., Delamarre, A., Chou, T., Gabbard, J. & Morgan, E. (2018). Early career teacher professional development: Bridging simulation technology with evidence-based behavior management. Journal of Technology and Teacher Education.

Shinn, M. R., Ramsey, E., Walker, H. M., Stieber, S., & O'Neill, R. E. (1987). Antisocial behavior in school settings: Initial differences in an at-risk and normal population. *The Journal of Special Education*, **21**, 69-84.

Simon, N. S., & Johnson, S. M. (2015). Teacher turnover in high-poverty schools: What we know and can do. *Teachers College Record*, *117*(3), 1-36.

Sitzmann, T. (2011). A meta-analytic examination of the instructional effectiveness of computer-based simulation games. Personnel Psychology, 64(2), 489-528.

Sørlie, M. A., Ogden, T., & Olseth, A. R. (2015). Preventing problem behavior in school through school-wide staff empowerment: Intervention outcomes. *World Journal of Educational Research*, *2*.

Straub, C., Dieker, L., Hynes, M., & Hughes, C. (2014). Using virtual rehearsal in TLE TeachLivE™ mixed reality classroom simulator to determine the effects on the performance of mathematics teachers (2014 TeachLivE National Research Project: Year 1 Findings). Orlando: University of Central Florida

Trahan, M. H., Smith, K. S., & Talbot, T. B. (2019). Past, present, and future: Editorial on virtual reality applications to human services. *Journal of Technology in Human Services*, *37*(1), 1-12.

Venkatesh, V., & Bala, H. (2008). Technology acceptance model 3 and a research agenda on interventions. *Decision sciences*, *39*(2), 273-315.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS quarterly*, 425-478.

Vince Garland, K. M., Holden, K., & Garland, D. P. (2016). Individualized clinical coaching in the TLE TeachLivE lab: Enhancing fidelity of implementation of system of least prompts among novice teachers of students with autism. *Teacher Education and Special Education*, *39*(1), 47-59.

Vince Garland, K. M., Holden, K., & Garland, D. P. (2016). Individualized clinical coaching in the TLE TeachLivE lab: Enhancing fidelity of implementation of system of least prompts

among novice teachers of students with autism. *Teacher Education and Special Education*, *39*(1), 47-59.

Vince Garland, K. M., Vasquez, E., & Pearl, C. (2012). Efficacy of individualized clinical coaching in a virtual reality classroom for increasing teachers' fidelity of implementation of discrete trial teaching. Education and Training in Autism and Developmental Disabilities, 47, 502-515

Vince Garland, K., & Garland, D. (2020). TeachLivE™ and Teach Well: Simulations in Teacher Education. In *Games and Simulations in Teacher Education* (pp. 183-195). Springer, Cham.

Vygotsky, L. (1978). Interaction between learning and development. *Readings on the development of children*, *23*(3), 34-41.

Wei, R. C., Darling-Hammond, L., & Adamson, F. (2010). *Professional development in the United States: Trends and challenges* (Vol. 28). Dallas, TX: National Staff Development Council.

White, G. C., & Bennetts, R. E. (1996). Analysis of frequency count data using the negative binomial distribution. *Ecology*, *77*(8), 2549-2557.

Zar, J. H. (2005). Spearman rank correlation. *Encyclopedia of biostatistics*, *7*.

Zhai, F., Raver, C. C., Jones, S. M., Li-Grining, C. P., Pressler, E., & Gao, Q. (2010). Dosage effects on school readiness: Evidence from a randomized classroom-based intervention. *Social Service Review*, *84*(4), 615-655.

Table 1

Teacher Characteristics

Variable	*%*	*M*	*SD*	*Range*
*Gender				
Female	75.00			
Male	22.90			
Age		44.06	12.12	24.00 – 64.00
Ethnicity				
African American	43.80			
Hispanic/Latino(a)	8.30			
European American	16.70			
Asian American	2.10			
American Indian/Alaskan Native	2.10			
Other	8.30			
Total years teaching		14.53	9.39	1.00 – 36.00
Highest degree earned				
Associates	2.10			
BA/BS	25.00			
MA/MS/Ed.M.	45.80			
Ph.D./Ed.D.	6.30			

Note. $N = 48$.

Table 2

Descriptive Statistics for Independent and Dependent Variables for Full Dataset

Variable	*M*	*SD*	*Range*
Total IVT-T mins	195.7	206.2	6.0 – 952.0
IVT-T weeks	4.6	5.2	1.0 – 21.0
System Usability Scale	65.5	17.3	30.0 – 97.5
BP			
Time 1	4.2	4.7	0.0 – 17.0
Time 2	8.8	9.9	0.0 – 40.0
BCF			
Time 1	29.3	20.5	2.0 – 71.0
Time 2	20.2	12.1	4.0 – 51.0

Note: BP= Behavioral Praise. BCF= Behavioral Corrective Feedback.

Table 3

Descriptive Statistics for RQ3

Variable	*M*	*SD*	*Range*
Total IVT-T mins	267.38	308.55	6.00 – 1289.00
SUS	65.51	17.32	30.00 – 97.50

Note. N = 39.

Table 4

Intercorrelations for Study Variables

		n	1	2	3	4	5	6
1.	Total IVT-T mins	46	—					
2.	IVT-T weeks	46	.95**	—				
3.	BP Time 1	38	.15	.02	—			
4.	BCF Time 1	38	-.30	-.06	.09	—		
5.	BP Time 2	34	.40*	.53**	.11	-.33	—	
6.	BCF Time 2	34	-.20	-.14	.14	.47**	.07	—

Note. SUS= System Usability Scale. BP= Behavioral Praise. BCF= Behavioral Corrective Feedback.

*$p < .05$, two tailed. **$p < .01$, two tailed.

Table 5

Number of Teachers Playing IVT-T Each Week

Total weeks played	# Of teachers
0	0
1	14
2	9
3	5
4	3
5	6
6	2
7	3
8	0
9	1
10	1
11	0
12	0
13	0
14	0
15	0
16	0
17	0
18	1
19	0
20	1
21	2
22	0
23	0

Note. $N = 48$

Table 6

Regression Analysis for Dosage as a Predictor of CSAS Strategy Counts

Model		*b*	*SE*	*p*-value
BP				
	Total IVT-T minutes	0.002	0.001	.061
	BP Time 1	0.045	0.048	.351
BCF				
	Total IVT-T minutes	-0.001	0.001	.770
	BCF Time 1	0.015	0.005	.002

Note. N=30. BP= Behavioral Praise. BCF= Behavioral Corrective Feedback.

Table 7

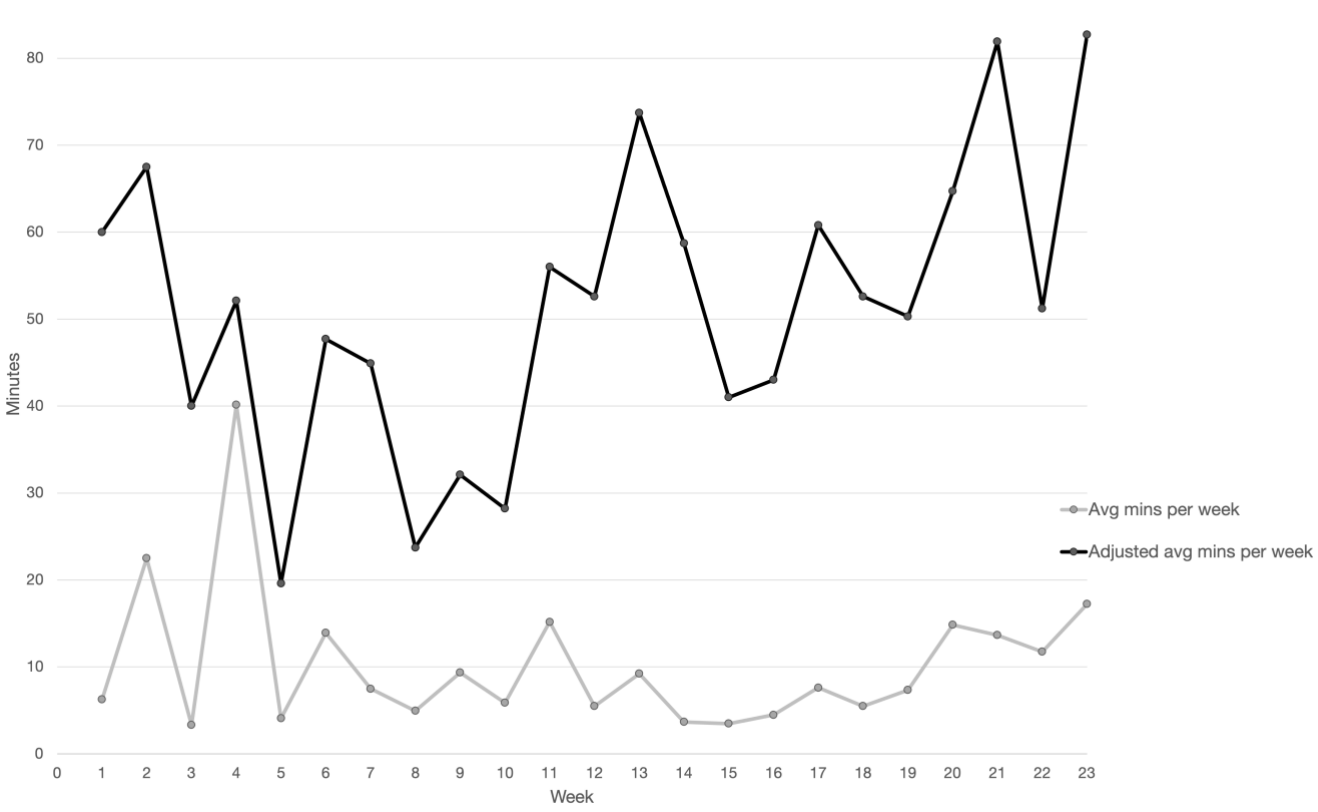
Correlation Analysis of Usability and IVT-T Mins

		1	2
1.	SUS	--	.403*
2.	IVT-T mins		--

Note. $N = 39$. Spearman's rho.
* $p < .05$. ** $p < .01$.

Figure 1

Average Time Teachers Played IVT-T per Week

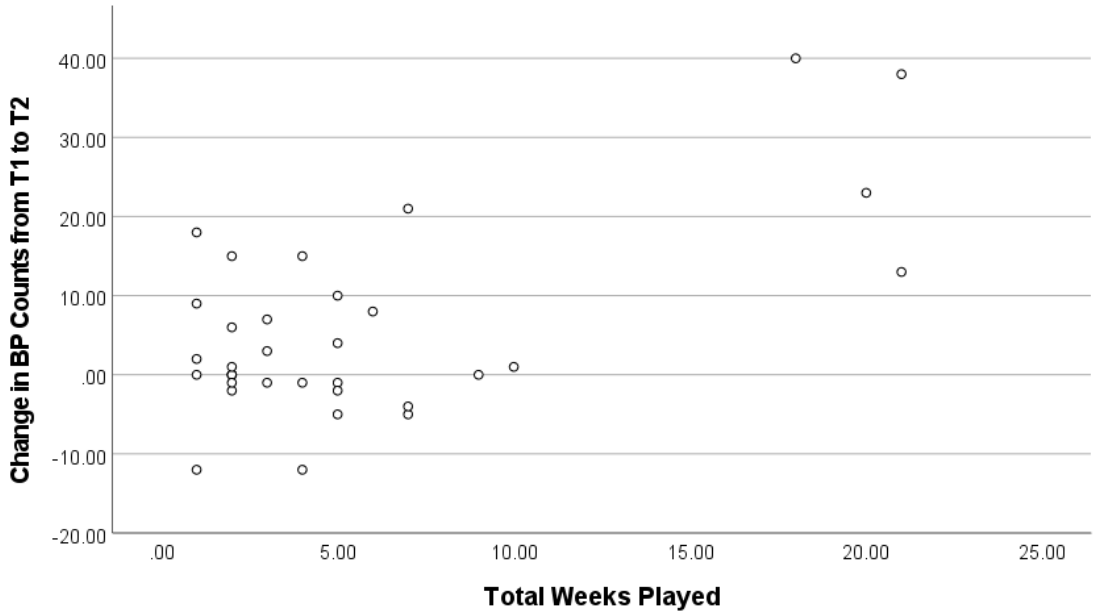


Note. Avg mins per week = average minutes all teachers played IVT-T (N = 48), SD = 8.3 minutes. Adjusted avg mins per week = average minutes teachers played IVT-T per week of those that played IVT-T, SD = 16.7 minutes.

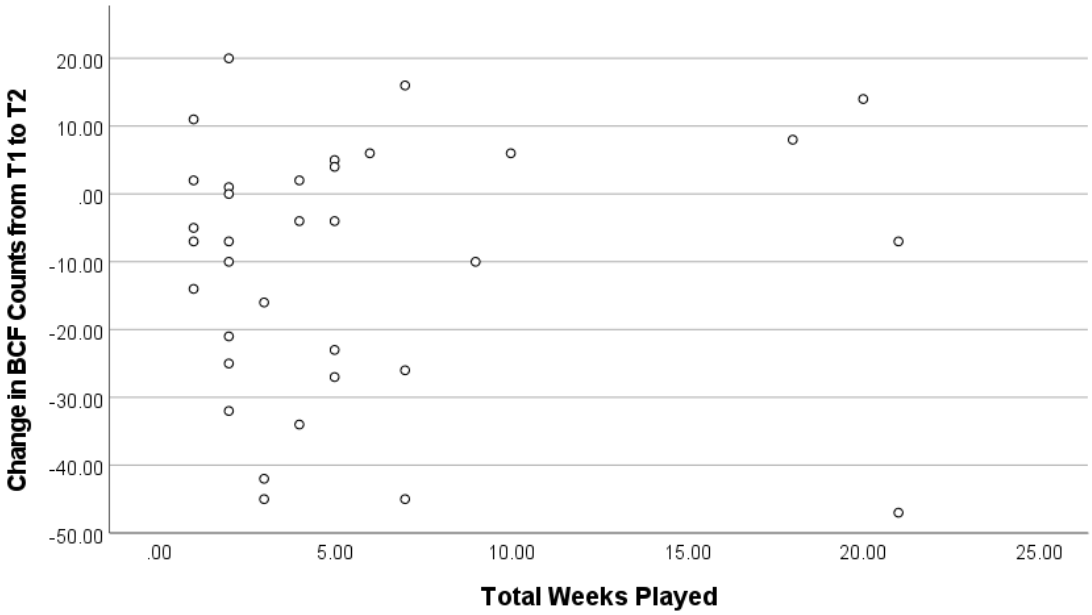
Figure 2

Relationship Between Total Weeks Played and Change in CSAS Strategy Counts

A



B



Note. Panel A: Relationship between change in behavioral praise and total weeks played. Panel B: Relationship between change in behavioral corrective feedback and total weeks played.