FROM TEACHLIVE™ TO THE CLASSROOM: BUILDING PRESERVICE
SPECIAL EDUCATORS’ PROFICIENCY WITH ESSENTIAL
TEACHING SKILLS

by

Melanie Rees Dawson

A dissertation submitted in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Disability Disciplines

Approved:

Benjamin Lignugaris/Kraft, Ph.D.
Major Professor

Charles L. Salzberg, Ph.D.
Committee Member

Timothy A. Slocum, Ph.D.
Committee Member

Marilyn Likins, Ph.D.
Committee Member

Sylvia Read, Ph.D.
Committee Member

Mark R. McLellan, Ph.D.
Vice President for Research and
Dean of the School of Graduate Studies

UTAH STATE UNIVERSITY
Logan, Utah

2016
ABSTRACT

From TeachLivE™ to the Classroom: Building Preservice Special Educators’ Proficiency with Essential Teaching Skills

by

Melanie Rees Dawson, Doctor of Philosophy

Utah State University, 2016

Preservice special education teachers need to develop essential teaching skills to competently address student academics and behavior in the classroom. TeachLivE™ is a sophisticated virtual simulation that has recently emerged in teacher preparation programs to supplement traditional didactic instruction and field experiences. Teacher educators can engineer scenarios in TeachLivE™ to cumulatively build in complexity, allowing preservice teachers to incrementally interleave target skills in increasingly difficult situations.

The purpose of this study was to investigate the effectiveness of TeachLivE™ on preservice special education teachers’ delivery of error correction, specific praise, and praise around in the virtual environment and in authentic classroom settings. Four preservice special educators who were teaching on provisional licenses in upper elementary language arts classrooms participated in this multiple baseline study across
target skills. Participants attended weekly TeachLivE™ sessions as a group, where they engaged in three short teaching turns followed by structured feedback. Participants’ proficiency with the target skills was analyzed on three weekly assessments. First, participants’ mastery of current and previous target skills was measured during their third teaching turn of the intervention session (i.e., TeachLivE™ training assessment). Next, participants’ proficiency with all skills, including those that had not been targeted yet in intervention, were measured immediately following intervention sessions (i.e., TeachLivE™ comprehensive assessment). Finally, teachers submitted a weekly video recording of a lesson in their real classroom (i.e. classroom generalization assessment).

Repeated practice and feedback in TeachLivE™ promoted participants’ mastery of essential target skills. Specifically, all participants demonstrated proficiency with error correction, specific praise, and praise around on both the TeachLivE™ training assessment and the more complex TeachLivE™ comprehensive assessment, with a strong pattern of generalized performance to authentic classroom settings. Participants maintained proficiency with the majority of the target skills in both environments when assessed approximately one month after intervention was discontinued. Implications of the study are discussed, including the power of interleaved practice in TeachLivE™ and how generalization and maintenance may be impacted by the degree of alignment between virtual and real teaching scenarios.

(241 pages)
PUBLIC ABSTRACT

From TeachLivE™ to the Classroom: Building Preservice Special Educators’ Proficiency with Essential Teaching Skills

Melanie Rees Dawson

Preservice special education teachers need to develop essential teaching skills to competently address student academics and behavior in the classroom. TeachLivE™ is a sophisticated virtual simulation that has recently emerged in teacher preparation programs to supplement traditional didactic instruction and field experiences. Teacher educators can engineer scenarios in TeachLivE™ to cumulatively build in complexity, allowing preservice teachers to incrementally interleave target skills in increasingly difficult situations.

The purpose of this study was to investigate the effectiveness of TeachLivE™ on preservice special education teachers’ delivery of error correction, specific praise, and praise around in the virtual environment and in authentic classroom settings. Four preservice special educators who were teaching on provisional licenses in upper elementary language arts classrooms participated in this multiple baseline study across target skills. Participants attended weekly TeachLivE™ sessions as a group, where they engaged in three short teaching turns followed by structured feedback. Participants’ proficiency with the target skills was analyzed on three weekly assessments. First, participants’ mastery of current and previous target skills was measured during their third teaching turn of the intervention session (i.e., TeachLivE™ training assessment). Next,
participants’ proficiency with all skills, including those that had not been targeted yet in intervention, were measured immediately following intervention sessions (i.e., TeachLivE™ comprehensive assessment). Finally, teachers submitted a weekly video recording of a lesson in their real classroom (i.e. classroom generalization assessment).

Repeated practice and feedback in TeachLivE™ promoted participants’ mastery of essential target skills. Specifically, all participants demonstrated proficiency with error correction, specific praise, and praise around on both the TeachLivE™ training assessment and the more complex TeachLivE™ comprehensive assessment, with a strong pattern of generalized performance to authentic classroom settings. Participants maintained proficiency with the majority of the target skills in both environments when assessed approximately one month after intervention was discontinued. Implications of the study are discussed, including the power of interleaved practice in TeachLivE™ and how generalization and maintenance may be impacted by the degree of alignment between virtual and real teaching scenarios.
ACKNOWLEDGMENTS

I would like to thank a number of individuals who have been influential in my graduate school journey. First, I extend appreciation to those who served on my dissertation committee. To my major professor, Benjamin Lignugaris/Kraft, thank you for encouraging me to pursue my doctoral degree and advising me throughout the process. I am grateful you recognized my potential and invested so much time and energy in me during my undergraduate and doctoral studies. A special thank you to Marilyn Likins for encouraging me to find balance between my professional and personal endeavors. I appreciate your mentorship and friendship. I am also grateful to Sylvia Read, Charles Salzburg, and Timothy Slocum for your dedication to my dissertation and your role in shaping my writing and research abilities.

I feel honored to have met such high-caliber colleagues, researchers, and educators while pursuing my graduate degree. Thank you to Marie Orton, my research assistant, for your competence and willingness to devote so much time to this project. I am grateful you endured two pilot studies and the dissertation. Also, I gratefully acknowledge the teachers who participated in this study. Thank you for your dedication to the research process and your desire to improve your teaching for the benefit of your students. You each have a promising career ahead.

Most of all, I would like to thank my family for their part in this milestone. To my mother, Grayce Rees, and my late father, Dean Rees, thank you for instilling in me the importance of education and a desire for continuous self-improvement. To my parents-in-law, Tom and Deanna Dawson, thank you for accepting me as one of your own and
supporting my endeavors in higher education. I especially extend love and gratitude to my husband, Kevin, for always encouraging me to become the best possible version of myself. It’s been an adventurous five years of marriage and graduate school, and I look forward to writing many more chapters of our life together. Finally, I am grateful for my children, both of whom were born during my doctoral program. To Genevieve, whose arrival changed the timing of my dissertation, and to William Charles, whose birth prevented me from attending graduation, thank you for teaching me to be flexible and for bringing renewed purpose to my work. I am blessed beyond measure.

Melanie Rees Dawson
# CONTENTS

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT .......................................................... iii</td>
</tr>
<tr>
<td>PUBLIC ABSTRACT .................................................. v</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS ................................................... vii</td>
</tr>
<tr>
<td>LIST OF TABLES ................................................... xi</td>
</tr>
<tr>
<td>LIST OF FIGURES ................................................... xii</td>
</tr>
</tbody>
</table>

## CHAPTER

I. INTRODUCTION ........................................................................... 1

- Field Placements ................................................................. 2
- Situated Learning Methods ....................................................... 3
- Purpose and Research Questions ............................................... 8

II. LITERATURE REVIEW ........................................................... 10

- Essential Teaching Skills .......................................................... 10
- Virtual Simulations and Teacher Preparation ............................. 20
- Conclusions ............................................................................. 48

III. METHODS ............................................................................ 52

- Participants ............................................................................. 52
- Settings ................................................................................... 54
- Lesson Materials ...................................................................... 57
- Dependent Variables and Measurement System ...................... 58
- Independent Variable ............................................................... 69
- Experimental Design and Procedures ......................................... 74
- Reliability ................................................................................. 82
- Treatment Fidelity ..................................................................... 88

IV. RESULTS ................................................................................ 90

- Opportunity to Respond Results ............................................. 90
- Error Correction, Specific Praise, and Praise Around Results .... 93
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Virtual Simulation Participant and Setting Characteristics ................................ 27</td>
</tr>
<tr>
<td>2. Virtual Simulation Independent Variable Characteristics ......................................... 30</td>
</tr>
<tr>
<td>4. Virtual Simulation Methodological Quality Indicator Scores ........................................ 46</td>
</tr>
<tr>
<td>5. Participant Background Information ............................................................................ 54</td>
</tr>
<tr>
<td>6. Participant Classroom Generalization Setting Information ........................................... 56</td>
</tr>
<tr>
<td>7. Characteristics of Each Weekly Assessment ................................................................. 62</td>
</tr>
<tr>
<td>8. Mean IOA Scores and Ranges for Target Skills in each Assessment Setting .................. 83</td>
</tr>
<tr>
<td>9. Participants’ OTR Rates per Min in Each Assessment Setting ....................................... 91</td>
</tr>
<tr>
<td>10. Results of Social Validity Items addressing Realism of TeachLivE .............................. 115</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>A visual representation of the complexity of teaching scenarios based on the frequency and intensity of academic errors and problem behavior</td>
<td>18</td>
</tr>
<tr>
<td>2.</td>
<td>Systematic searching and screening process</td>
<td>23</td>
</tr>
<tr>
<td>3.</td>
<td>The realism of classroom simulators on a continuum from less realistic to more realistic</td>
<td>34</td>
</tr>
<tr>
<td>4.</td>
<td>TeachLivE middle school students</td>
<td>56</td>
</tr>
<tr>
<td>5.</td>
<td>Lisa’s percentage of correct steps for error correction, specific praise rate per min, and percentage of correct steps for praise around on the TeachLivE training assessment, TeachLivE comprehensive assessment, and classroom generalization assessment</td>
<td>94</td>
</tr>
<tr>
<td>6.</td>
<td>Grayce’s percentage of correct steps for error correction, specific praise rate per min, and percentage of correct steps for praise around on the TeachLivE training assessment, TeachLivE comprehensive assessment, and classroom generalization assessment</td>
<td>99</td>
</tr>
<tr>
<td>7.</td>
<td>Deanna’s percentage of correct steps for error correction, specific praise rate per min, and percentage of correct steps for praise around on the TeachLivE training assessment, TeachLivE comprehensive assessment, and classroom generalization assessment</td>
<td>104</td>
</tr>
<tr>
<td>8.</td>
<td>Marie’s percentage of correct steps for error correction, specific praise rate per min, and percentage of correct steps for praise around on the TeachLivE training assessment, TeachLivE comprehensive assessment, and classroom generalization assessment</td>
<td>109</td>
</tr>
</tbody>
</table>
CHAPTER I

INTRODUCTION

The goal of preservice teacher training programs is to develop educators’ knowledge and skills and to help them apply essential skills in authentic classroom settings (Allsopp, DeMarie, Alvarez-McHatton, & Doone, 2006). Teacher preparation programs are currently under a great deal of pressure to produce competent teachers, especially in this era of high-stakes testing and legislation focused on student outcomes (Girod & Girod, 2006; Lignugaris/Kraft, Sindelar, McCray, & Kimerling, 2014). Effectively preparing special education teachers is of particular concern because the number of students with disabilities is increasing each year, and the demand for special educators is expected to increase over time (McLeskey, Tyler, & Flippin, 2004; Tyler & Brunner, 2014). High attrition in special education exacerbates the demand for teachers and suggests that teachers may be exiting their training programs inadequately prepared to meet the demands of full-time teaching positions.

Critical competencies for special educators can be categorized in two core domains of student outcomes: academics and behavior. Special educators need to competently initiate academic and behavioral engagement with students and capably respond to students’ academic errors or behavioral difficulties. In the academic domain, teachers must learn to proficiently ask high quality questions that engage students in the content (see Greenwood, Delquadri, & Hall, 1984), deliver questions with adequate pacing (see Stichter et al., 2009), and respond to academic errors to increase student performance (see Lignugaris/Kraft & Harris, 2014). In the behavior domain, educators
need to increase specific praise for positive student behavior and strategically deliver this feedback to shape desired behaviors in the classroom (see Dawson & Lignugaris/Kraft, 2015).

Field Placements

Teacher educators have long acknowledged the gap between teachers’ knowledge and their application of essential skills (March, 2002; Pretti-Frontczak, Brown, Senderak, & Walsh, 2005) and attempted to structure preparation programs to close this gap (Allsopp, DeMarie, Alvarez-McHatton, & Doone, 2006; Brownell, Ross, Colón, & McCallum, 2005). Teacher education programs typically include two major components: coursework and field experiences. Coursework is where preservice educators are exposed to the basic theories of teaching and learning, and field placements are where they apply related strategies (Rosenberg, Jackson, & Yeh, 1996).

Field placements are beneficial because they give preservice teachers an opportunity to interact with students, colleagues, and administrators. These interactions may positively influence preservice teachers’ understanding of diversity and their ability to communicate with individuals who are different (Lin, Lake, & Rice, 2008). Additionally, field experiences help preservice educators understand how factors such as school culture, district policies, and state legislation influence daily classroom functions. Finally, field placements give preservice teachers opportunities to balance academic and behavioral issues in actual classrooms (Cruickshank, 1986). Techniques for concurrently managing academic and behavior issues in the classroom cannot be represented with
adequate complexity through didactic instruction alone.

In spite of the documented benefits of field experiences, there are substantial limitations that must be considered as well. It is difficult for teacher educators to align the field experience with the intended purposes of the placement because many classroom and school factors cannot be controlled (e.g. curriculum, diversity of the students, school culture, quality of administration; Cruickshank, 1986). Furthermore, it is impossible for teacher educators to match the complexity of field placements with the performance level of the teacher. If preservice teachers are not ready for the cognitive demands of the field experience they are not likely to benefit fully from the placement (Hixon & So, 2009). For example, Moore (2003) noted that teachers often shift their focus to routine tasks and procedural issues and away from high-quality teaching skills during practicum placements. Similarly, Girod and Girod (2006, 2008) explain that the complexities of real classrooms often require teachers to pick and choose what to focus on, especially when they are not yet proficient at juggling multiple skills. For example, teachers may be so overwhelmed with elements of classroom management that they are unable to focus on crucial instructional skills, such as engaging students in active responding and providing high-quality feedback for academic and behavioral performance.

**Situated Learning Methods**

Grossman and McDonald (2008) suggested that a critical component in teacher preparation programs is the opportunity to practice complex teaching skills in classroom situations that successively approximate actual practice. Similarly, according to
Thomassen and Rive (2010), it is necessary to create an “enabling context” (p. 159) for learners to become proficient with target skills. Traditional field placements may be too complex for novice teachers to learn new skills. Thus, novice educators may need additional simplified contexts for practicing essential teaching skills. In response to this need, teacher educators often utilize strategies that simulate real classroom scenarios. These methods emerged from the theory of situated learning, which proposes that knowledge acquisition requires realistic context and complexity and that knowledge transfer depends on how closely practice opportunities match the situation in which the learner is to apply the information (J. S. Brown, Collins, & Duguid, 1989). Utley (2006) summarizes the critical aspects of situated learning approaches as (a) interacting socially, (b) solving problems in realistic contexts, and (c) creating a community of learners.

Situated learning methods such as case-based instruction, roleplaying, and virtual simulations are implemented in teacher preparation programs to supplement didactic instruction and to prepare educators to succeed in authentic classroom environments.

**Case-Based Instruction**

Various case-based instructional strategies are currently implemented in preservice teacher training programs, including written case studies, video cases, problem-based learning, and web-based instruction. Although the titles and instructional mediums may differ, these case-based strategies are all similar in that they provide realistic classroom examples and dilemmas that promote discussion and problem solving for future educators (Levin, Hibbard, & Rock, 2002). Case-based instruction supplements traditional instruction and exposes preservice educators to the complex dilemmas
teachers face in the field.

Compared to traditional lectures, case-based instruction may promote better knowledge acquisition (Yadav, Bouck, Da Fonte, & Patton, 2009), longer retention of information (Langone, Malone, & Clinton, 1999; Malone & Langone, 2005), and better performance on process-oriented tasks, such as essay prompts (Langone et al., 1999). However, the effectiveness of case-based instruction is limited because it does not require active application of discrete teaching skills.

**Roleplaying**

Roleplaying is perhaps the oldest form of classroom simulation, dating back to the 1800s (A. H. Brown, 1999). In the 1960s, the term “microteaching” emerged (Amobi & Irwin, 2009), which refers to a type of structured roleplaying used in training settings that involves planning instruction, delivering instruction, and reflecting on instruction (Diana, 2013). Roleplaying and microteaching differ from case-based instruction because they require preservice teachers to apply knowledge by actively practicing specific skills. In this way, roleplaying requires engagement in behaviors more closely aligned to actual classroom teaching than does case-based instruction.

However, Brownell, Chard, Benedict, and Lignugaris/Kraft (in press) and Grossman (2005) noted that microteaching has produced mixed outcomes for teachers, including questionable transfer to authentic teaching environments. Another concern is that it may cultivate beliefs about instruction that are too simplistic for the complexities of actual classroom teaching (Bell, 2007; Brownell et al., in press). These limitations are not surprising when considering the theory of situated learning. The degree to which
roleplaying scenarios match real classrooms is limited because preservice teachers typically deliver lessons to their university peers (see Brownell et al., in press; Diana, 2013), which does not require them to realistically respond to the range of academic and behavioral challenges they will encounter in the classroom.

**Virtual Simulations**

Virtual technologies have emerged in the field of teacher education, largely because of their success in medical and military training (Dieker, Hynes, Hughes, & Smith, 2008). Virtual simulations differ from roleplaying scenarios because teachers interact with virtual students instead of adult colleagues, and, therefore, the interactions may more realistically approximate authentic teaching situations. Virtual classrooms are designed to represent real classrooms, but in contrast to traditional field placements, virtual environments provide a safe and controlled environment for teachers to practice skills (Billingsley & Scheuermann, 2014; Powers & Darrow, 1994).

Virtual simulations are safe for students because instructional time in a simulated classroom does not take away from instructional time in actual classrooms, and teacher mistakes do not harm real people or negatively impact student outcomes as they would in an authentic field setting (Strang, Landrum, & Lynch, 1989). The environment is also safe for teachers because the complexity of the simulation can be programmed to match the teacher’s zone of proximal development, which can be summarized as the difference between an individual’s current level of performance and the potential level of performance that can be achieved with guidance (Levykh, 2008). Adequately challenging simulated practice opportunities, coupled with immediate mentoring, provide a
nonthreatening learning experience for preservice educators.

The ability to control simulations to represent realistic classroom interactions without containing all the complexity of actual classrooms is one of the greatest strengths of virtual technology. Eliminating some of complexities that may be present in real field placements permits learners to focus on a few relevant details at a time before moving on to more complex environments (Lloyd, 1983). As preservice educators master basic skills in virtual classrooms, they may better allocate their time and attention to issues of classroom organization, relationships with colleagues, and procedural concerns in real classrooms, without sacrificing effective teaching. Practicing specific skills in controlled virtual environments may prepare teachers to benefit more fully from field placements and promote success in real classrooms (Hixon & So, 2009).

Several technology-based classroom simulations have been developed over the years, with varying levels of realism and complexity. Most classroom simulations used in teacher preparation programs (e.g., simSchool, Cook School District Simulation, and Curry Simulation) rely on artificial intelligence or pre-programmed student actions and comments triggered in response to teacher behavior (see Badiee & Kaufman, 2014; Girod & Girod, 2006, 2008; simSchool, 2015; Strang, Landrum, & Ulmer, 1991; Strang, Murphy, Kauffman, Badt, & Loper, 1986; Zibit, & Gibson, 2005). Recently, the University of Central Florida developed a more sophisticated classroom simulator, the TLE TeachLivETM classroom simulation laboratory (TeachLivE), which is considered a mixed-reality environment because it utilizes both artificial and human intelligence (Dieker, Straub, Hughes, Hynes, & Hardin, 2014). Some of the basic student behaviors
are preprogrammed and can be triggered manually (e.g. students laughing, raising their hands, talking to one another), but most of the interactions occur in real time by an actor playing the part of the students (i.e. interactor). This “human-in-the-loop” (Nagendran, Pillat, Kavanaugh, Welch, & Hughes, 2014, p. 112) creates authentic responses to teacher behavior based on extensive student profiles, providing a rich simulation that realistically represents many of the nuances of an actual classroom. According to situated learning theory, practice opportunities in this realistic simulation should promote both skill development and skill transfer to authentic environments.

**Purpose and Research Questions**

Teacher educators are concerned with developing preservice teachers’ understanding of critical competencies and promoting their implementation of these skills in classrooms. Virtual simulations are a promising method for helping new teachers learn critical instruction and management skills. In addition, virtual simulations provide a safe and controlled context in which teachers may practice target skills. TeachLivE is an especially powerful simulator because students respond to teacher behavior in real time, making the interactions authentic and realistically complex. Based on the theory of situated learning, this close alignment between practice opportunities in TeachLivE and real interactions in the classroom increases the likelihood that teachers will demonstrate proficiency with essential skills in both environments. The purpose of this study is to investigate the effectiveness of TeachLivE simulations on teachers’ development of essential teaching skills and the extent to which they generalize these skills to authentic
special education classrooms. The research questions for this investigation were as follows.

1. To what extent will preservice special educators develop proficiency with error correction, specific praise, and praise around during TeachLivE sessions, as measured by their performance during the third practice opportunity during each session (i.e., TeachLivE training assessment)?

2. To what extent will preservice special educators demonstrate proficiency with error correction, specific praise, and praise around when teaching in more complex simulated scenarios, as measured by their performance immediately following each TeachLivE session and during a maintenance session approximately 1.5 months following intervention (i.e., TeachLivE comprehensive assessment)?

3. To what extent will preservice special educators generalize proficiency with error correction, specific praise, and praise around from the TeachLivE comprehensive assessments to authentic teaching situations, as measured by their weekly performance in their own classroom and during maintenance sessions approximately one month following intervention (i.e., classroom generalization assessment)?

4. How do preservice special educators perceive the realism of the TeachLivE classroom and value the intervention and assessment procedures, as measured by a social validity survey administered at the conclusion of the intervention sessions?
CHAPTER II
LITERATURE REVIEW

The dual purpose of this literature review is (1) to briefly summarize the research supporting essential teaching skills and (2) to provide a systematic research synthesis of how interacting with virtual students has been used to support educators’ development of teaching techniques.

Essential Teaching Skills

There are a number of pedagogical skills that create a strong foundation for effective classroom teaching. The most vital competencies, particularly for teachers of students with special needs, are foundation skills that can be applied across a range of instructional settings and content areas (Brownell et al., in press; Hiebert & Morris, 2012; Lignugaris/Kraft & Harris, 2014). Special educators’ ability to flexibly apply generic teaching techniques in a variety of contexts may be more critical than content knowledge in promoting student achievement (Brownell et al., 2009). Essential pedagogical skills typically target student outcomes in either the academic domain or behavior domain. A brief overview of research supporting selected teaching skills in each domain is provided below, followed by a discussion of how the target skills are interrelated and how teacher educators can promote the development of a balanced repertoire of target skills.

Academic Domain

Concern about student achievement has gained political attention in recent years
and is the primary focus of many current reform efforts, including the shift towards response-to-intervention frameworks (Wixsom, 2011). Academic achievement is difficult to connect directly to an individual teacher’s performance. This is especially true for students with disabilities who receive services across general and special education settings with multiple service providers (Lignugaris/Kraft et al. 2014), or when analyzing global measures of achievement, such as standardized test scores. However, when examining academic responding of individual students at the classroom level, two effective teaching skills include providing opportunities for students to respond and correcting academic errors.

**Opportunity to respond (OTR).** The term “opportunity to respond” emerged from the work of Greenwood et al. (1984) as they attempted to shift the discussion of students’ academic performance from personal variables, such as intelligence and attitude, to environmental variables, such as teacher delivery and modification of instruction. They defined an OTR as a stimulus provided by the teacher (e.g. asking a question or initiating a task) that results in an observable student response (e.g. reading, writing, discussing, motor behaviors). Researchers in every decade since the 1970s have documented that a teacher’s ability to produce active student responding is critical for promoting engagement and achievement (e.g. Carnine, 1976; Greenwood et al., 1984; Heward et al., 1996; MacSuga-Gage & Simonsen, 2015; Simonson, Myers, & DeLuca, 2010; Stanley & Greenwood, 1983; Sutherland, Alder, & Gunter, 2003; Sutherland & Wehby, 2001).

Delivering OTRs with appropriate pacing creates instructional momentum and
maximizes students’ practice with the content. High OTR rates are linked to improved student performance for students with varying abilities (Carnine, 1976; Lignugaris/Kraft & Harris, 2014; MacSuga-Gage & Simonsen, 2015; Sutherland et al., 2003; Stichter et al., 2009), including children with learning disabilities, mental retardation, or emotional and behavioral disorders (Stichter et al., 2009; Sutherland & Wehby, 2001). Optimal OTR rates vary depending on the complexity of the task, newness of the material, ability level of students, and interresponse time (Lignugaris/Kraft & Harris, 2014; Lignugaris/Kraft & Rousseau, 1982), but according to Stichter et al. (2009), “3.5 prompts per min during active instruction with students could serve as a ‘tipping point’ at which increased student engagement and achievement are supported” (p. 69). MacSuga-Gage and Simonsen corroborated this target rate based on their systematic review of the literature, concluding that desired student outcomes are observed when teachers deliver around three to five OTRs per min.

Teachers can structure OTRs with varying difficulty, from questions that require low-level recall to tasks that require high-level analysis and synthesis of information (see Bloom & Krathwohl, 1956). One technique that has garnered some research support is to first ask low-level questions and then strategically shift to higher-level questions based on correct student responding (Lignugaris/Kraft & Harris, 2014). This direction from low- to high-level questions increases student accuracy by establishing basic knowledge to support success on subsequent, more difficult questions (Pressley et al., 1992). Supporting students’ accuracy on complex questions is especially critical with the recent adoption of Common Core State Standards because the standards predominantly require
higher order processes (Lignugaris/Kraft et al., 2014).

**Error correction.** Identifying and correcting student errors consistently and immediately is an essential teaching competency for promoting student accuracy on academic tasks. Error correction is a foundational skill that teachers can flexibly apply across a variety of content areas. Researchers have demonstrated the effectiveness of error correction procedures for students with varying ability levels across a range of target skills, such as passage reading (Jones, Lignugaris/Kraft, & Peterson, 2007; Watson, Fore, & Boon, 2009), sight words (Barbetta, Heron, & Heward, 1993; Barbetta, Heward, Bradley, & Miller, 1994), math facts (Bennett & Cavanaugh, 1998) and geography concepts (Barbetta, & Heward, 1993).

Lignugaris/Kraft and Harris (2014) reviewed a number of error correction approaches used in the classroom and concluded that the procedure with the strongest empirical support includes a model, test, and delayed test. This means that the teacher (a) demonstrates the correct response immediately following the error, (b) provides an opportunity for the student to independently respond to the question following the model, and (c) delivers the question to the student again later in the lesson. Other procedures they reviewed include a cue or prompt instead of a direct model, a model-only procedure, or model/test without a delayed test. Addressing errors with a complete model/test/delayed test procedure places low demand on a student because the child can simply repeat the teacher’s model during the test (Jones et al., 2007), which increases the probability that the student will also accurately respond to the question during the delayed test (Bennett & Cavanaugh, 1998).
Behavior Domain

Teachers commonly cite student misbehavior as a major cause of burnout and attrition and also indicate that they receive the least amount of training in this area (Maag, 2001; Reinke, Herman, & Stormont, 2013). Classroom management is a classic source of stress (Pisacreta, Tincani, Connell, & Axelrod, 2011), especially for beginning teachers (Wolff, van den Bogert, Jarodzka, & Boshuizen, 2015). School-Wide Positive Behavioral Interventions and Supports (SW-PBIS) is a universal approach for improving student behavior that has gained popularity in schools and produced promising reductions in problem behavior (Bradshaw, Mitchell, & Leaf, 2010). However, Reinke et al. pointed out that teachers need to implement effective behavior management strategies at the classroom level for SW-PBIS to be as successful as possible. Two teaching strategies for promoting desired behavior in the classroom are specific praise and praise around.

Specific praise. According to Hattie (2009), feedback is the single most powerful practice influencing student achievement. Feedback can serve as a corrective measure (such as error correction procedures) or as an affirmation of desired behavior or academic performance. Praise is among the most commonly recommended type of teacher feedback, perhaps because it is so readily accessible in any instructional setting. Unfortunately, praise is also a common skill deficit among educators (Kalis, Vannest, & Parker, 2007), especially for teachers of older children (Anderson, Evertson, & Brophy, 1979), or teachers of students with challenging behaviors (Fullerton, Conroy, & Correa, 2009; Sutherland, Wehby, & Copeland, 2000). Furthermore, teachers are unlikely to implement praise in the classroom without receiving targeted intervention (see Fullerton...
et al., 2009; Pisacreta et al., 2011; Sutherland et al., 2000).

Researchers recommend that teacher praise should be contingent and specific, meaning the statement follows appropriate student behavior (Brophy, 1981) and explicitly identifies what aspect of the behavior was performed well (e.g. Allday et al., 2012; Brophy, 1981, Chalk & Bizo, 2004; Duchaine, Jolivette, & Fredrick, 2011; Filcheck, McNeil, & Herschell, 2001; Haydon & Musti-Rao, 2011; Hemmeter, Snyder, Kinder, & Artman, 2011; Myers, Simonsen, & Sugai, 2011; Reinke, Lewis-Palmer, & Martin, 2007; Sutherland et al., 2000). In a systematic research review of experimental studies that address specific praise, Dawson and Lignugaris/Kraft (2015) concluded that delivering specific praise directly to target students is an empirically supported treatment for improving student outcomes, the majority of which were behavior outcomes (for guidelines for determining the evidence base of an intervention and qualifying it as an empirically supported treatment, see Cook et al., 2014; Horner et al., 2005; Horner, Swaminathan, Sugai, & Smolkowski, 2012). Specific praise is a powerful, accessible, and relatively simple intervention for increasing desired behavior and decreasing problem behavior for students who are reinforced by adult attention.

**Praise around.** Praise around is a nonconfrontational technique for addressing common behavior challenges in the classroom. It is appropriate to use with students for whom teacher attention and praise are conditioned reinforcers, in response to a range of problem behaviors that are not self-injurious or dangerous to others. Praise around is a two-part procedure that leverages the effectiveness of both vicarious reinforcement and direct reinforcement. The vicarious reinforcement component occurs first, when the
teacher delivers a specific praise statement to a student who is engaged appropriately while the misbehaving student observes the interaction. According to Bandura (1965), this may motivate the observing student to modify their behavior because they believe they will receive reinforcement for behaving similarly. The direct reinforcement component occurs next, when the teacher delivers a specific praise statement to the target student once they exhibit the desired behavior. This step capitalizes on the power of specific praise and ensures that the target student receives positive attention for behaving appropriately.

In their review on specific praise, Dawson and Lignugaris/Kraft (2015) discovered a pattern of student outcomes that supports the use of the praise around procedure. They found when specific praise is delivered vicariously desired student behavior or academic responses increase initially but that the positive effects tend to deteriorate over time (see Ollendick & Shapiro, 1984; Ollendick, Shaprio, & Barrett, 1982; Strain & Pierce, 1977). In contrast, when students observe vicarious specific praise and later receive direct specific praise, desired responses will likely maintain or increase (e.g. see Ollendick, Dailey, & Shaprio, 1983).

The praise around procedure is conceptually consistent with differential reinforcement of incompatible or alternative behaviors, a technique that effectively reduces problem behavior by withholding reinforcement for such behavior and providing reinforcement for a preferred behavior instead (Cooper, Heron, & Heward, 2007; Rhode, Jenson, & Reavis, 1992). An important benefit of praise around is that it prompts appropriate behavior without the use of reprimands, which is a typical teacher response to
misbehavior (Gable, Hester, Rock, & Hughes, 2009; Maag, 2001). Reprimanding is problematic because it may maintain inappropriate behavior with attention (Pisacreta et al., 2011) and does little to teach socially appropriate replacement behaviors (Maag, 2001). In contrast, when teachers use praise around they explicitly state the desired behavior as they differentially attend to students, contingent on students’ performance of the specified behavior.

**Developing a Balanced Repertoire of Essential Teaching Skills**

In practice, OTRs, error correction, specific praise, and praise around are interconnected techniques that contribute to the overall learning environment and impact student outcomes across academic and behavior domains. For example, high-quality OTRs engage students in the content and maximize their time on-task. OTRs also establish the basic teacher-student interaction (i.e. question followed by a response) that lays the foundation for many other critical teaching interactions, such as delivering praise for correct responses or error correction for incorrect responses. Likewise, correcting errors increases OTRs, promotes accurate responding, and occasions praise for academic performance. Cyclically, effectively managing behavior with specific praise and praise around decreases problem behavior and increases time available for content instruction, potentially resulting in higher OTR rates. Other researchers have similarly noted interactions among these essential teaching skills (e.g. see Gable et al., 2009; Lignugaris/Kraft & Harris, 2014; Sutherland, Wehby, & Yoder, 2002).

Simple classroom interactions necessitate only a limited number of basic teaching
skills, but more complex scenarios require teachers to balance several skill sets simultaneously. For example, the simplest teaching interaction involves the teacher asking a question, the student answering correctly, and the teacher providing positive feedback. This scenario puts low demand on teachers and only requires the delivery of OTRs and praise. Teaching interactions become more complex when they are disrupted by student errors and/or challenging behavior, requiring teachers to deliver multi-step error correction and behavior redirection (e.g. praise around) procedures. Moreover, the complexity of teaching scenarios is amplified as the frequency and intensity of disruptors (i.e. student errors and problem behavior) increase, and if disruptors overlap. Figure 1 illustrates how the overall frequency and intensity of disruptors may impact the difficulty of teaching interactions.

![Figure 1](image-url)

**Figure 1.** A visual representation of the complexity of teaching scenarios based on the frequency and intensity of academic errors and problem behavior. As the overall frequency and intensity of academic and behavioral disruptors increase, so does the difficulty of the teaching interaction.
Academic and behavioral disruptors increase the cognitive load on teachers, especially novice educators who are not yet proficient with essential teaching skills (Moos & Pitton, 2014). According to cognitive load theory, dealing with complex tasks requires that limited working memory functions with unlimited long-term memory (Paas & Ayres, 2014). Cognitive load is optimized by strategically combining previously learned information or skills with new information or skills, resulting in an individuals’ ability to deal with increasingly complex problems. Ideally, teacher educators should manage preservice teachers’ cognitive load by first building success during simple practice scenarios and incrementally increasing the complexity of the practice opportunities over time (Grossman & McDonald, 2008).

Ultimately, teachers need to proficiently implement multiple skills simultaneously, which is called interleaving (Dunlosky, Rawson, Marsh, Nathan, & Willingham, 2013). Correctly interleaving instruction and behavior management skills requires teachers to discriminate what strategies to use, and how to use them, depending on the situation (Brownell et al., in press). Importantly, in most educational contexts, focused or blocked instruction is provided first on a specific skill or strategy, and then the target skill is interleaved with other skills in the learner’s repertoire (Dunlosky et al., 2013). Researchers have demonstrated that interleaving improves retention and application of skills, especially on delayed assessments (Dunlosky et al., 2013; Taylor & Rohrer, 2010). In teacher education, interleaved practice may promote the flexible application of essential skills appropriate to the teaching context, which is what teachers are required to do in authentic teaching situations.
In summary, effective teachers strategically deliver interrelated instruction and management techniques appropriate to the learning context, instead of relying on isolated or rigid skill sets (Lignugaris/Kraft & Harris, 2014; Stronge, Ward, & Grant, 2011). Thus, the goal is to promote teachers’ development of a balanced repertoire of essential target skills that can be applied to situations of varying complexity. Teacher educators can facilitate this process by cumulatively programming practice opportunities that involve both the development of new skills and the maintenance of formerly practiced skills. This approach incrementally increases teacher’s cognitive load as they interleave an expanding array of teaching skills in increasingly disruptive instructional contexts. Creating a sequence of increasingly complex practice opportunities is difficult to achieve with traditional instruction and field-based approaches, but virtual technology makes it possible to engineer such practice opportunities.

**Virtual Simulations and Teacher Preparation**

Historically, simulated teaching experiences involved either roleplaying with colleagues or engaging in cased-based problems presented via print. In the mid-1970s, technology-based simulations emerged in teacher preparation programs (Cruickshank, 1988), allowing teachers to interact with virtual students and solve problems presented via computers. Technology-based simulations have become increasingly accessible in teacher preparation programs, largely due to the development and evolution of microcomputers (now known as desktop or personal computers).

Most virtual classrooms are accessed from a personal computer, including game-
like environments and Second Life simulations. Game-like simulators allow teachers to instruct students and make ongoing instructional decisions (Dieker, Rodriguez, Lignugaris/Kraft, Hynes, & Hughes, 2014), usually by selecting from a menu of options, which triggers a range of preprogrammed student responses. Interactions in Second Life are more fluid because the user embodies an avatar and interacts with other human-generated avatars in an online environment by typing or talking (Brownell et al., in press; Dieker et al., 2014).

The most sophisticated type of virtual classroom available is a full-immersion simulator (such as TeachLivE). Interactions during full-immersion simulations differ from game-like classrooms and Second Life environments because teachers enter a physical classroom with life-size student avatars projected on a screen. They engage in realistic teaching dialogue as they physically move through the classroom, while the human-in-the-loop produces authentic student responses. Practicing target skills in a full-immersion simulator is more like real teaching because teachers verbally instruct students instead of typing or selecting responses, and they interact as themselves instead of assuming the identity of an avatar.

As technology-based simulations gain traction in teacher preparation programs, it is important to understand how they can be used most effectively to maximize teacher outcomes. Therefore, the purpose of this research synthesis was to analyze the available evidence on virtual simulations’ impact on teachers’ pedagogical skills. The research question driving this literature review was: How are virtual simulations of K-12 students used with preservice or inservice teachers, and how do they impact teacher performance
in virtual and real classrooms?

Procedures

Searching and screening. Qualifying studies for this synthesis were identified through database searching, screening procedures, and ancestral searches. First, systematic electronic searches were conducted in three common education databases (i.e., Academic Search Premier, ERIC, and PsychINFO) to identify potential studies using virtual simulations with teachers. Searches were conducted using Boolean operators and truncation. Search terms included seven descriptors for the independent variable (i.e. virtual reality, virtual simulat*, virtual classroom, simulated environment, classroom simulation, and second life) and six terms for the purpose of the simulation training or the target population (i.e., teacher education, teacher preparation, teacher train*, professional development, preservice teach*, inservice teach*). Searches were conducted by combining each term for the independent variable with each term for the purpose/population, resulting in 42 total searches. Next, eight additional searches were conducted using the proper names of classroom simulations found in the literature (i.e., TeachLivE, simSchool, Classroom SIMS, Cook School District Simulation, Curry Simulation, Class SIM, VirtualPREX, and UTAS virtual class). These 50 initial searches yielded 1,097 articles for review.

As illustrated in Figure 2, these articles were screened in the following phases:

1. Title screening: The titles of all 1,097 potential articles were reviewed to remove all duplicates.

2. Abstract screening: The abstracts of 613 unique articles were reviewed to identify all relevant studies that reported teacher outcomes.
**Database Search**
Records identified by searching Academic Search Premier, ERIC, and PsycINFO using preselected search terms

**Title Screening**
n = 1097
All duplicates removed

**Abstract Screening**
n = 613
Identified relevant studies with teacher outcomes

**Article Screening**
n = 33
Full articles screened for all inclusion criteria

**Qualified studies**
n = 9

**Ancestral Search**
n = 341
(Round 1 = 277, Round 2 = 66)
Repeated screening process from the beginning for all references cited in qualified studies

**Additional Qualified Studies**
n = 2

**Total Qualified studies**
(Database + Ancestral)
n = 11

*Figure 2. Systematic searching and screening process.*
3. Article screening: The full articles for 33 studies were reviewed to locate experimental studies that met all inclusion criteria. Nine articles qualified at this point in the review process.

Next, ancestral searches were conducted for the nine qualifying articles. The reference lists contained 277 additional records to review, and two additional studies qualified for the synthesis. Finally, an ancestral search of these two articles was conducted, resulting in 66 records to review and no new qualifying studies. In total, 11 studies (eight group design, three single subject design) qualified for the synthesis based on the following inclusion/exclusion criteria.

1. The study was published in a peer-reviewed journal in the English language between 1965 and spring of 2016.

2. The study was a group design or single subject study. Because of the limited number of available studies, group design studies with only one group were allowed, as long as pretest and posttest measures were reported. Qualitative studies and correlation studies were excluded.

3. The simulation was technology-based. Simulations involving traditional role playing, discussing cases, or reflecting on pedagogical skills were excluded.

4. Participants in the simulation were preservice or inservice teachers of general education or students with disabilities in preschool through high school.

5. The simulation required active instruction or decision-making from the teachers to virtual students. Simulations that did not require active interaction (e.g., observing a teaching scenario, learning content knowledge) with virtual students (e.g., virtual parents, colleagues, or administrators) were excluded.

6. At least one outcome was reported for teachers’ performance during the simulation, during an assessment outside of the simulation, or during authentic classroom instruction. Studies were excluded if the only reported measures were content knowledge, self-perception of skills, attitudes about teaching, or reflections about the simulated experience. (For studies with multiple outcome measures, only those relevant to the review are reported in the synthesis).
**Coding.** Coding documents were created to organize data from the qualified studies in three areas: study characteristics, experimental outcomes, and methodological quality. Study characteristic codes included information about (a) participants, (b) setting/context, (c) independent variable, including a description of the virtual simulation and all other components used in the training (e.g., didactic instruction, feedback), (d) dependent variables, including the primary outcomes as well as any generalization or maintenance measures, and (e) study design.

Coding for experimental outcomes occurred in two steps. First, the outcomes were analyzed and reported using statistical analysis for group design studies and visual analysis for single subject studies. For group design studies, the $p$ value for statistical significance and Cohen’s $d$ effect sizes for the relevant analyses were reported. If the effect size was not reported in the study, it was calculated from the provided information. For single subject studies, each point of contrast was visually analyzed and reported as an effect, non-effect, or negative effect. These codes were based on the change (if any) in level, trend, and variability between phases, and if the change favored the treatment (see Kratochwill et al., 2010). Second, a summary of effects was determined for each study, based on the guidelines established by Cook et al. (2014). Positive effects were defined as $d \geq 0.40$ for group studies and as 75% or more of coded phase changes showing a change in the therapeutic direction for single subject studies. Negative effects were defined as $d \leq -0.40$ for group studies and as 75% or more of the coded phase changes showing a change in the non-therapeutic direction for single subject studies. Mixed/neutral effects were defined as $-0.40 \leq d \geq 0.40$ for group studies and as not meeting the criteria for
either positive effects or negative effects for single subject studies.

Methodological quality codes were based on Cook et al.’s (2014) guidelines and included information about (a) context and setting, (b) participants, (c) intervention agent, (d) implementation fidelity, (e) internal validity, (f) outcome measures/dependent variables, and (g) data analysis. Each dichotomous indicator was scored for meeting the criterion (+) or not meeting the criterion (-). Several indicators applied to both group design and single subject studies, but some indicators were specific to one design type. There were 24 indicators for group design studies and 22 indicators for single subject.

Findings

The 11 qualified studies were published between 1986 and 2016. Although this covers 30 years, there is a distinct split in publication dates with a 15-year gap in between. Five studies (46%) were published between 1986 and 1991 and six studies (55%) were published between 2006 and 2016. Although virtual simulations were utilized and under development between 1991 and 2006, no published studies during that timeframe met the inclusion criteria for the current synthesis. The results of the qualified studies are synthesized below based on the three coding categories. First, pertinent study characteristics are summarized. Next, outcome patterns are reported. Finally, methodological quality is discussed.

Study characteristics. The numerous study characteristics are detailed in the following sections.

Participants. Cumulatively, 344 teachers participated in the studies, of which 284 were preservice teachers and 60 were inservice teachers. As shown in Table 1, the
Table 1

Virtual Simulation Participant and Setting Characteristics

<table>
<thead>
<tr>
<th>Study</th>
<th>Participants</th>
<th>Age</th>
<th>Specialization</th>
<th>Experience</th>
<th>Setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Badiee &amp; Kaufman (2014)</td>
<td>N = 22</td>
<td>NR</td>
<td>Preservice elementary or secondary student teachers</td>
<td>NR</td>
<td>Recruited from classes and computer labs in a teacher education program</td>
</tr>
<tr>
<td></td>
<td>F = 19; M = 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brubacher et al. (2015)</td>
<td>N = 36</td>
<td>NR</td>
<td>Inservice elementary general education</td>
<td>1-41 years</td>
<td>Recruited from local Canadian elementary schools</td>
</tr>
<tr>
<td></td>
<td>F = 24; M = 12</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fisher et al. (2010)</td>
<td>Teachers:</td>
<td>23-55</td>
<td>Inservice middle school general education teachers</td>
<td>1-31 years</td>
<td>Volunteers from three middle schools in the same city</td>
</tr>
<tr>
<td></td>
<td>N = 8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Treatment = 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Comparison = 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>F = 7; M = 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Students:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>N = 125</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Treatment = 76</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Comparison = 49</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>F = 57; M = 68</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Girod &amp; Girod (2006)</td>
<td>N = 71</td>
<td>&lt; 25 to &gt; 31</td>
<td>Preservice secondary teachers (various content areas)</td>
<td>NR</td>
<td>Self-selected into groups from a teaching master’s program</td>
</tr>
<tr>
<td></td>
<td>Treatment = 33</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Comparison = 38</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>F = 33; M = 38</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Murphy et al. (1987)</td>
<td>N = 18</td>
<td>NR</td>
<td>Preservice general or special education teachers</td>
<td>0-3 years</td>
<td>Enrolled in general education or special education undergraduate methods courses</td>
</tr>
<tr>
<td></td>
<td>F = 17; M = 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*(table continues)*
<table>
<thead>
<tr>
<th>Study</th>
<th>Participants</th>
<th>Age</th>
<th>Specialization</th>
<th>Experience</th>
<th>Setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strang et al. (1987)</td>
<td>N = 32</td>
<td>NR</td>
<td>Preservice teachers (emphasis area NR)</td>
<td>2-8 years</td>
<td>Preservice teachers were enrolled in undergraduate teaching methods course</td>
</tr>
<tr>
<td></td>
<td>Preservice = 20</td>
<td></td>
<td></td>
<td>(inservice</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Inservice = 12</td>
<td></td>
<td></td>
<td>teachers)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(Gender NR)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Strang et al. (1986)</td>
<td>N = 34</td>
<td>NR</td>
<td>Preservice teachers (emphasis area NR)</td>
<td>NR</td>
<td>Enrolled in introductory education course</td>
</tr>
<tr>
<td></td>
<td>(Gender NR)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Strang et al. (1989)</td>
<td>N = 61</td>
<td>NR</td>
<td>Preservice teachers (emphasis area NR)</td>
<td>None</td>
<td>Participation in simulation was part of an introductory education course</td>
</tr>
<tr>
<td></td>
<td>F = 53; M = 9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Strang et al. (1991)</td>
<td>N = 52</td>
<td>NR</td>
<td>Preservice teachers (emphasis area NR)</td>
<td>NR</td>
<td>Participation in simulation was part of an introductory education course</td>
</tr>
<tr>
<td></td>
<td>(Gender NR)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vince Garland et al. (2012)</td>
<td>N = 4</td>
<td>23-54</td>
<td>Inservice Special Education</td>
<td>2-15 years</td>
<td>Recruited from program for Graduate Certificate for Autism Spectrum Disorder</td>
</tr>
<tr>
<td></td>
<td>F = 4; M = 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vince Garland et al. (2016)</td>
<td>N = 6</td>
<td>23-30</td>
<td>Preservice Special Education</td>
<td>NR</td>
<td>Enrolled in Master’s curriculum for individuals with moderate and severe disabilities course</td>
</tr>
<tr>
<td></td>
<td>F = 3; M = 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note. F = Females; M = Males; NR = Not reported.*
majority of research teams (91%) isolated their samples to either preservice or inservice teachers, with the exception of Strang, Kauffman, Badt, Murphy, and Loper (1987) who used inservice teachers as a baseline comparison group for the preservice teachers. Years of teaching experience for inservice teachers across studies ranged from 1-41 years. Only Murphy, Kauffman, and Strang (1987) reported experience of preservice teachers, which was 0-3 years. Studies included teachers specializing in general education (27%) or special education (27%), with two studies (18%) that included teachers from both specializations. Teachers’ emphasis area was not reported in three studies (27%). Participants’ ages ranged from 23 to 55 years old, as reported in four studies (36%). Researchers from one study also included student participants, consisting of 125 middle school children (Fisher et al., 2010).

Settings/context. Nine studies (82%) were conducted at universities in connection with undergraduate preparation programs (n = 6) or graduate level coursework (n = 3; see Table 1). Participation was embedded into the course requirements in two of these studies (Strang et al., 1989; Strang et al., 1991). In the remaining two studies (18%), virtual simulations were provided as professional development to inservice teacher volunteers from local schools (Brubacher, Powell, Skouteris, & Guadagno, 2015; Fisher et al., 2010).

Independent variable. Simulated practice in a virtual classroom was the primary independent variable in each study. Often, researchers also provided related instruction and feedback to facilitate the experience. The various simulations and related intervention components are summarized below and outlined in Table 2.
<table>
<thead>
<tr>
<th>Study</th>
<th>Simulation</th>
<th>Interaction</th>
<th>Sessions</th>
<th>Instruction</th>
<th>Feedback</th>
</tr>
</thead>
<tbody>
<tr>
<td>Badiee &amp; Kaufman</td>
<td>simSchool (Version 1) Online simulation with 1-5 student avatars, accessed on a personal computer</td>
<td><em>Teacher:</em> Assigns tasks and makes comments from a menu of choices</td>
<td>3 sessions</td>
<td>Introduction to simulation</td>
<td>Computer-generated results and verbal debriefing</td>
</tr>
<tr>
<td>(2014)</td>
<td></td>
<td><em>Students:</em> Respond with text in a speech bubble</td>
<td>(same day)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>20-25 min each</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brubacher et al.</td>
<td>Unreal Interviewing Online simulation with 5 year old avatar, accessed on a personal computer</td>
<td><em>Teacher:</em> Selects the best possible interview question from a menu of choices</td>
<td>3 sessions</td>
<td>Article about best practices in questioning students</td>
<td>Visual feedback on screen after every question during simulation</td>
</tr>
<tr>
<td>(2015)</td>
<td></td>
<td><em>Student:</em> Responds to selected questions with pre-recorded vocalization</td>
<td>(in 7 days)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>≈ 30 min each</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fisher et al.</td>
<td>Classroom simulator for Concept Mastery Routine Software simulation with 3 students, accessed on a personal computer</td>
<td><em>Teacher:</em> Instructs virtual students by typing actions using the keyboard</td>
<td>1 session</td>
<td>Computer training about Concept Mastery Routine (including both training and simulation)</td>
<td>Feedback from “virtual coach” during simulation*</td>
</tr>
<tr>
<td>(2010)</td>
<td></td>
<td><em>Student:</em> Responds to teacher behavior with pre-recorded video clips</td>
<td>≤ 3 hr</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Girod &amp; Girod</td>
<td>Cook School District Simulation Online simulation of instructional decisions based on student profiles (no physical representation of students), accessed on a personal computer</td>
<td><em>Teacher:</em> Inputs instructional information and makes ongoing adaptations based on student responses</td>
<td>2 sessions</td>
<td>2 hr simulation training for treatment group.</td>
<td>Computer-generated data and whole-group debriefing</td>
</tr>
<tr>
<td>(2006)</td>
<td></td>
<td><em>Students:</em> Behavior and academic information in response to teacher input available throughout lesson</td>
<td>(in 3 weeks)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2 hr each</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*(table continues)*
<table>
<thead>
<tr>
<th>Study</th>
<th>Simulation</th>
<th>Interaction</th>
<th>Sessions</th>
<th>Instruction</th>
<th>Feedback</th>
</tr>
</thead>
</table>
| Murphy et al. (1987) | Curry Simulation | *Teacher:* Verbally instructs students  
*Student:* Human system operator produces verbal responses, as cued by the computer program | 2 sessions (same day)  
10-20 min each  
1 follow-up (13-57 days later) | Introduction to simulation | One of the following:  
Misbehavior Record:  
Computer-generated data  
Misbehavior Profile:  
List of effective management strategies |
| Strang et al. (1986) | Curry Simulation | *Teacher:* Verbally instructs students  
*Student:* Human system operator enters codes for teacher behavior and then produces verbal responses prompted by the computer program | 2 sessions (same day)  
11-21 min each | NR | Computer-generated data about teaching interaction |
| Strang et al. (1987) | Curry Simulation | *Teacher:* Verbally instructs students  
*Student:* Human system operator enters codes for teacher behavior and then produces verbal responses prompted by the computer program | 2 sessions (days NR)  
= 20 min each  
1 follow-up (76-139 days later) | Introduction to simulation | Computer-generated data  
High Feedback: View record immediately after Session 1, and prior to session 2  
Low Feedback: View record only prior to session 2 |
| Strang et al. (1989) | Curry Simulation | *Teacher:* Verbally instructs students  
*Student:* Human system operator enters codes for teacher behavior, which triggers computer-synthesized student responses. | 2 sessions (same day)  
13 min each  
1 follow-up (> 74 days later) | Introduction to simulation | Computer-generated data about teaching interaction (verbal assistance available for interpreting data, but no mentoring) |

*(table continues)*
<table>
<thead>
<tr>
<th>Study</th>
<th>Simulation</th>
<th>Interaction</th>
<th>Sessions</th>
<th>Instruction</th>
<th>Feedback</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strang et al. (1991)</td>
<td>Self-Administered version of Curry Simulation</td>
<td><em>Teacher:</em> Types comments and selects responses from available choices.</td>
<td>2 sessions</td>
<td>NR</td>
<td>Computer-generated data about teaching interaction</td>
</tr>
<tr>
<td></td>
<td>Computer simulation with 12 students, pre-programmed by course instructor, accessed on a personal computer</td>
<td><em>Student:</em> Comments appear as text on screen, and facial expressions change to show behavior</td>
<td>(same day)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vince Garland et al. (2012)</td>
<td>TeachLivE (Severe setting)</td>
<td><em>Teacher:</em> Active verbal instruction</td>
<td>6 sessions</td>
<td>Written instructions, definitions, and ongoing modeling and coaching</td>
<td>Visual and verbal performance feedback</td>
</tr>
<tr>
<td></td>
<td>Full-immersion simulation with life-size avatar projected on screen</td>
<td><em>Student:</em> Human interactor playing avatar with autism verbally responds.</td>
<td>(multiple days)</td>
<td>≤ 15 min each</td>
<td></td>
</tr>
<tr>
<td>Vince Garland et al. (2016)</td>
<td>TeachLivE (Severe setting)</td>
<td><em>Teacher:</em> Active verbal instruction</td>
<td>3-5 sessions</td>
<td>Written instructions, definitions, and individualized clinical coaching</td>
<td>Visual and verbal performance feedback</td>
</tr>
<tr>
<td></td>
<td>Full-immersion simulation with life-size avatar projected on screen</td>
<td><em>Student:</em> Human interactor playing avatar with autism verbally responds.</td>
<td>(multiple days)</td>
<td>≥ 15 min each, included 5 teaching trials</td>
<td></td>
</tr>
</tbody>
</table>
Virtual simulations. Six simulations were utilized across the 11 studies. In chronological order from oldest to newest publication date of the first study using the simulation, they are: (1) the Curry Simulation, (2) the Cook School District Simulation, (3) an unnamed virtual workshop, (4) TeachLivE, (5) simSchool, and (6) Unreal Interviewing. A similar team of researchers investigated the effectiveness of the Curry Simulation in all five older studies (46%) published from 1986 to 1991. In the six studies (54%) published since 2006, the same first author examined the effectiveness of TeachLivE in two studies (18%) and distinct research teams examined the effectiveness of the simulations in the remaining four studies (36%).

The sophistication of the virtual simulations varied substantially across systems and directly impacted the realism of teacher interactions with virtual students within each system. The degree to which a simulation approximated the realism of authentic classroom environments was influenced by the following four factors: (a) the stimulus display from which the teacher received input and made decisions, (b) the teacher response mode, (c) how responsive the system was to teacher behavior, and (d) the extent to which system responses represented realistic student behavior. Figure 3 illustrates how well various simulation platforms approximated the realism of classroom environments and a description of each simulation based on the four factors is provided below.

All virtual classrooms reviewed, with the exception of TeachLivE (Vince Garland, Holden, & Garland, 2016; Vince Garland, Vasquez, & Pearl, 2012), were navigated from a personal computer. In four of the personal computer simulations (Cook School District Simulation, unnamed virtual workshop, simSchool, and Unreal
Figure 3. The realism of classroom simulators on a continuum from less realistic to more realistic.

Interviewing), teachers selected actions from a list of options or typed text using the keyboard. The teacher responses then triggered preprogrammed student responses. Across these studies, student responses were conveyed in a variety of ways, including ongoing textual information about student behavior and academic performance (Cook School District Simulation, see Girod & Girod, 2006), comments in a speech bubble (simSchool, see Badiee & Kaufman, 2014), prerecorded verbalizations (Unreal Interviewing, see Brubacher et al., 2015), or video clips of student responses (virtual workshop, see Fisher et al., 2010).

The first version of the Curry Simulation differed from the other desktop virtual classrooms because the teachers instructed the students verbally (Murphy et al., 1987;
Strang et al., 1986, 1987). A human operator behind a partition coded each teacher behavior as it happened, and the system then produced a scripted response that the operator, acting as the student, read out loud. Strang et al. (1989) expanded the technology in the Curry Simulation so the virtual classroom was projected onto a large screen and teachers were no longer restricted to sitting behind a computer. This technique more closely approximated how teachers actually move in a classroom and brought the simulation closer to a full-immersion environment. However, in the latest version of the Curry Simulation included in this synthesis, Strang et al. (1991) prioritized full-automation over full-immersion. The human-in-the-loop was removed and teachers typed comments and responded to textual information from the students, much like the interaction mechanisms of the other desktop simulators.

TeachLivE is the only full-immersion simulator included in the synthesis (see Vince Garland et al., 2012, 2016). In this mixed-reality environment, teachers interacted directly with a virtual student projected on a large screen, similar to how they would interact with students in an actual classroom. While a small percentage of common student behaviors were pre-programmed, Vince Garland et al. (2012, 2016) primarily relied on a human-in-the-loop, or interactor, who responded directly to the classroom teacher with a distinct student personality. Verbal teacher instruction and student responses occurred in real time and were not limited by pre-programmed options, which allowed interactions to unfold authentically as they would in a real classroom.

Sessions. In all studies, teachers engaged in simulation activities individually. There was no indication that teachers observed one another or cotaught in the virtual
classrooms. Across studies, there was a large range in the number of sessions and total time teachers engaged with virtual students. Teachers participated in one to six sessions, with a total of 20 min to 4 hr of simulated practice. In the majority of studies (54%), researchers provided two or three practice sessions that were 10 to 30 min long. Vince Garland et al. (2012; 2016) provided up to six 15-min sessions. Therefore, most teachers interacted with virtual students for 20-90 total min. Outlier times were observed in Girod and Girod (2006), who required teachers to complete entire work sample lessons in two 2-hr sessions, and in Fisher et al. (2010), who provided a one-time professional development experience that lasted up to 3 hr but also included didactic instruction. Strang et al. (1991) did not report the number of min teachers interacted in the simulation and instead defined a session as 80 teaching events. (Note: data reported here reflect intervention sessions only. Baseline, orientation, and maintenance sessions are not included in the totals).

Instruction. In nine studies (82%), teachers received instruction before engaging with virtual students. However, researchers in only five of these studies focused their instruction on the target skills addressed in the simulation. Teachers who participated in Unreal Interviewing read an article outlining best practices in questioning students prior to interviewing a virtual student in the simulation (Brubacher et al., 2015). Similarly, teachers who engaged in the virtual workshop on the Concept Mastery Routine completed a software program about the instructional technique prior to practicing it in the simulation (Fisher et al., 2010). In both the Vince Garland et al. (2012, 2016) studies, teachers received written instructions for the steps in Discrete Trial Training or System-
of-Least Prompts, respectively, prior to their participation in TeachLivE, where they also received additional coaching during each session. Finally, Girod and Girod (2006) provided work sample methodology during university coursework to teachers who participated in the Cook School District Simulation, in addition to a detailed orientation to the simulation. In the remaining four studies, researchers instructed teachers on the mechanics of the virtual classroom, but did not explicitly indicate that teachers received instruction on the target skills. However, because these teachers were recruited from undergraduate education programs, or participated in the study as part of a course, it is assumed they were taught related instructional and behavior management techniques during coursework.

*Feedback.* Teachers in all studies received some form of visual feedback about their performance during the simulation. In seven studies (64%), teachers received a computer-generated report of their performance at the end of the session. In two studies (18%), researchers embedded ongoing visual feedback into the simulation and teachers did not receive any further feedback at the end of the session (Brubacher et al., 2015; Fisher et al., 2010). Vince Garland et al. (2012, 2016) recorded which skills the teacher performed correctly or incorrectly, and in one study (Vince Garland et al., 2016) showed the teacher at the end of the practice session, and in the other study (Vince Garland et al., 2012) presented it to the teacher at the beginning of the next intervention session. In addition to visual feedback, in four studies (36%) researchers delivered verbal feedback, debriefing, or coaching (Badiee & Kaufman, 2014; Girod & Girod, 2006; Vince Garland et al., 2012, 2016).
Dependent variables. As summarized in Table 3, the majority of research teams (73%) reported teacher performance during the simulation. In three studies (27%), researchers reported performance outside of the simulation, including during a mock interview (Brubacher et al., 2015), or during actual classroom teaching (Fisher et al., 2010; Girod & Girod, 2006). The dependent variables targeted in ten studies (91%) were instructional and/or behavior management skills. Instructional skills were addressed in seven of these studies, and included delivering lessons (Girod & Girod, 2006; Strang et al., 1991), aligning appropriate tasks to students based on their profiles (Badiee & Kaufman, 2014), addressing student errors (Strang et al., 1989), or implementing specific instructional techniques, such as Discrete Trial Training (Vince Garland et al., 2012), System-of-Least Prompts (Vince Garland et al., 2016), and the Concept Mastery Routine (Fisher et al., 2010). Behavior management skills were addressed in five studies (46%), all of which involved responding to misbehavior in the Curry Simulation (Murphy et al., 1987; Strang et al., 1986, 1987, 1989, 1991). Brubacher et al. (2015) is the only research team that reported a nonpedagogical outcome. They measured the type and number of questions teachers ask suspected victims of abuse, a skill not typical in daily classroom instruction, but important in some circumstances.

Study design. In eight studies (73%), researchers conducted group design investigations. In five of these studies, researchers used a one-group pretest/posttest design to investigate the primary research questions (Badiee & Kaufman, 2014; Brubacher et al., 2015; Murphy et al., 1987; Strang et al., 1989; Strang et al., 1991), and in two studies researchers utilized subgroups to address secondary research questions
## Table 3

**Virtual Simulation Research Designs, Dependent Variable Characteristics, Experimental Outcomes, and Summary of Effects**

<table>
<thead>
<tr>
<th>Study</th>
<th>Design</th>
<th>Dependent variable</th>
<th>Experimental outcomes</th>
<th>Summary of effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Badiee &amp; Kaufman (2014)</td>
<td>One Group: Pretest, posttest 1, &amp; posttest 2</td>
<td>Task Appropriateness: Computer-generated mean score for alignment between assigned tasks and student characteristics during simulation</td>
<td>Decrease from pretest to posttest 1 ($p = 0.950$); $d = -0.02$ Increase from posttest 1 to posttest 2 ($p &lt; .001$) $d = 0.96$</td>
<td>Mixed/Neutral Effects on alignment of tasks to student needs</td>
</tr>
<tr>
<td>Brubacher et al. (2015)</td>
<td>One Group: Pretest &amp; posttest</td>
<td>Teacher Questions: Number and type of questions asked of suspected victims of abuse during mock interviews before and after simulation</td>
<td>Decrease in total number of questions asked ($p &lt; .001$); $d = 0.76$ Increases in open-ended questions ($p &lt; .001$); $ #: d = 1.34$; $%: d = 1.99$ Decreases in specific questions ($p &lt; .001$); $ #: d = 0.99$; $ %: d = 1.00$ Decreases in leading questions ($p &lt; .001$); $ #: d = 1.602$; $%: d = 2.30$</td>
<td>Positive Effects on the number and type of questions</td>
</tr>
<tr>
<td>Fisher et al. (2010)</td>
<td>Single Subject: Two multiple baselines across teachers in different groups: 1. Virtual workshop 2. Traditional workshop</td>
<td>Implementation of Concept Mastery Routine (CMR): Percentage of implementation for 39 possible instructional behaviors Student mastery of concept: Percentage of correct responses to 22 short answer questions (Pretest &amp; posttest)</td>
<td>Virtual Workshop: 4 positive effects 0 non-effects 0 negative effects Traditional Workshop: 4 positive effects 0 non-effects 0 negative effects Increases in student performance from pretest to posttest: Virtual Workshop ($p = .002$); $d = 5.96$ Traditional workshop ($p = .004$); $d = 3.88$</td>
<td>Positive Effects on teacher delivery of CMR in real classroom Positive Effects on real students' mastery of concepts</td>
</tr>
<tr>
<td>Girod &amp; Girod (2006)</td>
<td>Two Groups: (Quasi-experimental) 1. Simulation 2. No simulation Pretest &amp; posttest</td>
<td>Work Samples: Mean work sample scores before and after simulation Lesson Delivery: Mean scores on each component of a lesson plan delivered in a real classroom before and after simulation</td>
<td>Between group differences in work sample scores from pre-test to post-test ($p &lt; .05$); $d = 0.44$ Between group differences in components of lesson evaluation from pre-test to post-test: Lesson planning ($p &lt; .05$); $d = 0.12$ Learning climate ($p &lt; .05$); $d = 0.29$ Lesson implementation: $d = -0.03$ Evaluation of achievement: $d = -0.05$ Impact on student learning: $d = -0.06$</td>
<td>Positive Effects on work sample scores Mixed/Neutral Effects on lesson delivery in a real classroom</td>
</tr>
</tbody>
</table>

*(table continues)*
<table>
<thead>
<tr>
<th>Study</th>
<th>Design</th>
<th>Dependent variable</th>
<th>Experimental outcomes</th>
<th>Summary of effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Murphy et al. (1987)</td>
<td>One Group: Pretest, posttest, &amp; follow-up</td>
<td>Responses to Misbehavior: Mean number of appropriate or inappropriate responses to 10 student talkouts during simulation</td>
<td>Increase in appropriate responses to talkouts ($p &lt; .001$): Pretest to posttest: $d = .86$ Pretest to follow-up: $d = .86$ Decrease in inappropriate responses to talkouts ($p &lt; .001$): Pretest to posttest: $d = 1.46$ Pretest to follow-up: $d = 1.59$ Decrease in time spent addressing talkouts ($p &lt; .0001$) Pretest to posttest: $d = 1.32$ Pretest to follow-up: $d = 1.48$</td>
<td>Positive Effects on response type and time spent addressing misbehavior</td>
</tr>
<tr>
<td></td>
<td>Two Subgroups: (Random) 1. Misbehavior Record (performance feedback) 2. Misbehavior Profile (strategy instruction)</td>
<td>Mean number of min spent addressing 10 talkouts during simulation</td>
<td>Feedback: No differences between feedback groups ($p &gt; .05$); $d = .37$ Maintenance: No differences between posttest and follow-up ($ps &gt; .05$)</td>
<td>Positive Effects on behavior management</td>
</tr>
<tr>
<td>Strang et al. (1986)</td>
<td>Two Groups: (Random) 1. Feedback 2. No Feedback</td>
<td>Pretest &amp; posttest Responses to Misbehavior: Mean percent of inappropriate responses to 10 student talkouts during simulation</td>
<td>Decrease in inappropriate responses to talkouts: Feedback group ($p &lt; .05$); $d = 2.61$ No Feedback ($p &lt; .05$); $d = 0.92$ Feedback group had greater decrease in percent of inappropriate responses to student talkouts ($p &lt; .01$); $d = 1.14$</td>
<td>Positive Effects on behavior management</td>
</tr>
<tr>
<td>Strang et al. (1987)</td>
<td>Two Groups: (Random) 1. High feedback 2. Low feedback</td>
<td>Pretest, posttest 1, posttest 2, &amp; follow-up Responses to Misbehavior: Mean number of attempts to successfully eliminate 8 talkouts during simulation</td>
<td>Inservice teachers required fewer attempts and less time per attempt to eliminate talkouts at baseline ($p &lt; .01$): Number of attempts: $d = 2.20$ Secs per attempt: $d = 1.73$ All preservice teachers decreased the number of attempts and time required to eliminate talkouts from pretest to posttest 2: High feedback: Number of attempts ($p &lt; .01$): $d = 1.40$ Secs per attempt ($p &lt; .05$): $d = 0.84$ Low feedback: Number of attempts ($p &lt; .01$): $d = 1.12$ Secs per attempt ($p &lt; .05$): $d = 0.94$ Feedback: No differences between feedback groups ($ps &gt; .05$) Maintenance: No differences between posttest and follow-up ($ps &gt; .05$)</td>
<td>Positive Effects on behavior management</td>
</tr>
<tr>
<td></td>
<td>Additional Group: Inservice teachers (comparison at baseline)</td>
<td>Mean secs spent per attempt addressing talkouts during simulation</td>
<td></td>
<td>Positive Effects on behavior management</td>
</tr>
</tbody>
</table>

*table continues*
<table>
<thead>
<tr>
<th>Study</th>
<th>Design</th>
<th>Dependent variable</th>
<th>Experimental outcomes</th>
<th>Summary of effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strang et al. (1989)</td>
<td>One Group: Pretest, posttest, &amp; follow-up</td>
<td>Responses to misspelling: Mean number of possible responses to misspellings, divided by the number of opportunities</td>
<td>Changes (in acceptable direction) for 4 of 6 responses to misspellings ($p &lt; .01$): $d$ values from 0.41 to 0.53</td>
<td>Positive Effects on instructional strategies</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Responses to misbehavior: Mean number of possible responses to misbehavior, divided the number of opportunities</td>
<td>Decrease in ignoring misspellings ($p &lt; .01$): $d = 0.43$</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Changes (in acceptable direction) for 4 of 4 responses to misbehavior ($p &lt; .01$): $d$ values from 0.26 to 0.39</td>
<td>Mixed/Neutral Effects on behavior management</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Decrease in ineffectively responding to misbehavior ($p &lt; .01$): $d = 0.56$</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Maintenance: No differences between posttest and follow-up scores ($p &gt; .05$)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Strang et al. (1991)</td>
<td>One Group: Pretest &amp; posttest Two Subgroups: (Random) 1. Answered 4 questions about lesson content after pretest 2. Answered 4 questions about misbehaviors after pretest</td>
<td>Instruction: Metric not defined, but included 2 possible effective strategies and 2 possible ineffective strategies</td>
<td>All teachers showed changes (in acceptable direction) for 4 of 4 teaching strategies ($p &lt; .05$ or &lt; .01): $d$ values from 0.63 to 1.42</td>
<td>Positive Effects on instructional strategies</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Behavior management: Metric not defined, but included 2 possible effective strategies and 2 possible ineffective strategies</td>
<td>Group 1 increased their use of think time (strategy 1) more than Group 2 ($p &lt; .01$): $d = 1.86$</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>All teachers showed changes (in acceptable direction) for 2 of 4 behavior strategies ($p &lt; .01$): $d$ values from 1.73 to 2.14</td>
<td>Mixed/Neutral Effects on behavior management</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Teachers did not show improvements in 2 of 4 behavior management strategies ($p &gt; .05$): $d$ values from .29 to .41</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Group 1 increased their use of touching students or asking them to recite relevant rules (strategy 2) more than Group 2 ($p &lt; .01$): $d = 1.05$</td>
<td></td>
</tr>
<tr>
<td>Vince Garland et al. (2012)</td>
<td>Single Subject: Multiple baseline across teachers</td>
<td>Discrete Trial Training: Percent proficiency with Discrete Trial steps while teaching in simulation</td>
<td>3 positive effects, 0 non-effects, 0 negative effects (Note: 1 participant not scored for having only one intervention data point)</td>
<td>Positive Effects on application of Discrete Trial Training</td>
</tr>
<tr>
<td>Vince Garland et al. (2012)</td>
<td>Single Subject: Multiple baseline across teachers</td>
<td>System-of-Least Prompts: Percent proficiency with steps of System-of-Least Prompts while teaching in simulation</td>
<td>6 positive effects, 0 non-effects, 0 negative effects</td>
<td>Positive Effects on application of System-of-Least Prompts</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Maintenance: All 6 teachers maintained proficiency from intervention to follow-up</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
(Murphy et al., 1987; Strang et al., 1991). In three of the group design studies, researchers compared the performance of two groups, including comparisons between teachers who received simulated practice and teachers who did not (Girod & Girod, 2006), as well as comparisons between teachers who received different types of feedback (Strang et al., 1986, 1989). Researchers of the remaining three studies (27%) conducted single subject experiments, all of which were multiple baselines across teachers (Fisher et al., 2010; Vince Garland et al., 2012, 2016).

**Experimental outcomes.** A strong pattern of positive effects on teacher performance was observed across studies as a result of simulated practice. Specific outcomes and summary of effects are displayed in Table 3, and synthesized below.

**Primary outcomes.** Most positive effects in this review reflect teacher improvement from pretest to posttest (or baseline to intervention) while interacting with virtual students. For dependent variables in the academic domain, the ratio of positive effects to mixed/neutral effects is 5:1, with no observed negative effects. Specifically, positive effects were observed for teachers’ work sample scores (Girod & Girod, 2006), delivery of appropriate instructional strategies (Strang et al., 1991), responses to student misspellings (Strang et al., 1989), implementation of Discrete Trial Training (Vince Garland et al., 2012), and System-of-Least Prompts (Vince Garland et al., 2016). Badiee and Kaufman (2014) produced mixed/neutral effects on teachers’ ability to individualize instructional tasks to students based on the profile of their strengths and needs.

For dependent variables in the behavior domain, the ratio of positive effects to mixed/neutral effects is 3:2, with no observed negative effects. All positive effects were
related to teacher responses to misbehavior, including increasing appropriate responses and decreasing inappropriate responses to misbehaving students (Murphy et al., 1987; Strang et al., 1986), decreasing the number of attempts to successfully eliminate talkouts (Strang et al., 1987), or decreasing the amount of instructional time spent addressing misbehavior (Murphy et al., 1987; Strang et al., 1986). The mixed/neutral effects from Strang et al. (1989) are due to effect sizes for teachers’ responses to misbehavior that ranged from $d = 0.26$ to $0.39$ ($d = .40$ is the minimum effect size required for a positive effects rating). In Strang et al. (1991) the mixed/neutral summary of effects reflects teacher improvement on some behavior management strategies, but not on others.

Finally, Brubacher et al. (2015) observed positive effects on teacher questioning techniques to be used in situations of expected abuse, which was the only non-pedagogical outcome in the review.

**Feedback.** The effect of feedback on teacher outcomes was investigated in three of the Curry Simulation studies. Strang et al. (1986) verified that teachers who received a copy of computer-generated data about their performance decreased inappropriate responses to misbehavior significantly more than teachers who received no feedback ($d = 1.14$). Next, Murphy et al. (1987) investigated the impact of two types of feedback and found no significant differences between teachers who received individualized feedback about their performance in the simulation (i.e., misbehavior record) and teachers who received generic information about effective behavior management strategies (i.e., misbehavior profile). Similarly, Strang et al. (1987) observed no significant differences in teacher responses to misbehavior based on the timing of feedback. Teachers who viewed
feedback about their performance immediately after posttest 1 and prior to posttest 2 (i.e., high feedback group) demonstrated similar outcomes to teachers who only viewed feedback about their performance prior to posttest 2 (i.e., low feedback group).

**Maintenance.** In four studies (36%), researchers measured teacher performance in the simulation after intervention. In three of the group design studies, one follow-up session was conducted between 13 and 159 days following the last intervention session (Murphy et al. 1987; Strang et al. 1987; Strang et al., 1989) during which researchers discovered no significant differences between performance at posttest and follow-up. In one single subject study, Vince Garland et al. (2016) conducted 2-4 weekly maintenance sessions per participant, during which all participants maintained proficiency with System-of-Least Prompts. Data from these four studies indicate that teachers maintained their performance of the target skills over time.

**Generalization.** Only two research teams (18%) investigated generalization of teaching skills in authentic classroom settings. Fisher et al. (2010) found positive effects on teachers’ application of the Concept Mastery Routine with real students, and also discovered significant increases in student performance before and after instruction. However, no significant differences in student outcomes were discovered between those instructed by teachers who received the virtual workshop and those instructed by teachers who received the traditional (i.e., face-to-face) workshop. Girod and Girod (2006) found mixed/neutral effects on lesson delivery in the real classroom. Teachers who practiced in the Cook School District Simulation earned significantly higher scores from pretest to posttest on some components of lesson delivery (lesson planning and learning climate).
than teachers who did not practice in the virtual classroom, but there were no differences in their scores on other lesson components (lesson implementation, evaluation of achievement, and impact on student learning).

**Methodological quality.** The studies in this synthesis were designed and conducted with varying degrees of methodological rigor. In general, the higher the methodological quality of a study, the more confidence can be placed in the results. Methodological quality indicator (MQI) scores for each study are presented in Table 4, with information about which indicators were satisfactorily met and which indicators were not met or not reported. Overall, the MQI scores ranged from 50-100% ($M = 75\%$).

The wide range of MQI scores may be partially indicative of the shift in methodological standards over time. The majority of studies with MQI scores below 70% were published 25-30 years ago. In these studies, researchers provided less detail about the participants, interventionists, and validity of the measures. Of the studies published in the last decade, only one received an MQI score below 70% (Badiee & Kaufman, 2014). In all studies, researchers satisfied the majority of indicators for the outcome measure/dependent variables and also met all indicators for data analysis. However, in five of the studies (46%), researchers used a one-group pretest/posttest research design to evaluate the effects of the virtual simulation, which is subject to a number of threats to internal validity, including maturation, history, and testing (Campbell & Stanley, 1963; Shadish, Cook, & Campbell, 2002). In addition, in seven studies attrition was an applicable indicator, and six researchers failed to report it. Finally, the most common omission among all studies was assessing and reporting implementation fidelity, with
Table 4

**Virtual Simulation Methodological Quality Indicator Scores**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0. Context and setting</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.1. Critical features of context/setting</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.0. Participants</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.1. Participant demographics</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.2. Disability or risk status</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.0. Intervention agent</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.1. Background of intervention agent and/or simulation</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.2. Qualifications of Interventionist</td>
<td>n/a</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td>4.0. Intervention</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.1. Intervention procedures</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>4.2. Description of materials</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>5.0. Implementation fidelity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.1. Implementation fidelity of treatment adherence</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>5.2. Dosage or exposure fidelity</td>
<td>n/a</td>
<td>+</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>5.3. If fidelity is assessed, it is done regularly and for each setting, interventionist etc.</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>6.0. Internal validity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.1. Systematic manipulation of the independent variable.</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.2. Baseline or control/comparison condition</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>6.3. Participants in baseline/control have no access to intervention</td>
<td>n/a</td>
<td>n/a</td>
<td>+</td>
<td>+</td>
<td>n/a</td>
<td>+</td>
<td>+</td>
<td>n/a</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

*(table continues)*
### Methodological quality indicators

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>6.4. Assignment to groups</td>
<td>n/a</td>
<td>n/a</td>
<td>+</td>
<td>-</td>
<td>n/a</td>
<td>-</td>
<td>-</td>
<td>n/a</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.5. ≥ 3 demonstrations of effect at 3 different times</td>
<td>-</td>
<td></td>
<td>n/a</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.6. Predictable baseline with ≥ 3 data points</td>
<td>+</td>
<td></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.7. Controls for threats to internal validity</td>
<td>+</td>
<td></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.8. Overall attrition is low</td>
<td>n/a</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>n/a</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.9. Differential attrition is low or controlled for</td>
<td>n/a</td>
<td>n/a</td>
<td>-</td>
<td>-</td>
<td>n/a</td>
<td>-</td>
<td>-</td>
<td>n/a</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7.0. Outcome measures/dependent variables</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7.1. Socially important outcomes</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7.2. Clear description of measurement of the DV</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7.3. Effects reported for all measures</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7.4. Appropriate frequency and timing of measures</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7.5. Adequate internal reliability</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7.6. Adequate evidence of validity</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8.0 Data analysis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8.1. Appropriate data analysis techniques for comparing two or more groups</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8.2. Clear graph showing data from each phase of study</td>
<td>+</td>
<td></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8.3. Effect sizes provided, or adequate data for calculations</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Total:                          | 10/16                     | 17/19                    | 18/25                 | 16/21                 | 13/17                 | 13/21                 | 14/21                 | 12/17                 | 10/20                 | 21/21                      | 21/21                      |
| % of Indicators Met:           | 63%                       | 90%                      | 72%                   | 73%                   | 77%                   | 62%                   | 67%                   | 71%                   | 50%                   | 100%                       | 100%                       |

**Note.** + = indicator met; - = indicator not met or not reported; not applicable to study. n/a = indicator not applicable to design type; n/a = indicator
Vince Garland et al. (2012, 2016) being the only research teams to satisfy the requirements of that indicator.

**Conclusions**

Virtual classrooms have been used for approximately 30 years in teacher preparation programs as a safe environment to practice a variety of target skills. The sophistication of the teacher and student interactions within simulations has increased in the last decade, as well as the available research investigating classroom simulations. This evolution is especially evident in the development of TeachLivE, a full-immersion simulator.

Overall, practicing in a virtual classroom has positive effects on teacher instruction and behavior management. The strongest pattern of positive effects was observed for instructional skills, although there are promising outcomes for behavior management skills as well. Importantly, no negative effects were discovered in any of the studies. There is limited evidence indicating the importance of providing feedback to teachers as part of simulated practice sessions, but it is not clear if some formats and timing of feedback are more efficacious than others in improving teacher performance. The available, albeit limited, evidence also suggests that positive effects may maintain over time and that certain components of simulated practice may transfer to real classroom teaching. These findings build an encouraging foundation for future investigations of virtual classrooms as teacher educators continue to explore how simulated practice opportunities can be best utilized to prepare teachers for the demands
of classroom instruction.

Though encouraging, the results of this synthesis should be considered in the context of some limitations. First, it is possible not all qualified studies were identified for the review, although systematic searching and screening procedures were employed. Also, the synthesis does not represent all simulations used in teacher preparation programs and professional development. For example, Second Life is a common type of virtual technology used with teachers, but no studies qualified for the review because the dependent variables were related to content knowledge or teacher attitudes and not performance outcomes. This was a common reason for excluding investigations of other types of simulations as well.

Additionally, there are several limitations of the virtual simulation literature base. First, methodological quality of several studies in this synthesis is a concern, especially in the areas of internal validity and implementation fidelity. In particular, there are multiple limitations of the one-group pretest/posttest design, which was utilized as the primary research design in almost half of the studies. Moreover, studies conducted in the early era of classroom simulations earned comparatively low MQI scores. Thus, the outcomes from these studies should be interpreted cautiously. Second, relatively few studies in the literature base focused on teacher performance outcomes, which is reflected in the small number of studies that qualified for this review. Although building teachers’ content knowledge and influencing attitudes about teaching may be important goals of a preparation program, it is critical to verify that teachers can proficiently deliver essential teaching skills that have a strong evidence base for improving student outcomes. Third,
most researchers provided no verbal feedback to accompany visual records of performance, and relatively few researchers provided participants with explicit instruction about the target skills prior to interacting with the virtual students (or it was not reported). Fourth, researchers typically targeted a limited range of skills and provided very few simulated practice sessions for the teachers. This precluded teachers’ opportunity to interleave multiple skill sets and deliver them in progressively more complex teaching scenarios that are representative of actual classrooms. Fifth, there is a conspicuous lack of generalization data to real classrooms before and after simulated practice. If the purpose of virtual simulations is to prepare teachers for the demands of instructing real students, then researchers need to verify that teacher performance not only improves in the virtual environment, but also transfers to authentic classroom settings.

The research study proposed in the subsequent chapter addresses many of these concerns and contributes new findings to the existing literature base. Methodological rigor was carefully considered, and piloted in two previous studies. Teachers received didactic instruction on essential target skills prior to sessions in the simulator. Multiple practice opportunities were provided within and across intervention sessions, followed by the delivery of a well-defined feedback protocol. Additionally, simulated sessions were conducted with a small group of teachers, maximizing their opportunity to observe others teach and to engage in discussion about the target skills. The complexity of virtual teaching scenarios was increased strategically, so that teachers interleaved a range of target skills from both the academic and behavior domain. Finally, outcomes were
measured in the virtual environment and in classrooms to verify the extent to which performance during simulated practice generalized and maintained in authentic classroom settings.
CHAPTER III

METHODS

Participants

Teacher participants were recruited from the 2015 cohort enrolled in the Mild/Moderate Alternative Teacher Preparation (M/M ATP) program at Utah State University (USU). All teachers in the cohort (n = 35) had a bachelor’s degree in another field and were placed in part- or full-time special education positions on letters of authorization while pursuing their licensure. These placements included any grade level or special education setting in the public or charter school system and served as the teachers’ required field experience. The M/M ATP program covered a geographic region of approximately 15 school districts and 8 charter schools.

All teachers in the M/M ATP cohort completed a required weeklong teaching course at the end of July 2015. During this course, teachers learned basic instructional and management skills to help them get started in their classrooms. Topics included classroom expectations and procedures, lesson plan development, core curriculum standards, and conducting small group reading and math sessions. On the last day of the course, they attended a 30-min orientation in the TeachLivE classroom. The primary researcher briefly introduced the simulated classroom and gave each teacher a short turn to interact with the virtual students. The purpose of this experience was to familiarize the teachers with the technology and give them a sense of the students’ personalities and the dynamics of the simulated classroom. (See Appendix A for orientation session instructions). At the end of the orientation, the primary researcher explained the
upcoming research study and distributed a sign-up sheet for teachers to indicate if they were interested in participating.

Priority for participation was given to volunteers teaching small group language arts to upper elementary or middle school students for at least part of the school day. The rationale was to maximize the extent to which the teachers’ classrooms matched the virtual environment. Teachers were excluded from participating if they worked as a paraeducator or taught only in inclusion co-teaching settings. Additionally, teachers were excluded from the study if they were not admitted into the M/M ATP program as part of the 2015 cohort, if they were not enrolled in the TeachLivE course, or if they indicated they were unwilling to participate in the study.

Nineteen M/M ATP teachers volunteered to be considered for the study. Of these volunteers, 12 met the criteria of teaching language arts in upper elementary or middle school classrooms. The primary researcher contacted the six volunteers in upper elementary classrooms first, because they were more likely to instruct small groups of students, which more closely aligned to the context of the virtual classroom. The primary researcher phoned each of these teachers to confirm their interest and answer their questions about the requirements of participating in the study. Two teachers were eliminated from consideration based on these phone calls. One teacher indicated she would not be enrolling in the M/M ATP program in the fall, and another teacher did not return the researcher’s phone calls. The remaining four teachers were identified as the research study participants. Approximately one week prior to the beginning of the study, the primary researcher met with each participant individually in her classroom to explain
the letter of information (see Appendix B), answer any further questions about the study, and obtain informed consent. All four participants were involved in the study from beginning to end, with no attrition. At the completion of the study, participants received a $300 stipend for their participation. Their performance in the study did not impact their grade in the TeachLivE course or their field placement evaluations.

Participants were all Caucasian females with diverse educational backgrounds. None had formally studied education before entering the M/M ATP program. Their ages at the beginning of the study ranged from 24-40 years old. All had previous experience as a paraeducator, with their time of service ranging from 2 months to 2 years. Additionally, Marie had experience as a substitute teacher. Participants were teaching in four different schools across three school districts and one charter school. Background information for each participant is displayed in Table 5.

Settings

TeachLivE Classroom

The TeachLivE classroom is located at the USU Salt Lake Extension at the Granite Education Center. The room consisted of a large screen that displayed the virtual

Table 5

<table>
<thead>
<tr>
<th>Participant</th>
<th>Gender</th>
<th>Age</th>
<th>Ethnicity</th>
<th>Undergraduate degree</th>
<th>Prior experience</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lisa</td>
<td>Female</td>
<td>24</td>
<td>Caucasian</td>
<td>Psychology</td>
<td>Paraeducator</td>
<td>1 year</td>
</tr>
<tr>
<td>Grayce</td>
<td>Female</td>
<td>27</td>
<td>Caucasian</td>
<td>Parks &amp; recreation</td>
<td>Paraeducator</td>
<td>2 months</td>
</tr>
<tr>
<td>Deanna</td>
<td>Female</td>
<td>40</td>
<td>Caucasian</td>
<td>Botany</td>
<td>Paraeducator</td>
<td>2 years</td>
</tr>
<tr>
<td>Marie</td>
<td>Female</td>
<td>38</td>
<td>Caucasian</td>
<td>Business</td>
<td>Substitute teacher paraeducator</td>
<td>7 years</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1 year</td>
</tr>
</tbody>
</table>
students, speakers to broadcast student responses, small pedestals to represent the physical placement of student desks, and an Xbox Kinect system to detect the physical position and movements of the participants.

The interactor played the part of the virtual students from a studio in Florida. The virtual students, presented on a large screen, could move their heads, upper bodies and arms, use gestures, and make eye contact. The interactor viewed and heard the participants’ behavior through a camera and responded by speaking into a microphone and triggering student movements by pushing buttons on a hand-held console. Interactor responses were transmitted back to the classroom in real time. It was not apparent to the participants that an adult actor was playing the roles of the students. The interactor’s identity and location were withheld to maintain the believability of the teaching experience.

The TeachLivE classroom consisted of five middle school students, with two students seated in the front row and three students in the second row. They were designed to represent a range of personalities and academic levels. Figure 4 shows the five middle school students seated in the TeachLivE classroom.

**Participants’ Classrooms**

Participants’ field placement classrooms served as the generalization setting for the study. All classes were resource language arts groups, with 3-6 students with mild/moderate special needs. Participants taught fourth- through sixth-grade students. Lisa, Grayce, and Deanna taught from scripted programs assigned by their district (see Table 6). Marie developed instructional materials to teach vocabulary from the reading text
The participants received varying amounts of training prior to implementing language arts instruction with their students. Lisa indicated that a colleague provided her with a 10-min explanation about the Phonics for Reading program. Grayce stated that her coach modeled a brief lesson from Reading Mastery III on two different occasions. Deanna reported that her district supervisor provided an informal training session on the Next Steps program. Finally, Marie received no inservice used in the general education class.
training on designing and implementing vocabulary instruction, but indicated that she structured her lesson delivery to be similar to the format she was practicing in TeachLivE.

**Lesson Materials**

The primary researcher provided lesson content and materials for the participants to deliver during all TeachLivE sessions and assessments. Three lessons were provided for TeachLivE sessions (i.e., Training Lessons A, B, and C) and 15 lessons were provided for TeachLivE comprehensive assessment sessions (Assessment Lessons 1-15). All lessons focused on developing vocabulary, and were adapted from the Bold Moves, Grade 6 materials from the STORYtown program (STORYtown, n.d.). Each lesson included 6-10 target vocabulary words with student-friendly definitions. The lessons were not scripted, but participants were given example teaching formats with strategies for initiating student responses. Participants were also given picture cards for each of the target vocabulary words, which displayed the definition of the vocabulary word, a relevant picture, and a sentence using the word (see Appendix C for sample vocabulary lesson materials). Additionally, they had access to the STORYtown story for each lesson, which used all vocabulary words for a given lesson in context. Participants were told to use these stories for additional context for teaching the vocabulary words but that they should not to spend instructional time reading the story with the virtual students. All lesson materials were available on the Canvas course website.
Dependent Variables and Measurement System

Target Skills

The essential skills targeted in this study were: error correction, specific praise, and praise around.

**Error correction.** Academic errors signaled an opportunity for participants to use an error correction procedure. Opportunities to deliver error corrections occurred when a student or group of students produced an incorrect answer or expressed confusion about the content when directed to answer academic questions. A complete error correction procedure included a model, test, and delayed test. Each step of the error correction sequence is described below (see Appendix D for complete definitions, examples, and non-examples of each step in the error correction sequence).

**Model.** The first step in the error correction sequence was demonstrating the correct answer to the question. Models also needed to include a meaningful prompt. For example, the participant was taught to say “p-r-e says ‘pre’” instead of saying “the answer is ‘pre.’” Additionally, negative comments, such as “No,” or “not quite,” even if they preceded a correct model, counted as an incorrect model.

**Test.** A test was delivered immediately following the model, and was defined as restating the original prompt to the student(s) who made the initial error.

**Delayed test.** After delivering a model and a test, the teacher needed to retest the student(s) who made the initial error. The delayed test had to be delivered after one or more intervening responses. The intervening response needed to be a different question than the test/delayed test question and could be directed to the same student who made
the error or to a different student or group of students. If the participant addressed misbehavior after delivering the test, then the participant’s reaction to the misbehavior could serve as the intervening response as long as the elapsed time was 10 s or more.

**Specific praise.** Teacher praise was defined as positive teacher statements and gestures referring to student work or behavior. Specific praise statements contained a praise word or phrase and also made direct reference to an academic skill or behavior. For example, specific praise statements for academic responses included, “Great example of the word ‘luxury,’” or “Thanks for sounding out each sound in the word and then saying it smoothly.” Specific praise statements for behavior included, “I appreciate Kevin opening his book quietly,” or “thank you all for following along with your eyes up here.” In contrast, general praise statements contained only the praise word or phrase without direct reference to the academic skill or behavior, such as, “Good job class,” “That’s correct,” “What a good student you are Ed,” or “I appreciate you.” Incorrect praise statements included comments that were negatively stated (e.g., “Thank you for not tapping your pen on the desk), or those that followed misbehavior (e.g., saying “I appreciate you sitting quietly” to a student who was talking out). Statements that did not reference academic work or behavior (e.g., “you look so nice today”), or corrective statements (e.g., “sit down right now,” or “I need you to follow along”) were not defined as either type of praise. (See Appendix D for complete definitions, examples, and non-examples of each type of praise).

**Praise around.** Participants were taught to implement the praise around procedure in response to student misbehavior. A complete praise around sequence
included ignoring a misbehaving student and delivering specific praise to another student who was exhibiting the desired behavior, and then delivering specific praise to the target student when he or she exhibited the desired behavior. Each step of the praise around sequence is described below, beginning with a description of what constituted an opportunity to use praise around (see Appendix D for complete definitions, examples, and non-examples of a praise around opportunity and each praise around step).

**Praise around opportunity.** Persistent or recurring student misbehavior signaled an opportunity to use praise around, and was defined as a disruptive off-task behavior that lasted for 5 s or longer, or lasted less than 5 s but recurred within 30 s. Examples of disruptive off-task behaviors that required praise around included talking out, engaging in a side conversation with another student, interrupting the teacher during instruction, producing disruptive noises with objects or mouth, verbally harassing another student or an adult, throwing a tantrum, screaming, or talking on a cell phone. Subtle off-task behavior counted as a praise around opportunity only if the participant responded to it, either by initiating a praise around sequence or by delivering a corrective statement. If a participant ignored or did not notice subtle misbehavior, and it did not escalate into disruptive off-task behavior, participants were not accountable to deliver the praise around procedure. Examples of subtle off-task behaviors included sitting inappropriately in the chair, putting head on desk, diverting eyes from the teacher or task, leaving one’s chair during instruction, or quietly texting on a cell phone.

**Praise around step 1.** The first step of the praise around procedure was to praise another student, or students, for exhibiting the desired behavior. The praise statement had
to be specific and identify the behavior that was incompatible with the problem behavior (e.g., if Sean was tapping his pen on the desk, an appropriate step 1 praise statement could be, “Thank you Maria for keeping your materials flat and quiet during the lesson.”) General praise statements, corrective statements, or statements that did not identify the behavior that was incompatible with the problem behavior did not count as correct delivery of step 1.

**Praise around step 2.** The next step of the praise around procedure was to praise the student who was previously misbehaving once they exhibited the desired behavior. Again, this statement needed to be specific and identify the behavior that was incompatible with the problem behavior. General praise statements, corrective statements, or statements that did not identify the incompatible behavior were not counted as correct delivery of step 2.

**Assessments**

Participants’ proficiency with the target skills were measured during three weekly assessments and reported as: (1) the TeachLivE training assessment, (2) the TeachLivE comprehensive assessment, and (3) the classroom generalization assessment. All weekly assessments were recorded using a Flip Video Camera. The characteristics of each assessment are summarized in Table 7 and described below.

**TeachLivE training assessment.** The purpose of the TeachLivE training assessment was to examine participants’ proficiency with the current skill targeted during the TeachLivE sessions, and to examine the extent to which they maintained proficiency with previously targeted skills (if applicable). Participants’ performance was scored
Table 7

*Characteristics of Each Weekly Assessment*

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>TeachLivE training</th>
<th>TeachLivE comprehensive</th>
<th>Classroom generalization</th>
</tr>
</thead>
<tbody>
<tr>
<td>When</td>
<td>Weekly sessions</td>
<td>Immediately after intervention session</td>
<td>1-7 days after intervention session</td>
</tr>
<tr>
<td>Where</td>
<td>TeachLivE middle school classroom</td>
<td>TeachLivE middle school classroom</td>
<td>Upper elementary language arts resource classrooms</td>
</tr>
<tr>
<td>Time</td>
<td>3 min</td>
<td>5 min</td>
<td>5-10 min</td>
</tr>
<tr>
<td>Lesson</td>
<td>Training Lesson A, B, or C</td>
<td>Assessment Lessons 1-15, randomly assigned</td>
<td>Lessons from designated programs or designed by the participant</td>
</tr>
<tr>
<td>Target skills</td>
<td>Cumulative target skills</td>
<td>All target skills</td>
<td>All target skills (opportunity dependent)</td>
</tr>
<tr>
<td>Present</td>
<td>All participants, primary researcher</td>
<td>One participant, primary researcher</td>
<td>Participant, students, sometimes paraeducators or other staff and students</td>
</tr>
</tbody>
</table>

During their final teaching turn in each weekly TeachLivE session. Each teacher’s final teaching turn was 3 min, and the primary researcher and other research study participants were present in the room. During the TeachLivE training assessment, the participants delivered the same training lesson they used during the previous teaching turns in the TeachLivE session.

**TeachLivE comprehensive assessment.** The purpose of the TeachLivE comprehensive assessment was to investigate the extent to which the participants demonstrated proficiency with past, current, and future skills targeted in the TeachLivE sessions. Weekly TeachLivE comprehensive assessments were held immediately following the TeachLivE sessions using the same simulated middle school classroom used during the TeachLivE sessions. The only individuals in the room during this 5-min assessment were the primary researcher and the participant being assessed. The order of
participants was rotated every week to prevent any bias based on order of assessment. Participants waited for their turn in the hallway or in an adjacent classroom. Participants had the opportunity to deliver all target skills during each TeachLivE comprehensive assessment, including skills that had not yet been taught during intervention.

During each TeachLivE comprehensive assessment, participants taught from the assessment lesson plan specified for that week. Assessment lesson plans followed the exact format and contained the same supplemental materials as the training lesson plans, but used different target vocabulary words. Prior to the beginning of the study, the primary researcher randomly assigned a different lesson to be taught during each TeachLivE comprehensive assessment session, and each lesson was available on the Canvas course website 5-6 days in advance. Requiring participants to deliver a different lesson plan from the lessons practiced during TeachLivE sessions prevented carryover effects that could have been due to lesson content instead of skill proficiency. Furthermore, requiring a different lesson for each TeachLivE comprehensive assessment session minimized potential practice effects related to lesson content.

The complexity of the teaching scenarios increased from the TeachLivE training assessment to the TeachLivE comprehensive assessment. During the TeachLivE training assessment, the primary focus was the skill targeted for that session. In addition, participants had opportunities to continue practicing previously targeted skills. In contrast, during TeachLivE comprehensive assessments, participants had opportunities to engage in all of the target skills, regardless of whether the skill was addressed in intervention (see Table 7). Specifically, each 5-min TeachLivE comprehensive
assessment session included approximately three to four academic errors and three to four misbehaviors, in addition to correct responses and appropriate behavior. The rationale for including all skills was twofold: (1) to assess to what extent participants demonstrated proficiency with skills from that day’s intervention session when they were required to implement them in a more complex teaching environment, and (2) to collect baseline data on skills that were not targeted yet. Also, the intent was to keep the complexity of the TeachLivE comprehensive assessments consistent across all phases of the study.

**Classroom generalization assessment.** The purpose of the classroom generalization assessment was to investigate the extent to which the participants generalized proficiency with the target skills when teaching students in their field placement classroom. Participants recorded themselves delivering a language arts lesson to the same instructional group each week. Classroom footage was captured from a Flip Video Camera positioned at the back of the classroom, with a view of the front of the participant and the backs of the students. The primary researcher provided instructions on the ideal placement of the camera when she met with participants in their classrooms to obtain informed consent prior to the beginning of the study. The participants recorded all classroom generalization sessions from the designated camera location without the researcher present in the room. Participants sent the footage to the primary researcher using USU’s big file transfer service (see Appendix E for video handling instructions provided for participants).

Typically, classroom generalization assessments were recorded 1-4 days following intervention ($M = 2$ days). In three instances, participants recorded themselves
7 days following intervention due to holidays or required professional development meetings, but the assessments still preceded the next intervention session (session 7 and 12 for Lisa, and session 12 for Deanna). Participants were instructed to record 5-10 min of active teaching, delivering instruction from the program required by their school or district. They submitted video footage that ranged from 5 min 2 s to 14 min 35 s (M = 7 min 34 s). The length of footage scored ranged from 5 min to 10 min (M = 6 min 33 s). Coding was initiated when the participant started teaching the students. Coding was temporarily suspended if instruction was interrupted by an announcement over the speaker or if another teacher came into the room to speak with the participant. Coding was terminated when the participant stopped teaching or after the first 10 min of active instruction.

**Coding Target Skills**

The recordings of each weekly assessment were scored using paper/pencil coding sheets (see Appendix F). A general description of the coding system for each target behavior is provided below. (Coding instructions that align to examples and non-examples of each target skill are provided in Appendix D).

**Error correction.** Each academic error was coded, including who made the error, what the error was, and what time the error occurred. Whenever participants initiated error correction, they were expected to deliver all steps in the error correction sequence. If the students did not respond on signal or if the participant indicated that they did not clearly hear the response, it did not count as an error. In these instances, the participant usually asked the question again, reminded the class that they needed to respond on
signal, or requested that the student repeat the answer more loudly. None of these responses counted as initiating an error correction procedure. Given an error correction opportunity, the observer coded the correct delivery of a model (+), the incorrect or incomplete delivery of a model (-), or failure to deliver a model (-), as well as the time the statement was delivered. The model was coded as not applicable (n/a) if the error occurred at the end of the observation and there was no time for the participant to deliver the model. The observer also coded the correct delivery of a test (+), the incorrect or incomplete delivery of a test (-), or failure to deliver a test (-), as well as the time the statement was delivered. The test was coded as not applicable (n/a) if the observation ended before the participant could deliver a test (e.g., time ran out during the model), or if the student answered the question immediately following the model without being prompted, and therefore the participant heard the student produce the correct answer without explicitly delivering the test. Additionally, the observer coded the correct delivery of a delayed test (+), the incorrect or incomplete delivery of a delayed test (-), or failure to deliver a delayed test (-), as well as the time the statement was delivered. The delayed test was coded as not applicable (n/a) if the observation ended less than a min after the delivery of the test and the participant did not deliver a delayed test within that timeframe, or if the student made another error when the teacher delivered the test. In that case, the student response was coded as a new error, and another complete error correction procedure was required. Correctly delivered error correction sequences are reported as the percentage of correctly implemented error correction steps per session. This was calculated by dividing the number of correctly delivered error correction steps...
by the total number of possible steps in the session, times 100.

At least five error correction steps were required to report a data point for a given session. If there were fewer than five possible error correction steps during an assessment in a given setting, the participant’s data were combined with their error correction data for the subsequent assessment session(s) in the same setting. At a phase change line, the data from sessions with fewer than five error correction steps were combined with the previous session(s). The rule for combining data across sessions was implemented to prevent extreme variability in the data paths that was an artifact of too few opportunities to demonstrate a target skill. This was especially relevant in participants’ classroom generalization settings where the frequency of student errors could not be programmed.

**Specific praise.** Correctly delivered specific praise statements (S+) were coded, as well as the time the statement was delivered. Incorrectly delivered specific praise statements were also coded (S-), and did not count in the praise rate calculation. Praise data are reported as the rate of specific praise statements per min, which was calculated by dividing the number of specific praise statements by the total number of min observed.

**Praise around.** Each persistent or recurring misbehavior was coded, including who misbehaved, the nature of the misbehavior, and what time the misbehavior occurred. If a praise around opportunity occurred, participants were accountable for delivering both steps in the praise around procedure. The observer then coded the correct delivery of step 1 (+), the incorrect or incomplete delivery of step 1(-), or omission of step 1 (-), as well as the time step 1 was delivered. Step 1 was coded as not applicable (n/a) if a student misbehaved at the end of the observation and the participant did not have time to deliver
step 1 before the time ended. Additionally, if the participant gave a corrective statement for a recurring misbehavior from the same student(s), step 1 was coded as not applicable (n/a) if the participant previously delivered step 1 and it was not effective at changing the behavior. The observer also coded the correct delivery of step 2 (+), the incorrect or incomplete delivery of step 2 (-), or omission of step 2 (-), as well as the time step 2 was delivered. Step 2 was coded as not applicable (n/a) if the student continued misbehaving so there was not an opportunity for the participant to deliver step 2, or if there was less than a min between the delivery of step 1 and the end of the observation and the participant did not deliver step 2 in that timeframe. Correctly delivered praise around sequences are reported as the percentage of correct praise around steps per session. This was calculated by dividing the number of correctly delivered praise around steps by the number of possible praise around steps in the session, times 100. Similar to error correction, at least five praise around steps were required to report a data point for a given session. If there were fewer than five possible praise around steps during an assessment session, data were combined across sessions in the same way as described for error correction.

**Social Validity Survey**

Participants completed an anonymous social validity survey at the conclusion of the intervention sessions (see Appendix G). The survey questions addressed the participants’ perception of the realism of the TeachLivE classrooms, students, interactions, and teaching scenarios (items 1-7; adapted from Hayes, Hardin, & Hughes, 2013). Participants were also asked to rate the relevance of each component of the
intervention and the extent to which they viewed the procedures as acceptable and valuable to their professional practice (items 8-29).

**Independent Variable**

The independent variable was a multicomponent package that included didactic instruction, practice and observation in the TeachLivE classroom, feedback, self-assessment, and written reflection. These components were delivered as part of a TeachLivE course completed by all teachers in M/M ATP program. The main independent variable was active practice in the TeachLivE classroom, and the other components were added to prepare the teachers to interact with the virtual students and to debrief their experiences in the simulation.

**Didactic Instruction**

Initial instruction on each target skill took place outside of the TeachLivE classroom to ensure that intervention sessions were reserved for active practice and feedback. The primary researcher and a teacher assistant recorded introductory instructional videos on each of the four target skills. Each video included a definition of the target skill, examples and non-examples of the target skill, and a short demonstration of how to deliver the target skill. The videos ranged in length from 7-12 min. (see Appendix H for the URLs for each training video). The primary researcher also created an instructional handout for each target skill to summarize the information in the video. Each handout included a definition of the target skill, the performance goal, as well as examples and non-examples of the skill (see Appendix I). After watching the video and
reading the handout, participants completed a short quiz assessing their understanding of the target skill (see Appendix J). The purpose of the quizzes was to increase participants’ accountability for interacting with the training materials prior to attending TeachLivE sessions.

**Practice and Observation in TeachLivE**

Prior to the beginning of each TeachLivE session, the primary researcher submitted session objectives to the interactor at the University of Central Florida (see Appendix K). On the session objectives form, the primary researcher identified the frequency of errors and/or misbehaviors the interactor needed to produce, based on the target skill for each session. The session objectives for each phase of the study specified opportunities for participants to demonstrate the current target skill as well as any previously targeted skills. For example, during baseline, the virtual students produced no academic errors and no misbehaviors. This allowed participants to become comfortable asking academic questions while holding all target skills in baseline. When intervention on error correction began, the virtual students were instructed to produce 2-3 academic errors but no misbehaviors during each participant’s turn. When specific praise was targeted, the virtual students continued to produce no misbehaviors during each teaching turn so there were numerous opportunities to praise student behaviors, and also produced 1-2 academic errors so the participants had intermittent occasions to practice error correction. Finally, when praise around was targeted, the virtual students produced 2-3 misbehaviors during each teaching turn to provide ample opportunities for participants to practice praise around, while still producing 1-2 academic errors so participants could
maintain proficiency with error correction.

At the beginning of each TeachLivE session, the primary researcher briefly conversed with the interactor to confirm her understanding of what was expected to provide fidelity for the session. In addition to clarifying the frequency of errors and/or misbehaviors needed for each turn, the interactor was also instructed to provide a range of simple to more intense errors and misbehaviors. If needed, the primary researcher also requested specific scenarios for certain participants, to provide individualized practice opportunities based on the type of academic errors and problem behaviors they were experiencing in their classrooms. For example, during her classroom generalization videos, Grayce sometimes corrected errors the first time a student answered a question incorrectly, but not if the error recurred. The interactor was asked to strategically recreate this scenario so Grayce could receive feedback on how to deal with this situation.

Another example was when Marie stated “thank you for not tapping your pencil,” when trying to praise around for pencil tapping during a classroom generalization assessment. During the next TeachLivE session, the interactor was told to provide Marie with an opportunity to address excessive pen clicking with one of the virtual students. This approach provided individualized practice opportunities in TeachLivE that aligned to scenarios observed in participants’ classrooms.

All participants in the research study attended each TeachLivE session together. Participants took turns teaching the virtual students and observed the other participants teach. There were three rounds of teaching, with each participant teaching once per round, for a total of three turns per participant per session.
The primary researcher served as the mentor for all TeachLivE sessions for the duration of the study. As needed, the mentor prompted the participants during the first two practice turns to improve their accuracy and proficiency with delivering the target skill. For example, the mentor might say, “turn that general praise statement to C.J. into a specific praise statement,” or “remember to come back to Sean for the delayed test,” or “it is time to praise Kevin now that he is sitting quietly.” The final teaching turn was used to assess each participant’s proficiency with the target skill(s) (i.e., the TeachLivE training assessment), thus there was no prompting on the final practice round to ensure it reflected each participant’s independent performance level. (See Appendix L for the complete TeachLivE session instructions, including directions for facilitating the three rounds of teaching turns).

**Feedback**

The mentor collected data on a feedback form for each participant’s performance during each teaching turn. The feedback form included data for the target skill specified for that session as well as data on skills practiced during previous sessions (see Appendix M). This cumulative feedback structure was implemented to encourage participants to interleave an expanding array of skills during TeachLivE sessions.

After each teaching turn, the mentor first asked the observing participants to verbalize, “what is one thing the teacher did well on (the current target skill).” Following feedback from the group, the mentor provided any additional positive or corrective feedback for the current target skill based on the collected data, as well as brief feedback for any prior target skills. The mentor then asked the participant who just taught, “What
would you like to improve on (current target skill).” During the third round this question was adapted to encourage the participant to explicitly make a connection between the current target skill and their own classroom. This prompt was individualized to each participant’s experience during the intervention session, and included questions such as, “what did you practice tonight that you will take to your students tomorrow,” or “how does what you’ve practiced and observed tonight with (current target skill) apply to the students you’re working with in your own classroom,” or “what do you need to implement (on current target skill) with your own students that will benefit them the most.” (See Appendix L for instructions for facilitating the feedback following each teaching turn).

**Participation Self-Assessment**

At the conclusion of each TeachLivE session, participants completed a participation self-assessment form, which is how participation points were assigned for the TeachLivE course. Participants rated their performance for the session based on punctuality, preparation, effort, and professionalism (see Appendix N).

**Written Reflection**

Following all intervention sessions on a target skill, the participants engaged in an online discussion via the Canvas course website. There were four online discussions required in the course, one for OTR and one for each target skill. Discussions included all other teachers in the M/M ATP program who also attended TeachLivE sessions but were not part of the research study. All M/M ATP teachers were prompted to reflect on their
successes and challenges with delivering the target skill during TeachLivE sessions, and to make relevant connections to their own classrooms. Teachers posted their initial reflection and commented on at least three of their peers’ posts as well. All participants in the research study participated fully in all online discussions, with the following exceptions: Marie failed to participate in the discussion on specific praise, and Deanna failed to participate in the discussion on praise around. Therefore, they were asked to submit a brief reflection paper responding to the same prompt given for the online discussion. They both completed this makeup assignment to compensate for missing the online discussion.

**Experimental Design and Procedures**

The research questions for this study were addressed using a multiple baseline design across target skills, replicated with four participants. The study included a baseline phase, an intervention phase for each essential target skill, and a maintenance phase. TeachLivE sessions were conducted weekly, on Monday evenings, with the following exceptions: Session 2 was held on a Tuesday evening to compensate for a Monday holiday, and two weeks elapsed between Session 11 and 12 because of a regional internet outage that prohibited connecting to the virtual classroom.

**Baseline**

The baseline phase served two purposes: (1) it provided an opportunity to collect baseline data for the primary target skills of error correction, specific praise, and praise around during the TeachLivE comprehensive assessment sessions; and (2) it was used to
provide instruction on opportunities to respond (OTRs). OTRs were not included as a primary dependent variable in the study because pilot data showed that participants’ delivery of OTRs, including rate and question type, was substantially impacted by the instructional program they were required to implement and less directly influenced by the training they received during TeachLivE sessions. Therefore, OTRs were taught during the baseline phase, not as a primary target skill, but in order to promote teachers’ proficiency asking a variety of questions. Teacher delivery of OTRs created an instructional foundation for the virtual students to answer questions, which allowed them to eventually make academic errors so participants could practice error correction. Also, correct OTRs increased opportunities for the participants to praise academic responding and on-task behavior. Thus, providing instruction on OTRs set the context for participants’ delivery of the other target skills. OTRs and question types are defined below, including a description of the basic coding procedures.

**OTR.** An OTR occurred when the participant asked an academic question and indicated if it was directed to the group or an individual. Examples of correct OTRs included, “what is the prefix in the first word, everyone?” or “please read the second word in the list, Maria.” Incorrect OTRs occurred when the participant delivered prompts as questions, such as “will you read the next sentence, C.J.?” or “Can you give me an example of the word communal, Sean?” Additionally, when the teacher did not direct questions to a specific student or group of students it did not count as a correct OTR, such as “who knows how to use the word apparent in a sentence?”

OTRs were divided into two questions types. Basic knowledge questions included
reading or repeating a word, sentence, or definition. Application questions included using a word in an original sentence, identifying or providing an example or non-example of a word, sharing a personal experience related to a word, or demonstrating the meaning of a word. (See Appendix D for complete definitions, examples and non-examples of OTRs, including examples of each question type).

On all assessments, the observer tallied the number of correctly delivered OTRs for each question type (B+ or A+) and incorrectly delivered OTRs for each question type (B- or A-). The overall rate of teacher-directed OTRs is reported, as well as the rate for each question type, and was calculated by dividing the number of correctly delivered OTRs by the total number of min observed.

**Baseline/OTR sessions.** Prior to the first session on OTRs, the participants accessed the OTR training video and handout (see Appendices H and I). After viewing these training resources, participants submitted their responses to the quiz (see Appendix J). In addition, participants were told to review the materials for Training Lesson A and bring them to the first session. The video, handout, and lesson materials were available on the Canvas course website six days prior to the first session.

At the beginning of each session focused on OTRs, the mentor invited all participants into the TeachLivE classroom. The mentor stated that OTR was the target skill for that session and then she read the definition of an OTR from the handout and reminded participants that the performance goal was at least 4 OTRs per min. In addition, she prompted participants to first establish basic knowledge (about 50% of questioning) and then move strategically to application questions (about 50% of questioning).
The participants then took turns teaching the virtual students in three rounds, during which they all taught Training Lesson A. During the first two rounds, each participant taught for 2 min and during the third round each participant taught for 3 min. The rationale for this structure was to provide shorter turns when scaffolding would likely be required, and build in a longer turn at the end of the session when participants would likely be more proficient with the target skill. The participants were allowed to “pause” the classroom during the first two teaching turns if they needed to ask a question or receive on-the-spot mentoring. The time for their turn was also paused and resumed when they began teaching again. No mentoring was provided during the third teaching turn.

The mentor collected data during each teaching turn (see Appendix M) and then provided feedback on OTRs only, using the feedback protocol described previously. Following the third round of teaching turns and feedback, the mentor asked if there were any remaining questions about OTRs and then ended the session with concise concluding remarks about the session. Participants completed the participation self-assessment form, received a copy of their feedback form, and then exited the classroom. Baseline sessions targeting OTRs lasted an average of 79 min (71-90 min).

**Intervention on Error Correction, Specific Praise and Praise Around**

Similar to baseline sessions, participants were required to view the training video and handout for the upcoming target skill and then submit their responses to the quiz prior to the first TeachLivE session on a new skill. The primary researcher unlocked the
Canvas module with the video and handout for each new target skill three to five days in advance. This ensured that participants had ample time to access the materials, but also guaranteed that didactic instruction did not precede the first intervention session by more than a few days. Videos and handouts for all future skills remained locked so training materials were only available at the appropriate intervention phase. Once training materials were unlocked, the participants had access to those materials for the duration of the study. They could re-watch the training video or download the handout again at any point.

In addition, participants were told which of the three training lessons they would deliver during each TeachLivE session, which were also available on the Canvas course website. All participants used the same training lesson for all teaching turns in a given session. Standardizing lesson content across participants helped control variability that might have resulted from teaching different content and lesson designs. Participants taught the same lesson for several consecutive TeachLivE sessions so they would become familiar with the content and focus primarily on developing fluency with the target skills. The primary researcher instructed the participants to move to a new training lesson when necessary to prevent boredom with the materials. They were not asked to move to a new training lesson at the same time they moved to a new target skill, except when they moved from OTRs to error correction. At that phase change, they were asked to move from Training Lesson A to Training Lesson B because it was more believable that students would start making errors when presented with a new list of vocabulary words.

Similar to baseline sessions, the mentor invited all participants into the TeachLivE
classroom at the beginning of each intervention session. The mentor reminded the participants of the target skill for that session and read or paraphrased the definition as well as the performance goal provided on the handout. Additionally, she briefly mentioned the target skills they practiced in previous sessions and explained that the goal was to maintain proficiency with prior skills while building fluency on the new skill.

The participants then took turns teaching the virtual students in three rounds. Similar to baseline sessions, the first two turns lasted 2 min and participants could pause the classroom if needed. The third turn lasted 3 min. The mentor collected data and facilitated feedback for each teaching turn. During intervention sessions on error correction, the mentor provided focused feedback on error correction and secondarily addressed OTRs. During intervention sessions on specific praise, the mentor provided focused feedback on specific praise with brief feedback on error correction. At this point, feedback on OTRs was discontinued. During intervention sessions on praise around, the mentor provided focused feedback on praise around while briefly addressing specific praise and error correction. Thus, as participants interleaved new target skills as the experimental phases unfolded, focused feedback was provided for the new target skill while cumulative feedback was provided for the previously targeted skills.

Following the third round of teaching turns and feedback, the mentor asked if there were any remaining questions about the target skill. She then provided concise concluding remarks about the target skill or session. Finally, she distributed the participation self-assessment forms for participants to fill out, as well as their feedback forms from the session. Intervention sessions lasted an average of 88 min (83-96 min).
Sessions targeting error correction lasted an average of 87 min (80-93 min), sessions targeting specific praise lasted an average of 89 min (87-90 min), and sessions targeting praise around lasted an average of 89 min (83-96 min).

**Phase Change Criteria**

Determining when to move to each new phase of the study was a complex decision-making process because it was based on data from four participants on multiple assessments (i.e., TeachLivE training assessment, TeachLivE comprehensive assessment, classroom generalization session). First, participants’ baseline data for the target skill needed to be stable or show a decreasing trend on the TeachLivE comprehensive assessment or the classroom generalization assessment. Second, the majority of participants needed to demonstrate proficiency with the current target skill on at least two of the three assessments. Proficiency for error correction and praise around was based on 80% correct steps or higher during focused feedback on two consecutive data points and an increasing or stable trend across data points. For specific praise, proficiency was defined as at least two or more specific praise statements per min during focused feedback for two consecutive data points with an increasing or stable trend.

**Maintenance**

**Classroom maintenance.** Classroom maintenance videos were collected between 29 and 44 days following the last TeachLivE session. The procedures for recording and submitting classroom maintenance videos were the same as described for classroom generalization videos, but participants did not attend any intervention sessions during this
time. All participants recorded in their classrooms twice a week for two weeks. Therefore, sessions during the maintenance phase represent biweekly recordings, while session data points in all other phases of the study represent weekly assessments. The length of videos submitted during the maintenance phase ranged from 5 min 12 s to 15 min 26 s ($M = 8$ min 11 s), and the length scored ranged from 5 min 3 s to 10 min ($M = 7$ min 17 s).

After Deanna and Marie submitted their fourth classroom maintenance video, the primary researcher gave brief feedback via email to remind them to deliver praise around steps with positively worded praise statements (e.g. “thank you for sitting quietly” instead of “thank you for not shouting out”). They both recorded additional maintenance videos the following week. Deanna submitted two additional videos, and Marie submitted one additional video. Importantly, this was the only time feedback was given for teacher performance in any setting other than the TeachLivE sessions.

**TeachLivE maintenance.** Access to TeachLivE was withheld from all participants during classroom maintenance data collection. Following submission of all classroom maintenance videos, participants came back to the TeachLivE classroom for one final comprehensive assessment session. The procedures for conducting this assessment, as well as the complexity of the session, were identical to those described for all other TeachLivE comprehensive assessments. Only one data point was collected in the TeachLivE comprehensive assessment setting because participants were scheduled to begin the spring semester TeachLivE course.
Reliability

The primary researcher scored 100% of the videos from the TeachLivE training assessments, TeachLivE comprehensive assessments, classroom generalization assessments, and maintenance assessments. The research assistant scored a total of 46% of all data points, which represented a cross-section of each experimental phase, assessment setting, and participant. At the conclusion of each week of baseline and intervention data collection, the primary researcher randomly selected one TeachLivE training assessment video, one TeachLivE comprehensive assessment video, and at least one classroom generalization assessment video for research assistant to double-code. The research assistant independently coded all target skills from each randomly assigned video. The research assistant was blind to the timing of phase changes and the order of skills targeted in intervention for the duration of the study. When intervention was complete, the researcher verified that at least one video for each participant was double-coded in each assessment setting for each experimental phase. Additional videos were randomly selected to satisfy this requirement and ensure the reported interobserver agreement (IOA) scores were representative of all phases, assessment settings, and participants.

At the conclusion of the maintenance data collection, one of the four videos for each participant from the first two weeks of maintenance was randomly selected for the research assistant to double code. Because additional videos were collected for Deanna’s and Marie’s classrooms following feedback, one of Deanna’s two additional videos was randomly selected for double coding, and Marie’s one additional video was assigned for
double coding.

Finally, if a video assigned for double-coding contained error correction or praise around data combined with another session or sessions because there were fewer than five possible steps, the additional videos were also assigned for double-coding for that skill only. Therefore, IOA scores for error correction and praise around were calculated by data point, and not necessarily by individual sessions.

IOA scores are summarized in Table 8 and reliability procedures for each target skill are described below.

**Reliability for Error Correction**

IOA was collected for 44% of all error correction data points across phases, assessment settings, and participants. It was scored point-by-point and calculated by dividing the number of agreements by the number of agreements plus disagreements, times 100 for each reliability session. An agreement was defined as both observers

<table>
<thead>
<tr>
<th>Assessment setting</th>
<th>Error correction</th>
<th>Specific praise</th>
<th>Praise around</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Opportunity</td>
<td>Steps</td>
<td>Occurrence</td>
</tr>
<tr>
<td>TeachLivE training</td>
<td>100% (100%-100%)</td>
<td>98% (89%-100%)</td>
<td>91% (75%-100%)</td>
</tr>
<tr>
<td>TeachLivE comprehensive</td>
<td>92% (67%-100%)</td>
<td>96% (83%-100%)</td>
<td>70% (0%-100%)</td>
</tr>
<tr>
<td>Classroom generalization</td>
<td>83% (40%-100%)</td>
<td>92% (73%-100%)</td>
<td>74% (0%-100%)</td>
</tr>
<tr>
<td>Total</td>
<td>89% (40%-100%)</td>
<td>95% (73%-100%)</td>
<td>75% (0%-100%)</td>
</tr>
</tbody>
</table>
coding the correct delivery or incorrect delivery of an error correction step within the same 5-s window, both coding the omission of the step, or both coding the step as not applicable. A disagreement was defined as the observers coding different responses for the same step, or coding the same response for the step outside the same 5-s window.

Reliability for error correction was determined using a two-stage process. After independently coding the assigned video, the primary researcher and research assistant reached consensus on student errors that counted as an error correction opportunity. This was done by first scoring all coded error correction opportunities for agreement or disagreement. An agreement was defined as both observers identifying the same academic error produced by the same student within a 5-s window. A disagreement was defined as one observer coding an error correction opportunity not coded by the other observer, or both observers coding the same error outside the 5-s window. All errors scored as an agreement remained on the coding sheet. All errors scored as a disagreement were discussed to reach consensus on if the situation warranted delivery of error correction. The primary researcher and research assistant referred to the definitions, examples, and non-examples to make this decision. Once they reached consensus on a particular instance, the error was then included or eliminated on the final coding sheet. Next, the primary researcher and research assistant independently coded each step for all error correction opportunities on the final coding sheet. The steps were then scored for agreements and disagreements. The mean percent agreement on opportunities to deliver error correction before reaching consensus was 89% (0%-100%). The mean percent agreement on error correction steps was 95% (73%-100%). (See Table 8 for the mean
IOA scores and ranges for error correction in each assessment setting.)

**Reliability for Specific Praise**

IOA was collected for 45% of all specific praise data points across phases, assessment settings, and participants. It was scored point-by-point and calculated by dividing the number of agreements by the number of agreements plus disagreements, times 100 for each reliability session. If both researchers agreed that no specific praise statements occurred during the assessment, the IOA score reported for that session was 100%. An agreement was defined as both observers coding delivery of a specific praise statement within the same 5-s window. A disagreement was defined as one observer coding correct delivery of a specific praise statement and the other observer not coding correct delivery of a specific praise statement within the same 5-s window. The mean percent agreement on point-by-point occurrence of specific praise statements was 75% (0%-100%). (See Table 8 for the mean IOA scores and ranges for specific praise in each assessment setting.)

The primary researcher and research assistant discussed all disagreements to reach consensus on if a statement should count as specific praise or not. They referred to the definitions, examples, and nonexamples to make this decision. If a disagreement occurred because the statement represented a new topography, or form of the statement, not encountered previously, a decision was made about the statement and a new rule was added to the list of definitions to be used when scoring future assessments.
Reliability for Praise Around

IOA was collected for 49% of all praise around data points across phases, assessment settings, and participants. It was scored point-by-point and calculated by dividing the number of agreements by the number of agreements plus disagreements, times 100 for each reliability session. An agreement was defined as both observers coding the correct delivery or incorrect delivery of a praise around step within the same 5-s window, both coding the omission of the step, or both coding the step as not applicable. A disagreement was defined as the observers identifying different codes for the step, or coding the same step outside of the same 5-s window.

Reliability for praise around was determined using the same two-stage process as described for error correction. The primary researcher and research assistant first reached consensus on each praise around opportunity, and then independently coded the steps for each opportunity. The mean percent agreement on opportunities to deliver praise around before reaching consensus was 79% (17%-100%). The mean percent agreement on praise around steps was 95% (67%-100%). (See Table 8 for the mean IOA scores and ranges for praise around in each assessment setting).

Building Consensus on Classroom Generalization Assessments

The type of errors and misbehaviors observed in the classroom generalization settings differed from those observed in the TeachLivE classroom due to several factors, some of which included the instructional program, age of students, and physical layout of the classroom. Therefore, both the primary researcher and research assistant scored all
classroom generalization assessment videos for at least the first five sessions to build consensus on which error types would count as error correction opportunities and which misbehaviors would count as praise around opportunities in each classroom. The criteria for adequate consensus was at least 75% agreement on error correction or praise around opportunities, and at least 75% agreement on steps for the target skill, for at least the last two consecutive data points. After five sessions, the consensus-building criteria were met for error correction in all classrooms except Deanna’s, which required one additional session to reach criteria. For praise around, the criteria for consensus were reached after five sessions in Grayce’s and Marie’s classrooms, and in Lisa’s classroom after session 6. The consensus-building process for praise around in Deanna’s classroom continued through session 12, partly because there were relatively few praise around opportunities in her classroom setting. Thus, both the primary researcher and research assistant scored all classroom generalization videos for all participants for sessions 1-5, and scored Lisa’s and Deanna’s classroom generalization videos for session 6 to continue building consensus on error correction and/or praise around. Starting with session 7, one of the four participants’ classroom generalization videos was randomly selected for the research assistant to double-code for all skills. Additionally, until session 12, Deanna’s classroom generalization video was assigned for the research assistant to double-code for praise around.

**Reliability Training**

The research assistant was trained to code each target skill using a three-step process. First, the primary researcher reviewed the definition, examples, and
nonexamples of the target skill with the research assistant. Second, the primary researcher and research assistant practiced coding the target skill side-by-side using video footage of teachers delivering lessons in the TeachLivE classroom from a previous study. Next, the primary researcher and research assistant independently coded the target skill from additional practice videos. When they reached an IOA score of 85% or higher for three consecutive videos, they moved on to the next target skill and repeated the process. If IOA fell below 80% for any target skill on a video, the primary researcher and research assistant met to discuss the disagreements, and to review the definition, examples, and non-examples for the target skill.

Treatment Fidelity

All TeachLivE sessions were recorded using a Flip Video Camera positioned at the back of the room so a percentage of them could be scored later to determine the extent to which the independent variable was administered as specified. A trained TeachLivE instructor served as the primary coder for treatment fidelity. She regularly conducted TeachLivE sessions for the M/M ATP program, but was not involved in the research study sessions. She was assigned to score the first session of each experimental phase, as well as one other randomly selected session from each phase, to ensure fidelity within and across phases. In total, she scored eight of the 14 TeachLivE sessions, or 57% of the total sessions.

A 69-point checklist was developed to score each of the crucial intervention components for each teaching turn for each participant (see Appendix O for the complete
treatment fidelity checklist). A component was scored as “1” if it did occur with fidelity and “0” if it did not occur with fidelity, or if it was unclear to the observer if it occurred. The fidelity of the interactor’s performance was also scored as part of the checklist to assess the extent to which she produced the appropriate level of difficulty for each phase of the study (i.e., she was scored for producing errors, misbehaviors, neither, or both as specified for the session). There was also a space for comments if the observer had questions or concerns to discuss with the primary researcher. The treatment fidelity score is reported as a percentage of correctly implemented intervention components. This score was calculated by dividing the number of correctly implemented components by the number of possible components for the session, times 100. The average treatment fidelity score across all scored sessions was 99.8% (98.6%-100%).

The research assistant independently double-coded 50% of the videos scored by the TeachLivE instructor to report an IOA score for treatment fidelity. One of the two scored videos from each phase was randomly selected for the research assistant to score. The research assistant scored treatment fidelity at the conclusion of the study so she stayed blind to the order of the target skills and the timing of phase changes. Treatment fidelity IOA was scored point by point, with the number of agreements divided by the number of agreements plus disagreements, times 100. An agreement was defined as both observers marking “1” for the same component, or both observers marking “0” for the same component. A disagreement was defined as one observer marking “1” and the other observer marking “0” for the same component. The average treatment fidelity IOA score was 99.6% (98.6%-100%).
CHAPTER IV
RESULTS

The research study results are reported in three parts. First, participants’ overall OTR rates are reported as a context for delivering the primary target skills of error correction, specific praise, and praise around. Next, each participant’s proficiency with the primary target skills is discussed across assessment settings and experimental phases. Third, outcomes of the social validity survey are addressed.

Opportunity to Respond Results

Initially, participants received instruction on OTRs as a foundation for error correction, specific praise, and praise around. Primary instruction and feedback on OTRs was provided during sessions 1-3. Cumulative feedback was given during intervention on error correction, and discontinued starting with session 7. IOA for OTR was collected for 43% of all assessment videos across phases, assessment settings, and participants. The primary researcher and research assistant independently coded the total frequency of OTRs in a session, including the frequency of each question type. IOA scores were calculated for total OTRs, basic knowledge questions, and application questions, by dividing the smaller frequency count by the larger frequency count, and multiplying by 100. The mean IOA for total OTRs was 92% (67%-100%), the mean IOA for basic knowledge questions was 88% (40%-100%), and the mean IOA for application questions was 89% (0%-100%).

Data on each participant’s average OTR rates, ranges, and standard deviations in
each setting, disaggregated by question type, are displayed in Table 9. Overall, all participants delivered a higher average rate of OTRs on the TeachLivE training assessment than on the TeachLivE comprehensive assessment. On average, Lisa, Grayce, and Deanna delivered a balance of basic knowledge questions and application questions, and Marie delivered more basic knowledge questions than application questions.

Table 9

Participants’ OTR Rates per Min in Each Assessment Setting

<table>
<thead>
<tr>
<th>Participant</th>
<th>Assessment setting</th>
<th>Basic</th>
<th></th>
<th>Application</th>
<th></th>
<th>Total OTR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>Lisa</td>
<td>TeachLivE training</td>
<td>1.8</td>
<td>0.6</td>
<td>0.7</td>
<td>3.0</td>
<td>2.0</td>
</tr>
<tr>
<td></td>
<td>TeachLivE comprehensive</td>
<td>1.8</td>
<td>0.6</td>
<td>1.2</td>
<td>2.4</td>
<td>1.6</td>
</tr>
<tr>
<td></td>
<td>Classroom generalization</td>
<td>11.0</td>
<td>3.7</td>
<td>1.3</td>
<td>16.8</td>
<td>0.1</td>
</tr>
<tr>
<td>Grayce</td>
<td>TeachLivE training</td>
<td>2.5</td>
<td>1.1</td>
<td>1.3</td>
<td>5.0</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td>TeachLivE comprehensive</td>
<td>1.8</td>
<td>0.8</td>
<td>0.6</td>
<td>3.6</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td>Classroom generalization</td>
<td>3.6</td>
<td>1.2</td>
<td>1.6</td>
<td>6.0</td>
<td>0.3</td>
</tr>
<tr>
<td>Deanna</td>
<td>TeachLivE training</td>
<td>2.0</td>
<td>0.7</td>
<td>0.7</td>
<td>3.0</td>
<td>2.0</td>
</tr>
<tr>
<td></td>
<td>TeachLivE comprehensive</td>
<td>1.8</td>
<td>0.6</td>
<td>1.0</td>
<td>3.4</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td>Classroom generalization</td>
<td>6.3</td>
<td>3.2</td>
<td>0.7</td>
<td>12.8</td>
<td>0.2</td>
</tr>
<tr>
<td>Marie</td>
<td>TeachLivE training</td>
<td>2.4</td>
<td>1.5</td>
<td>0.3</td>
<td>6.0</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td>TeachLivE comprehensive</td>
<td>2.4</td>
<td>1.0</td>
<td>0.2</td>
<td>3.8</td>
<td>1.2</td>
</tr>
<tr>
<td></td>
<td>Classroom generalization</td>
<td>1.1</td>
<td>0.9</td>
<td>0.0</td>
<td>3.8</td>
<td>1.9</td>
</tr>
</tbody>
</table>
However, there was a great deal of overlap in the distribution of basic and application questions throughout the study.

In their classrooms, Lisa and Deanna averaged a higher overall rate of OTRs than in the TeachLivE assessment settings. They almost exclusively delivered basic knowledge questions as part of the Phonics for Reading and Next Steps scripted programs respectively (e.g., tasks included reading sounds, words, and simple sentences). Grayce’s overall rate of OTRs in her classroom was similar to her average rate in the TeachLivE training setting, but with a wider distribution of rates. Similar to Lisa and Deanna, Grayce averaged a higher rate of basic knowledge questions in her classroom than on both TeachLivE assessments. (Grayce taught Reading Mastery III, which included word reading, passage reading, and basic recall comprehension questions).

In contrast to the other participants, Marie averaged a lower overall rate of OTRs on the classroom generalization assessment than on both TeachLivE assessments. In addition, she averaged a lower rate of basic knowledge questions in her classroom than on both TeachLivE assessments. She was also the only participant who averaged a higher rate of application questions than basic knowledge questions in her classroom, and was also the only participant who did not teach from a scripted program. Instead, she selected target words from the general education history textbook and structured vocabulary activities similar to those practiced during TeachLivE sessions (e.g., most tasks required students to provide original sentences and examples of the selected vocabulary words).
Error Correction, Specific Praise, and Praise Around Results

The first three research questions addressed the effects of TeachLivE intervention on the primary target skills of error correction, specific praise, and praise around. Specifically, they addressed the extent to which participants would: (1) develop proficiency with the essential target skills during TeachLivE sessions; (2) demonstrate proficiency with the essential target skills when teaching in a more complex TeachLivE assessment session and maintain proficiency approximately 1.5 months following intervention; (3) generalize essential target skills to an authentic classroom environment and maintain proficiency approximately one month following intervention sessions. Overall, participants developed high levels of proficiency with the target skills during TeachLivE sessions, continued to demonstrate proficiency with the target skills during the TeachLivE assessment sessions immediately following intervention, and generalized delivery of the target skills to authentic classroom environments. When performance feedback was withdrawn, participants maintained proficiency with the target skills at varying levels in both the virtual environment and classroom settings. Data for each participant’s performance on the essential target skills are reported below.

Lisa’s Results

**TeachLivE performance.** The first leg of the multiple baseline graph in Figure 5 illustrates Lisa’s results for error correction. Her correct delivery of error correction steps during baseline on the TeachLivE comprehensive assessment ranged from 47%-70% ($M = 56\%$) with a downward trend in performance. During focused feedback, her scores on
Figure 5. Lisa’s percentage of correct steps for error correction, specific praise rate per min, and percentage of correct steps for praise around on the TeachLivE training assessment, TeachLivE comprehensive assessment, and classroom generalization assessment.
the TeachLivE comprehensive assessment were higher than her baseline scores ($M = 87\%$) with an upward trend, and she reached 100\% proficiency by the third session. On the TeachLivE training assessment she consistently demonstrated 100\% correct steps. During cumulative feedback, Lisa’s proficiency on the TeachLivE training assessment was more variable than during focused feedback ($M = 93\% \text{ [67\%-100\%]}$). However, she improved her delivery of error correction on the TeachLivE comprehensive assessment, averaging 99\% correct steps, with only one data point below 100\% (91\%). Lisa demonstrated an average increase of 40\% correct error correction steps from baseline to intervention (focused feedback and cumulative feedback combined) on the TeachLivE comprehensive assessment. When assessed on the TeachLivE comprehensive assessment 44 days following intervention, Lisa delivered 89\% correct error correction steps.

The second leg of the multiple baseline graph in Figure 5 shows Lisa’s results for specific praise. During baseline, she delivered stable, low rates of specific praise on the TeachLivE comprehensive assessment ($M = 0.2 \text{ [0.0-0.6]}$). When focused feedback was initiated, Lisa substantially increased her rate of specific praise on both the TeachLivE comprehensive assessment ($M = 3.1 \text{ [2.2-3.6]}$), with a similar rate on the TeachLivE training assessment ($M = 3.2 \text{ [2.7-3.7]}$). During cumulative feedback, she continued to improve her rate of specific praise on the TeachLivE comprehensive assessment with a data path trending upward, and a mean rate of 3.9 (3.6-4.8) statements per min, and she maintained her rate of specific praise on the TeachLivE training assessment ($M = 3.3 \text{ [2.3-3.7]}$). Overall, Lisa improved her delivery of specific praise from baseline to intervention on the TeachLivE comprehensive assessment by an average of 3.4 specific
praise statements per min. Lisa maintained her delivery of specific praise at a rate of 3.2 statements per min on the TeachLivE comprehensive assessment 44 days later.

Lisa’s delivery of praise around is illustrated in the third leg of her multiple baseline graphs (see Figure 5). During baseline on the TeachLivE comprehensive assessment, she delivered no praise around steps for the first five data points. After she received intervention on specific praise, she started applying the skill to praise around opportunities, but did so with variability (0%-63%). After focused feedback on praise around, Lisa consistently delivered 100% correct praise around steps on both the TeachLivE training assessment and TeachLivE comprehensive assessment, with the exception of the second comprehensive assessment session when she delivered 89% correct praise around steps. On average, she delivered 88% more correct praise around steps during intervention than baseline on the TeachLivE comprehensive assessment. She maintained proficiency during the follow up session, with 93% correctly delivered praise around steps.

**Classroom generalization.** Overall, Lisa’s performance in her classroom correlated with her performance in TeachLivE (see Figure 5). Her mean baseline performance on error correction in her classroom was similar to her performance during baseline on the TeachLivE comprehensive assessment. She delivered a mean of 56% (38%-67%) correct error correction steps, with a distinct downward trend. When intervention was initiated in TeachLivE, she displayed a substantial level change with stable performance during both focused feedback ($M = 97$ [91%-100%]) and cumulative feedback ($M = 94$ [88%-100%]), representing an average increase of 39% from baseline
to intervention. Lisa was assessed in her classroom four times during the maintenance phase, between 30 and 38 days following intervention. She maintained proficiency with error correction, demonstrating a mean score of 98% (94%-100%) correct error correction steps.

Similar to her performance on error correction, Lisa generalized delivery of specific praise to the classroom at levels that correlated to those observed in TeachLivE (see Figure 5). During baseline, she delivered almost no specific praise in her classroom ($M = 0.1 [0.0-0.4]$). When intervention was introduced in TeachLivE, she demonstrated a substantial and stable level change ($M = 2.8 [2.3-3.5]$) during focused feedback, with an increasing trend during cumulative feedback ($M = 3.1 [2.3-4.2]$). Lisa delivered an average of 3.1 specific praise statements per min across the entire TeachLivE intervention phase, which was an average of 3.0 specific praise statements per min more than observed during baseline. During the four maintenance assessment sessions in her classroom, Lisa’s specific praise rate was similar to her intervention specific praise rate ($M = 3.2 [2.8-4.2]$) with a slight downward trend.

Lisa generalized proficiency with praise around to her own classroom, with some variations in the data patterns compared to her performance in the virtual environment (see Figure 5). Similar to her baseline performance in TeachLivE, Lisa demonstrated no praise around steps prior to TeachLivE intervention on specific praise. On the second session of TeachLivE intervention on specific praise, Lisa’s performance on praise around improved to 20% in the classroom. (Note: Lisa did not meet baseline phase change criteria because her performance on praise around improved prior to intervention
on both the TeachLivE comprehensive assessment and classroom generalization assessment. When focused feedback was initiated in TeachLivE, she delivered an average of 90% (67%-100%) correct praise around steps with an upward trend from the first data point to the last data point, and stable proficiency for the last three sessions of the phase (92%-100%). During the maintenance assessments, Lisa delivered 100% correct praise around steps for all three data points, representing stable proficiency and a 10% average increase from intervention to maintenance.

Grayce’s Results

**TeachLivE performance.** The first leg of the multiple baseline graph in Figure 6 shows Grayce’s results for error correction. Grayce’s performance during baseline sessions on the TeachLivE comprehensive assessment was similar to Lisa’s performance, but at a lower performance level. Grayce delivered a range of 25%-47% correct steps (\(M = 35\)) with an overall downward trend. When focused feedback on error correction was introduced in intervention, she increased her performance to a mean level of 78% (73%-83%) correct steps, with a slight upward trend. Her mean performance on the TeachLivE training assessment was even higher (\(M = 88\)), though more variable (75%-100%). Grayce continued to improve her performance on error correction during cumulative feedback on both the TeachLivE comprehensive assessment (\(M = 93\% \ [64\%-100\%]\)) and the TeachLivE training assessment (\(M = 91\% \ [67\%-100\%]\)). Overall, Grayce demonstrated an average increase of 54% from baseline to intervention on the TeachLivE comprehensive assessment. Grayce attempted all error correction steps during the follow-up assessment 44 days later, but correctly delivered 67% of the steps.
Figure 6. Grayce’s percentage of correct steps for error correction, specific praise rate per min, and percentage of correct steps for praise around on the TeachLivE training assessment, TeachLivE comprehensive assessment, and classroom generalization assessment.
Grayce’s performance on specific praise is illustrated in the second leg of the multiple baseline graph in Figure 6. Similar to Lisa, Grayce delivered low rates of specific praise during baseline on the TeachLivE comprehensive assessment (M = 0.4 [0.0-0.8]), with a trend descending to a rate of 0.0 for the final two data points. When focused feedback was introduced, Grayce increased her delivery of specific praise on the TeachLivE comprehensive assessment to an average rate of 2.5 statements per min (2.0-3.4), with an upward trend, and delivered comparatively higher rates on the TeachLivE training assessment (M = 3.6 [3.0-4.3]). During cumulative feedback, Grayce continued to improve her specific praise rate and demonstrate an increasing trend on the TeachLivE comprehensive assessment (M = 3.3 [2.6-4.0]). In contrast, Grayce showed a decreasing specific praise rate for the first four sessions on the TeachLivE training assessment and then increased her delivery of specific praise to earlier levels during the final session (M = 3.6 [2.7-4.3]). On average, Grayce improved her delivery of specific praise by 2.6 statements per min from baseline to intervention on the TeachLivE comprehensive assessment. She maintained specific praise at a rate of 3.0 statements per min on the TeachLivE comprehensive assessment 44 days later.

The third leg of the multiple baseline graph in Figure 6 shows Grayce’s proficiency with praise around. During baseline on the TeachLivE comprehensive assessment, she delivered an average of 13% correct praise around steps (0%-25%). In contrast to the pattern observed for Lisa, Grayce’s delivery of praise around steps did not appear to be impacted by intervention on specific praise. Following focused feedback during intervention, Grayce demonstrated a substantial level change on the TeachLivE
comprehensive assessment, delivering an average of 96% correct praise around steps. Similarly, she delivered a mean of 97% correct steps on the TeachLivE training assessment. The data paths for both assessments were almost identical. Grayce demonstrated 100% proficiency with praise around on all data points on both assessments, with the exception of the fourth intervention session where she delivered 80% correct steps on the TeachLivE comprehensive assessment and 86% correct steps on the TeachLivE training assessment. Overall, Grayce improved her average proficiency with praise around 83% from baseline to intervention on the TeachLivE comprehensive assessment. She maintained her performance during the follow up session, with 100% proficiency with praise around.

**Classroom generalization.** Similar to the generalization patterns observed for Lisa, Grayce’s classroom performance correlated with her performance in TeachLivE (see Figure 6), Grayce’s mean baseline performance on error correction in her classroom ($M = 33\%$) was similar to her performance on the TeachLivE comprehensive assessment, however, she delivered a wider range of correct error correction steps (0%-53%) with an upward trend. When the TeachLivE intervention was introduced, she demonstrated an upward trend during focused feedback and increased her mean level of proficiency with error correction to 84% (72%-90%). She continued to improve her performance in the classroom when receiving cumulative feedback in TeachLivE ($M = 89\%$ [83%-100%]). Grayce demonstrated an average improvement of 55% on correctly delivered error correction steps in her classroom from baseline to intervention, which was similar to the difference observed between the two phases on the TeachLivE comprehensive
Grayce was assessed in her classroom four times during the maintenance phase, from 29 to 38 days following intervention, and correctly delivered an average of 76% (71%-79%) error correction steps. Her performance initially dropped during this phase, but then trended upward.

Grayce delivered similar mean rates of specific praise in her classroom as she delivered in the corresponding phases in TeachLivE (see Figure 6). During baseline her mean rate of specific praise was 0.3 (0.0-0.6). When the TeachLivE intervention was introduced, she increased her specific praise rate to an average of 2.1 (1.7-2.4) and demonstrated a consistently increasing trend during focused feedback. She improved her performance during cumulative feedback, but demonstrated more variability than observed in TeachLivE. Grayce’s specific praise rate ranged from 1.9 to 6.5 statements per min, with a mean of 3.6, and a straight upward trend for the final three data points. On average, she improved her delivery of specific praise from baseline to intervention by 2.7 specific praise statements per min. During follow-up, Grayce maintained a mean rate of 3.1 (1.8-5.1) specific praise statements per min and demonstrated similar variability as observed in the prior phase.

Overall, Grayce generalized proficiency with praise around from TeachLivE to her classroom, but did so at lower mean rates and with more variability than observed in the virtual setting, as illustrated in Figure 6. During baseline, her mean delivery of praise around steps (7% [0-18%]) was lower than observed on the TeachLivE comprehensive assessment. She increased her average proficiency to 77% correct praise around steps when the TeachLivE intervention was introduced, with a range of 50% to 91% and an
ascending trend. Her average improvement from baseline to intervention on praise around was 70%. During the maintenance assessments in her classroom, Grayce improved her proficiency with praise around to an average of 87% correct steps (70%-100%), and demonstrated an upward performance trend ending with a data point of 100% proficiency.

Deanna’s Results

TeachLivE performance. The first leg of the multiple baseline graph in Figure 7 shows Deanna’s performance on error correction. Her correct delivery of error correction steps on the TeachLivE comprehensive assessment in baseline averaged 43%, with a range of 37%-48%. She increased her mean proficiency with error correction to 91% (80%-100%) correct steps during focused feedback, and similar to Lisa and Grayce, displayed an increasing trend. She demonstrated a similar average proficiency and trend on the TeachLivE training assessment ($M = 93\% [89\%-100\%]$). Deanna maintained proficiency with error correction for all sessions during cumulative feedback on both the TeachLivE comprehensive assessment ($M = 94\% [88\%-100\%]$) and TeachLivE training assessment ($M = 98\% [86\%-100\%]$). Overall, she improved her mean delivery of correct error correction steps 50% from baseline to intervention on the TeachLivE comprehensive assessment and demonstrated 100% proficiency with error correction 44 days later.

Deanna’s performance on specific praise is illustrated in the second leg of the multiple baseline graph in Figure 7. Similar to Lisa and Grayce, she delivered a low specific praise rate during baseline ($M = 0.4 [0.0-0.8]$) on the TeachLivE comprehensive
Figure 7. Deanna’s percentage of correct steps for error correction, specific praise rate per min, and percentage of correct steps for praise around on the TeachLivE training assessment, TeachLivE comprehensive assessment, and classroom generalization assessment.
assessment. She steadily increased this rate during focused feedback from 1.6 to 3.4 specific praise statements per min, with an average rate of 2.4 statements per min on the TeachLivE comprehensive assessment, which closely paralleled her performance on the TeachLivE training assessment ($M=3\ [2.0-3.7]$). She delivered even higher mean rates of specific praise with stable data paths during cumulative feedback, on both the TeachLivE comprehensive assessment ($M=3.8\ [3.4-4.0]$) and TeachLivE training assessment ($M=3.7\ [3.3-4.3]$). Deanna increased her overall specific praise rate by 2.9 statements per min from baseline to intervention on the TeachLivE comprehensive assessment. During the follow-up session 44 days after intervention, she maintained the same specific praise rate ($M=3.8$) as observed during cumulative feedback.

The third leg of the multiple baseline graph in Figure 7 shows Deanna’s proficiency with praise around. During baseline on the TeachLivE comprehensive assessments, she delivered no praise around steps for five of the six first data points, with 14% correct steps during session 3. Similar to Lisa, Deanna began to deliver some specific praise statements to address praise around opportunities once intervention began for specific praise and before intervention on praise around. This impacted the last three data points of baseline, which ranged from 25%-36% correct praise around steps and trended upward. Deanna demonstrated a substantial and stable level change during intervention on praise around. She consistently delivered 100% correct praise around steps on the TeachLivE comprehensive assessment, with only one data point below 100% on the TeachLivE training assessment (83% on the third data point). Her mean improvement from baseline to intervention on the TeachLivE comprehensive assessment
was 88%. Like Grayce, Deanna maintained 100% proficiency with praise around at follow-up.

**Classroom generalization.** Deanna generalized proficiency with error correction to the classroom with similar mean rates, ranges, and stable trends as observed in TeachLivE for corresponding phases (see Figure 7). This close correlation of Deanna’s error correction performance between the two environments closely resembles the pattern observed for Lisa. During baseline, Deanna delivered an average of 43% (30%-56%) correct error correction steps, with a steep downward trend. When the TeachLivE intervention was introduced, she immediately demonstrated a distinct level change within a stable range during focused feedback ($M = 94\%$ [83%-100%]), which she maintained during cumulative feedback ($M = 95\%$ [88%-100%]). These scores represent a mean increase of 51% from baseline to intervention, which closely corresponds to the increase observed from baseline to intervention in TeachLivE. Deanna was assessed six times in her classroom during the maintenance phase, between 29 and 44 days after the TeachLivE intervention was discontinued. Deanna maintained proficiency with error correction with a mean of 99% (92%-100%) correct steps, similar to the maintenance scores observed for Lisa.

As shown in Figure 7, Deanna delivered similar overall rates of specific praise in her classroom compared to TeachLivE. During baseline, Deanna delivered low and stable rates of specific praise, averaging 0.3 (0.0-0.4) statements per min. Once the TeachLivE intervention was introduced, she increased her specific praise rate to a mean of 3.8 (2.2-6.4) statements per min during focused feedback and 3.1 (2.4-4.2) statements per min
during cumulative feedback. Overall, Deanna’s rate of specific praise was more variable in the classroom than in TeachLivE, similar to the pattern observed for Grayce. Deanna’s mean specific praise rate of 3.4 (2.2-6.4) statements per min across focused feedback and cumulative feedback of the TeachLivE intervention phase represents an average improvement of 3.1 statements per min over baseline. Deanna maintained her rate of specific praise during the six follow-up assessments, delivering an average of 3.1 (2.6-4.0) statements per min.

Similar to Lisa and Grayce, Deanna generalized proficiency with praise around from the virtual setting to her real classroom (see Figure 7). During baseline, she delivered no praise around steps for all sessions. In contrast to her performance on the TeachLivE comprehensive assessment, her performance in the classroom was not impacted by intervention on specific praise. Once the TeachLivE intervention was initiated, Deanna increased her mean proficiency with praise around to 89% (70%-100%), with an upward trend ending in 100% for the last two data points. Deanna had few opportunities to deliver praise around during the first four sessions of the maintenance phase (between days 29-38 following intervention), resulting in one combined data point at 67% correct steps. During the last assessment video, she delivered each step in the praise around process, but worded some of her praise around statements negatively (e.g., “Thank you for not having anything in your hands”). The primary researcher provided feedback to Deanna about how to correctly word praise around statements, and collected two more classroom assessments (days 43 and 44 following intervention). Deanna demonstrated 100% proficiency with praise around for both
sessions following feedback.

Marie’s Results

TeachLivE performance. Marie’s performance on error correction is illustrated in the first leg of the multiple baseline graph in Figure 8. On average, she delivered 39% correct error correction steps on the TeachLivE comprehensive assessment, with a trend descending from 50% to 27% correct steps. Only two baseline data points were collected because she missed session 2 due to a family emergency. Similar to Lisa, Grayce, and Deanna, when focused feedback was initiated, Marie demonstrated an immediate level change on the TeachLivE comprehensive assessment ($M = 80\%$ [75%-88%]). Moreover, her average proficiency with error correction on the TeachLivE training assessment ($M = 96\%$ [89%-100%]) was higher than on the TeachLivE comprehensive assessment during focused feedback, similar to Lisa and Grayce. During cumulative feedback, Marie increased her average proficiency on the TeachLivE comprehensive assessment to 90% (75%-100%). However, similar to Grayce, she showed an initial decrease in performance on the TeachLivE training assessment (60%), which trended back up by the second session and resulted in similar average proficiency as observed during focused feedback ($M = 91\%$). Overall, Marie improved her mean delivery of correct error correction steps 49% from baseline to intervention on the TeachLivE comprehensive assessment, which was similar to the improvement observed for both Grayce and Deanna. Marie demonstrated 100% proficiency with error correction on the TeachLivE comprehensive assessment 44 days later, just as observed for Deanna.

The second leg of the multiple baseline graph in Figure 8 shows Marie’s
Figure 8. Marie’s percentage of correct steps for error correction, specific praise rate per min, and percentage of correct steps for praise around on the TeachLivE training assessment, TeachLivE comprehensive assessment, and classroom generalization assessment.
performance on specific praise. Similar to the other three participants, Marie delivered a low rate of specific praise statements per min during baseline sessions on the TeachLivE comprehensive assessment ($M = 0.2 \ [0.0-0.4]$), and then demonstrated a distinct level change when focused feedback was introduced ($M = 2.6 \ [2.0-3.0]$). This data path was parallel to, and slightly lower than, her performance on the TeachLivE training assessment ($M = 3.1 \ [2.3-4.0]$), which was also similar to the pattern observed for the other participants. During cumulative feedback, Marie continued to increase her rate of specific praise on both assessments. On the TeachLivE comprehensive assessment she delivered an average of 3.6 ($2.8-4.0$) specific praise statements per min, and on the TeachLivE training assessment she delivered an average of 4.7 ($3.7-6.0$) specific praise statements per min. Marie’s performance during cumulative feedback differed from the other three participants because she delivered higher specific praise rates on the TeachLivE training assessment than on the TeachLivE comprehensive assessment. However, when considering performance during focused feedback and cumulative feedback combined, Marie delivered an average of 3.0 more specific statements per min during intervention than baseline on the TeachLivE comprehensive assessment, which was similar to the improvement observed for the other participants. Marie increased her delivery of specific praise statements to 4 per min on the TeachLivE comprehensive assessment 44 days following intervention.

The third leg of Marie’s multiple baseline graph displays her results for praise around (see Figure 8). Marie delivered no praise around steps for the first three baseline data points on the TeachLivE comprehensive assessment, and increased to 17% correct
steps for the fourth data point, which was the session prior to intervention on specific praise. Similar to Lisa and Deanna, Marie began to deliver more praise around steps after intervention on specific praise, but showed a descending trend from 33% to 9% correct steps. When praise around was targeted with focused feedback, Marie demonstrated a substantial level change on the TeachLivE comprehensive assessment ($M = 88\%$ [67%-100%]), with even higher proficiency on the TeachLivE training assessment ($M = 98\%$ [88%-100%]). Overall, she delivered an average of 78% more correct praise around steps during intervention than baseline on the TeachLivE comprehensive assessment. Like the other participants, Marie demonstrated proficiency with praise around on the TeachLivE comprehensive assessment when assessed 44 days post-intervention. She delivered 100% correct praise around step during this session.

**Classroom generalization.** Marie generalized proficiency with error correction to the classroom. However, in contrast to the overall patterns observed for the other participants, Marie’s performance levels and trends were different from those observed in corresponding phases in the virtual environment (see Figure 8). During baseline, Marie’s delivery of error correction was 7% correct steps for the single data point collected, which was much lower than her baseline delivery in TeachLivE. Only one baseline data point was reported because she did not teach the week of session 2, and data were combined for sessions 1 and 3 to include at least five possible error correction steps. (Note: Because Marie had fewer than three error correction baseline data points on both the TeachLivE comprehensive assessment and classroom generalization assessment, she did not meet the baseline phase change criteria.) As soon as the TeachLivE intervention
was initiated, Marie began proficiently delivering error correction in her classroom.

During focused feedback she averaged 93% (89%-100%) correct error correction steps, which represented a more substantial performance increase from baseline to intervention in the classroom than in TeachLivE. However, during cumulative feedback, Marie’s average proficiency with error correction in the classroom dropped to 78% (67%-89%), with a descending trend. Marie is the only participant who demonstrated a decreasing performance trend for error correction in the classroom during cumulative feedback. Moreover, the observed descending performance trend in the classroom contrasts with her stable, but variable, performance pattern on both TeachLivE assessments during cumulative feedback. Nevertheless, when considering her average performance across focused feedback and cumulative feedback, she delivered 77% more correct error correction steps during intervention than during baseline, which was the highest percent increase from baseline to intervention in the classroom for all participants. During the maintenance phase, Marie was assessed in her classroom five times between 29 and 44 days after the TeachLivE intervention, and demonstrated a similar average proficiency with error correction as Lisa and Deanna ($M = 93\%$ [80%-100%]). Marie generalized the skill of specific praise to her classroom, but did so at lower rates than she demonstrated during each corresponding phase in TeachLivE (see Figure 8), and at lower rates than the other participants delivered in their classrooms. During baseline, she delivered an average of 0.2 (0.1-0.2) specific praise statements per min, which was the same average rate she delivered during baseline sessions on the TeachLivE comprehensive assessment. Once the TeachLivE intervention was initiated,
Marie increased her average specific praise rate in the classroom during focused feedback to 2.0 (1.9-2.0) statements per min, and maintained her specific praise rate during cumulative feedback \((M = 2.1 [1.3-4.4])\), but demonstrated more variability. On average, Marie delivered 1.9 specific praise statements per min more during intervention than baseline in her classroom. She maintained the same average specific praise rate \((M = 2.1)\) when assessed five times between 29 and 44 days following the TeachLivE intervention, with specific praise rates ranging from 1.2 to 2.8 statements per min.

Overall, Marie generalized proficiency with praise around to her classroom, but demonstrated different performance patterns than observed in TeachLivE (see Figure 8). She delivered no correct praise around steps in her classroom during baseline, in contrast to her performance on the TeachLivE comprehensive assessment. When the TeachLivE intervention was introduced, Marie demonstrated an immediate level change with a steep upward trend, starting at 46% correct praise around steps, and ending at 100% correct praise around steps for the last data point of the phase \((M = 71\%)\). On the first maintenance assessment, collected 29 days after the last TeachLivE intervention session, Marie delivered 67% correct praise around steps, which was slightly lower than her average classroom performance during TeachLivE intervention. She then improved her performance to 100% correct steps, and then dropped to 0% correct steps for the last data point. Similar to Deanna, Marie attempted each step in the praise around sequence, but worded the statements negatively (e.g. “Thank you for not shouting out”). At this point, the primary researcher gave the same feedback to Marie as she gave to Deanna about how to correctly word praise around statements, and then collected one more classroom
assessments 44 days following intervention. Following feedback, Marie demonstrated 100% proficiency with praise around, just as observed for Deanna.

**Social Validity Results**

The fourth research question addressed social validity, and asked: How do preservice special educators perceive the realism of the TeachLivE classroom and value the intervention and assessment procedures? Overall response patterns from the social validity survey are summarized below.

**Realism of TeachLivE**

Results from the first seven items of the social validity survey are displayed in Table 10, and indicated that participants perceived TeachLivE as an authentic classroom where they engaged in realistic teaching situations. All participants agreed or strongly agreed that TeachLivE felt like a real classroom, that they understood the students’ different personalities, and that they interacted with them as they would in a real classroom. One participant noted that she did not think of the virtual students as real children, but also indicated that she interacted with them as if they were. Also, two participants were undecided about whether or not the Xbox Kinect feature that visually zoomed in as they approached a student and zoomed out when they moved away from a student enhanced their interactions in the virtual classroom. All four participants felt the scenarios they encountered in the TeachLivE sessions were very similar to situations they faced in their own classroom. One participant identified specific aspects of the virtual experience that were realistic, including “Talkouts, different personalities, almost word
Table 10

Results of Social Validity Items addressing Realism of TeachLivE

<table>
<thead>
<tr>
<th>Social validity item</th>
<th>Percentage of participants’ ratings</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. The TeachLivE classroom feels like a real classroom.</td>
<td>50  50</td>
</tr>
<tr>
<td>2. During my interaction with the TeachLivE students, I began to understand their different personalities.</td>
<td>75  25</td>
</tr>
<tr>
<td>3. During my interaction with the TeachLivE students, I thought of them as real kids.</td>
<td>50  25  25</td>
</tr>
<tr>
<td>4. I was able to interact with the TeachLivE students like I would in a physical classroom.</td>
<td>100</td>
</tr>
<tr>
<td>5. My visual proximity to the TeachLivE students (zooming in when I approached a student and zooming out when I moved away) enhanced my interactions in the classroom.</td>
<td>25  25  50</td>
</tr>
<tr>
<td>6. The scenarios I encountered during the training sessions were similar to situations I encounter in my own classroom.</td>
<td>100</td>
</tr>
<tr>
<td>7. The scenarios I encountered during the assessment sessions were similar to situations I encounter in my own classroom.</td>
<td>100</td>
</tr>
</tbody>
</table>

Note. SA = strongly agree; A = agree; U = undecided; D = disagree; SD = strongly disagree

for word what my students say.” Another participant wrote, “Personalities were similar. Content and behavior issues were similar too.” Another focused on specific behaviors she encountered, such as “talkouts, getting out of seat, clicking pen, were behaviors I deal with daily.” Moreover, all participants strongly agreed the scenarios encountered in the TeachLivE comprehensive assessment sessions were similar to their own classroom, though some also pointed out slight differences. One participant indicated, “...error correction happens very similarly. The only difference is that my students tend to repeat
errors a little more.” Another wrote, “I don’t have cell phone problems, my kids are too little. But talkouts, conversations, etc. were right on with what I see in my class.”

Comments from two participants suggested they viewed the increased complexity of the TeachLivE comprehensive assessments to be even more realistic to their classrooms than the simplified TeachLivE training scenarios that incrementally increased in complexity. One participant wrote, “These sessions were more realistic to my school.” The other noted, “by the end of the training sessions it was the same as the assessment sessions.”

**Value of the Intervention and Assessment Procedures**

The remaining 22 items of the social validity survey addressed participants’ perceptions of the value and acceptability of the intervention and assessment procedures. Overall, results indicated strong consensus that the participants felt the intervention components were valuable in becoming proficient with the target skills and that they viewed the assessment procedures as acceptable. All participants agreed or strongly agreed that the didactic instruction and quizzes completed prior to training sessions were helpful in introducing, defining, and solidifying their knowledge of each target skill. Additionally, with only two exceptions, they all agreed or strongly agreed that each specific feedback procedure employed during training sessions helped them improve their performance on the target skills (i.e., hearing feedback from the mentor, providing feedback to their peers, verbalizing their own goals for improvement, explicitly making connections between the virtual environment and their classroom, and receiving a copy of their feedback form). The only exceptions were one participant who disagreed that
verbalizing her goals for improvement was helpful, and another who disagreed that receiving a copy of the feedback form was helpful. Overall, participant responses indicated that the lesson materials provided by the primary researcher were useful, that most spent 1-5 min preparing to deliver them during training sessions and comprehensive assessment sessions, and that all participants prepared alone. Participants were split in their perceptions of the value of the online discussions. All agreed or strongly agreed that the procedures for recording in their classroom and submitting videos to the primary researcher were straightforward and took a reasonable amount of time. When presented with a hypothetical situation of earning money for mastering a new skill in their classroom within a week, all participants indicated they would choose to first practice the skill in TeachLivE, as opposed to practicing with a peer, or foregoing prior practice altogether.

When asked for any additional comments, all participants responded positively to the TeachLivE experience. The following response by one participant was representative of the other three participants’ reactions: “I was hesitant to participate in TeachLivE, but I ended up loving it. It was a great way to practice and refine my skills in a safe environment, and then be able to go back to my classroom and apply it and see immediate results there as well. I also strongly enjoyed videoing [in my classroom]. It helped me make sure I was applying the skills I learned.”
CHAPTER V

DISCUSSION

The critical and ongoing need for highly skilled special educators is evident in the widespread teacher shortages in the field (Tyler & Brunner, 2014) as well as the concern about outcomes for students with disabilities (Mason-Williams, 2015). High attrition indicates traditional instruction and field placements may be insufficient to produce competent teachers who are capable of dealing with the multifaceted demands of today’s special education classrooms. Situated learning approaches, such as virtual simulations, may help bridge the gap between teachers’ knowledge about best teaching practices and their classroom application. However, outcomes likely depend on the sophistication of the virtual environment and the extent to which the practice opportunities match the demands of authentic teaching situations (see J. S. Brown et al., 1989; Spiro, Feltovich, & Coulson, 1996), as well as the alignment of the practice tasks to the learners’ current abilities as they move from novice to proficient with a skill set (see Jacobson & Spiro, 1991). TeachLivE is a particularly powerful simulation because it provides a full-immersion virtual teaching experience with scenarios that can be strategically increased to match the complexities that exist in a real classroom, enabling learners to acquire and combine essential skills incrementally.

The purpose of this study was to utilize the TeachLivE middle school simulation to promote participants’ mastery of critical competencies aimed at student academics and behavior. The results from this study indicate that TeachLivE is a powerful platform for providing repeated practice and feedback on essential target skills. All participants
improved performance on error correction, specific praise, and praise around on both the TeachLivE training assessment and the more complex TeachLivE comprehensive assessment. Additionally, there was a strong pattern of generalization from TeachLivE to the classroom. Moreover, participants maintained proficiency with the majority of the target skills when assessed in TeachLivE and their classroom 1-1.5 months following intervention. On the social validity survey, participants indicated that the TeachLivE middle school students and teaching scenarios were realistic representations of their own students and classroom settings, and that the intervention and assessment procedures were acceptable and helpful in mastering the essential target skills.

**Implications for Teacher Preparation**

The current study makes several contributions to the existing literature base, and the results reveal several issues worth considering in teacher preparation. First, this study was designed to look at cumulative proficiency with multiple target skills, demonstrating the power of incrementally increasing the learners’ cognitive load through interleaved practice in a virtual environment. Second, generalization data were collected across all phases of the study, which addressed a critical omission evident in the available virtual simulation literature. Third, maintenance data were collected in the classroom and the virtual environment, showing the extent to which participants proficiently delivered target skills after intervention in the virtual setting was discontinued. These contributions are explored below.
Interleaving Target Skills

When instruction on target skills is interleaved, learners have opportunities to cumulatively build proficiency with skills (Dunlosky et al., 2013). When skills are introduced cumulatively, the practice environment increases in complexity with the introduction of each new skill, thus gradually increasing the learners’ cognitive load. Applying these principles to teacher preparation requires controlling situational contexts so various teaching repertoires may be introduced systematically. TeachLivE provides a medium for controlling situational contexts and interleaving essential skills from both the academic and behavior domains. Specifically, Vince Garland et al. (2016) recommended that researchers should investigate mastery of skills in TeachLivE by first expecting novice teachers to demonstrate proficiency when an avatar’s behavior is simplified, and again when an avatar’s behavior is increased to simulate more complex student behavior or classroom situations. In the current study, the avatars’ academic errors and problem behaviors were systematically increased in TeachLivE. This increased the complexity of student behavior and created practice opportunities for participants to interleave target skills. The TeachLivE training assessment was structured to emphasize the current skill targeted during intervention and simultaneously provide opportunities to practice previously targeted skills. In contrast, the TeachLivE comprehensive assessment was designed to provide opportunities on all skills, whether or not they had been targeted yet in intervention. Thus, the TeachLivE comprehensive assessment introduced a consistent level of complexity, which was higher than participants encountered on the TeachLivE training assessment, until the last intervention phase when the complexity was similar in
both settings.

The value of interleaved practice is evident when comparing participants’ performance on the TeachLivE training assessment and on the TeachLivE comprehensive assessment. Participants demonstrated proficiency with the current target skill on both assessments during focused feedback, but typically at higher levels in the simplified training environment. For example, Lisa, Grayce, and Marie delivered a higher percentage of correct error correction steps on the TeachLivE training assessment than the TeachLivE comprehensive assessment. Similarly, Grayce, Deanna, and Marie delivered higher rates of specific praise during focused feedback on the TeachLivE training assessment than the TeachLivE comprehensive assessment. However, during cumulative feedback, when the complexity of the training environment increased, participants closed these performance gaps for both error correction and specific praise and demonstrated similar proficiency on the two assessments. The only exception was for Marie who continued to deliver a higher specific praise rate on the TeachLivE training assessment than the TeachLivE comprehensive assessment. Interestingly, participants demonstrated similar proficiency with praise around on both assessments during intervention, which is the phase of the study when the complexity of the TeachLivE training assessment aligned most closely with the TeachLivE comprehensive assessment. These patterns indicate that participants acquired a more sophisticated teaching repertoire as skills were strategically interleaved during intervention.

The effect of interleaved practice might also promote generalization of target skills to the classroom. As each new skill was targeted in intervention, participants not
only demonstrated proficiency with the new skill but also maintained or improved proficiency with prior skills on the classroom generalization assessment. The only exception to this pattern was Marie’s downward trend on error correction during cumulative feedback on the classroom generalization assessment. It is possible that skills were interleaved too quickly for Marie, impacting her ability to maintain prior performance levels on error correction in the classroom. Teacher educators should be aware that some teachers may require more interleaved practice than others, or different timing of the introduction of new skills, before they can proficiently balance the delivery of multiple skills in authentic environments.

These results contrast with the patterns observed in a study by Dawson and Lignugaris/Kraft (in press) where they intervened on target skills in TeachLivE in isolation. Although participants demonstrated a similar overall pattern of higher performance on the TeachLivE training assessment than the TeachLivE comprehensive assessment, the participants had difficulty maintaining previous levels of performance when feedback was withdrawn, especially delivering specific praise in the classroom setting. It is important to note, however, that another difference between the two studies was the order of the target skills. Dawson and Lignugaris/Kraft focused on basic teaching skills first (i.e., OTR and praise), and then introduced more complex skills for addressing behavioral and academic disruptors to the instructional flow (i.e., praise around and error correction). The frequency and intensity of these disruptions amplified the complexity of the instructional interaction (see Chapter II, Figure 1). The rationale for this training sequence was to first consolidate basic skills and then introduce complexity in the
simulated teaching scenarios. In contrast, in the current study, error correction was introduced immediately after participants were taught how to deliver OTRs, thus introducing academic disruptors (i.e., errors) immediately after practicing OTRs, the basic component skill. Next, specific praise was targeted while opportunities to correct errors continued, allowing participants to interleave a basic skill in the more complex instructional environment that included academic errors. The difficulty of the instructional environment was then increased again, when behavioral disruptors (e.g., talking out, pencil tapping, and other persistent problem behaviors) were added to academic disruptors, allowing participants to interleave praise around with error correction and specific praise. It is possible that sustained teacher proficiency with essential skills may depend not only on interleaved practice structures but also on the order skills are targeted in intervention. Teacher educators must find the delicate balance between introducing disruptors when teachers are proficient with basic skills and interleaving complex skills early enough to provide teachers with ample time to develop mastery within increasingly difficult classroom situations. Additionally, teacher educators may need to plan for longer intervention sessions as more skills are interleaved. For instance, in this study, intervention sessions were approximately 8 to 10 min longer than baseline sessions, which was likely due to the cumulative feedback structure that was implemented during each leg of intervention.

**Generalizing Proficiency to the Classroom**

Teacher educators continuously struggle with how to transfer elements of effective pedagogical practice from methods classes to the complex interactive
instructional environments in classrooms (Grossman, 2005; Grossman & McDonald, 2008; Lignugaris/Kraft & Harris, 2014). In a review of pedagogical practices in teacher education, Grossman found that there are a number of studies in which teacher educators examined the effects of simplified simulated opportunities for novice teachers to practice instructional pedagogy, however they found no studies in which teacher educators examined the effects of simplified simulated interactive teaching on actual classroom practice.

This study extends the available teacher education literature as well as the virtual simulation literature, including the two studies conducted in TeachLivE (Vince Garland et al., 2012, 2016), by demonstrating that practice in simplified instructional settings that progressively increase instructional complexity results in generalized performance of identified teaching practices to authentic classrooms. One variable that may affect teachers’ generalization from TeachLivE to the actual classroom is the similarity between the practice environment and the actual classroom environment. That is, when situations in teachers’ classrooms are similar to those encountered in TeachLivE and their students respond similarly to the virtual students, teachers are more likely to engage in the behaviors previously reinforced in TeachLivE.

For example, one factor that might have contributed to Lisa’s, Grayce’s, and Deanna’s generalized performance of error correction from TeachLivE to their classrooms is that the types of errors their students produced were similar to the types of errors the TeachLivE students produced and required similar error correction topographies. Moreover, the error correction typically resulted in a correct student
response and the participant could move forward with instruction. In contrast, Marie encountered different types of errors in her classroom than in TeachLivE. The majority of errors she faced in her classroom were on application questions that were typically more difficult to correct than the most complex errors she encountered with the virtual students. For instance, the TeachLivE students might overtly give an incorrect example of a word, whereas Marie’s own students might make multiple errors on one trial (e.g., give an incorrect example and also conjugate a word in the sentence incorrectly). Although Marie attempted to correct the majority of these errors, students sometimes repeated errors or produced new errors on the initial test or delayed test. This resulted in more complex instructional interactions than she was prepared to handle because her error correction was not successful. There are several potential reasons for the complexity of errors Marie encountered in her classroom. For example, many of her students were English Language Learners, the content may have been too advanced for the students, or perhaps the delayed test needed to be delivered more quickly and then repeated on successively longer intervals to firm the student’s response. Regardless of the reasons, it is clear that many of the error correction situations Marie encountered with her own students were more difficult to address than those she encountered in TeachLivE. These differences might have contributed to her downward performance trend on error correction in her classroom during cumulative feedback in TeachLivE. Importantly, although some situations were explicitly programmed in intervention to align with participants’ classrooms, these particular scenarios were not. In the future, specific situations that more closely match individual participants’ classrooms could be
programmed to increase the stimulus similarity between the virtual classroom and actual classroom environments.

Another factor that might have influenced participants’ generalization of performance from TeachLivE to the classroom is the frequency of opportunities to deliver a skill in their classroom. For instance, all participants’ demonstrated more variable performance on praise around in their classroom than in TeachLivE. Lisa and Grayce had a similar rate of opportunities to practice praise around in their classrooms as in TeachLivE, but Deanna and Marie encountered fewer opportunities to practice the skill in their classroom than in TeachLivE. The lower rate of misbehavior coupled with behavior topographies that were subtler than those presented in TeachLivE may have influenced their generalization of praise around to the classroom.

The analysis of similarity between TeachLivE and participants’ actual classroom supports J. S. Brown et al.’s (1989) theory that the extent to which learners apply skills in an authentic setting depends on how closely the stimulus situations in the practice environment align to those in the authentic settings. Differences in stimuli and the frequency of opportunities to practice a skill should be expected between the virtual environment and the classroom, even when utilizing a simulation as sophisticated as TeachLivE. Teacher educators need to recognize that this will affect the extent to which generalization is observed across settings, and when possible they should program stimuli in the virtual setting to promote success in real classrooms.

**Maintenance of Target Skills**

Few researchers in the virtual simulation literature collected maintenance data,
and none measured teachers’ maintenance of skills in the classroom. The maintenance data collected in this study show promising proficiency with essential target skills following intervention. Lisa maintained proficiency with each target skill on the TeachLivE comprehensive assessments and in her classroom. Grayce maintained proficiency with specific praise and praise around on the TeachLivE comprehensive assessment and in her classroom. Grayce’s decreased performance on error correction on the TeachLivE comprehensive assessment was the result of delivering some of the tests and delayed tests with incorrect wording (e.g., “will you give me another example of apparent”). Both Deanna and Marie maintained proficiency with error correction, specific praise, and praise around on the TeachLivE comprehensive assessment, and maintained proficiency with error correction and specific praise on the classroom generalization assessment. Similar to Grayce, their decreased proficiency with praise around on the classroom generalization assessment was because of using incorrect wording (e.g., “Thank you for not talking out.”). Grayce, Deanna, and Marie discriminated opportunities to correct errors or to praise around for desired behavior, but the topography with which they delivered the skill began to drift. Importantly, Grayce’s delivery of the error correction steps led to correct student responding, and Deanna’s and Marie’s praise around statements evoked the desired behavior from the target students. Therefore, their delivery of the skills remained functionally effective. However, it is possible that over time the observed drift could become more pronounced and impact the effectiveness of the skills in their classrooms. When the primary researcher provided brief feedback to Deanna and Marie about the topography of their praise around
statements in the classroom, both achieved 100% proficiency for the remaining data point(s). This pattern suggests that teacher educators may need to provide occasional feedback on previously mastered skills so teachers continue delivering them correctly over time.

**Limitations**

This study makes substantial contributions to the existing literature on using virtual simulations in teacher preparation. However, there are some limitations to consider as well, such as the interdependence of specific praise and praise around, the limited number of data points in some phases, concerns with some reliability scores, the narrow scope of generalization, the inclusion of multiple intervention components, and potential confounding variables.

**Interdependence of Specific Praise and Praise Around**

One assumption of a multiple baseline design is that target skills are independent, so intervening on one skill will not impact participants’ performance on a skill still in baseline (Cooper et al., 2007). The exception in the current study was that specific praise is a component of praise around, making the two target skills interrelated. This limitation was minimized in the current study by targeting specific praise prior to praise around. In previous research, Dawson and Lignugaris/Kraft (in press) found that participants did not improve their performance on praise around as a function of practicing specific praise. Nevertheless, increases in correct praise around steps during baseline were observed for
Lisa and Deanna on the TeachLivE comprehensive assessment that corresponded with the timing of intervention on specific praise. Also, Lisa improved delivery of praise around during the final baseline session on the classroom generalization assessment. However, it is not likely that either participant would have become proficient with praise around as quickly as observed during intervention given the patterns established on the last three baseline data points. Specifically, both Lisa and Deanna demonstrated level changes for praise around on the TeachLivE comprehensive assessment and on the classroom generalization assessment during intervention well above the trend line established by their last three baseline data points. In contrast to the patterns observed for Lisa and Deanna, Grayce’s and Marie’s baseline performance on praise around was not directly impacted by intervention on specific praise. They both delivered some correct praise around steps during baseline on the TeachLivE comprehensive assessment prior to and after intervention on specific praise, and Grayce demonstrated the same pattern on the classroom generalization assessment.

It is possible that participants’ specific praise rate during cumulative feedback was impacted by intervention on praise around because they strategically practiced delivering specific praise to target misbehavior. This may explain why all participants increased their average specific praise rate on the TeachLivE comprehensive assessment during intervention on praise around. Similarly, during cumulative feedback, Deanna and Marie increased their average specific praise rate on the TeachLivE training assessment, and Lisa and Grayce increased their average specific praise rate on the classroom generalization assessment.
**Number of Data Points per Phase**

At least three data points per phase are recommended to demonstrate experimental effects in a single subject study (Cook et al., 2014). This criterion was met in the majority of phases and assessments for each participant. Phases with fewer than three data points were due to participant absences, inability to record in the classroom, or lack of opportunity to engage in a skill in the classroom. Finally, only one maintenance data point was collected on the TeachLivE comprehensive assessment in order to withhold participants’ exposure to the virtual classroom while collecting classroom generalization data and also to conclude data collection prior to participants receiving training on new skills at the beginning of the subsequent semester.

**Reliability of Reported Data**

Overall, IOA scores were high for the target skills, which increases confidence that reported changes in participant behavior represented actual changes in their behavior (Cooper et al., 2007). However, there were some low IOA scores that must be considered when interpreting the data. First, there was a large range of IOA scores for the occurrence of specific praise on the TeachLivE comprehensive assessment and classroom generalization assessment. Sessions with 0% IOA on these assessments occurred most often during baseline when participants delivered few specific praise statements. For example, when a participant delivered one specific praise statement in a given session, and only one observer coded it as such, it resulted in an IOA score of 0% for that session. This contributed to the large range and somewhat low average IOA score for specific praise. Similarly, low IOA scores for error correction opportunities and praise around
opportunities were often the result of few errors or misbehaviors occurring in a session, especially on the classroom generalization assessment when the number of opportunities could not be programmed. Moreover, IOA scores for error correction opportunities and praise around opportunities were lower on the classroom generalization assessment than either TeachLivE assessment because errors and problem behaviors were less standardized in the real classrooms than in the virtual environment. The implication of these disagreements is that the primary researcher may not have detected all error correction or praise around opportunities. Thus, it is possible that the data presented is an overestimate of participants’ proficiency with a skill. Detecting error correction and praise around opportunities in the classroom was particularly challenging because sometimes novel student topographies were observed that set the occasion for the participants to engage in the error correction or praise around process. To minimize this limitation, a two-step recursive process was employed when a new student topography was identified. First, the primary researcher and research assistant agreed on the new definition of an error correction or praise around opportunity. Next, the primary researcher examined prior video recordings to identify instances of the newly defined situations. These situations were coded and participant data were updated accordingly. This process ensured consistency in definitions of error correction and praise around opportunities throughout all phases of the study. Importantly, IOA for the delivery of steps was high for both error correction and praise around in all settings.

The classroom recordings provided a somewhat narrow lens for viewing student behavior and a participant’s response to that behavior. It is possible that coding
participant behavior via video recordings impacted reliability of the reported data. For example, on some occasions it was difficult to clearly hear everything on a recording. This increased the probability of a coding error. However, the benefits of using permanent products to score participant performance outweighed potential drawbacks, especially because the video recordings eliminated the need for both researchers to be present during assessments and allowed them to watch sessions multiple times if needed to verify participants’ proficiency with the target skills.

**Limited Scope of Generalization**

Participants recorded themselves delivering a consistent type of instruction to the same group of students each week. Although this minimized variability across content areas and instructional groups, it also limited the scope of generalization data. It’s unknown how well participants generalized proficiency with target skills to lesson structures, content areas, or groups of students beyond what was observed on their classroom recordings. For example, Lisa always recorded at the beginning of her lesson, when students were reading isolated sounds, reading words with those sounds, and occasionally reading sentences with the target words. Similarly, Marie’s classroom recordings were limited to students reading vocabulary words and then giving examples of the words or using them in novel sentences. Grayce and Deanna included more instructional variety in their recordings. For instance, Grayce’s lessons included word attack tasks, passage reading, and comprehension questions, and Deanna’s lessons incorporated sight word flash cards, sound/spelling word sorts, dictation, and finding target words in a passage. However, even for participants who captured footage with a
wider range of instructional variety, the classroom generalization assessments represented a narrow scope of the content and students that participants taught in a given day. Therefore, it is impossible to surmise how well they delivered the target skills when they were not recording. Also, the very act of recording likely caused heightened awareness of their teaching behaviors and motivated them to do their best. Thus, it is possible the classroom generalization data are representative of participants’ best possible performance outcomes and not how they consistently taught throughout the day. Ideally, as participants improved their proficiency with the target skills they came in contact with natural communities of reinforcement (Cooper et al., 2007), such as increased student learning and compliance, which motivated them to continue delivering the target skills even when they were not recording for the purpose of the research study.

**Multiple Intervention Components**

Participants engaged in multiple activities as part of the intervention, including didactic instruction, teaching in the simulated classroom, observing peers teaching, hearing and providing feedback, and reflecting on the training experiences. It is difficult to determine how these components interacted to impact teacher performance both in the simulator and in their classroom setting. However, the available research suggests that skill transfer to authentic classroom situations is mixed as a result of traditional coursework and field experiences, which typically includes elements of didactic instruction, observing teaching, and written or oral reflections about teaching (Clift & Brady, 2005). Furthermore, most of these reported outcomes focused on teachers’ understanding of concepts or beliefs about teaching, as opposed to their ability to deliver
specific teaching techniques in the classroom. These data suggest that active teaching and ongoing feedback in TeachLivE were critical intervention components in participants’ mastery of skills and their ability to generalize proficiency to the classroom.

**Potential Confounding Variables**

Isolating independent variables in teacher preparation programs presents a particularly difficult methodological challenge. For instance, participants in the current study were exposed to several resources in the M/M ATP program beyond TeachLivE sessions that may have impacted their performance on the target skills. Specifically, participants received feedback from their district coaches throughout the semester, they attended weekly classes and completed coursework related to instructional techniques and behavior management strategies, and they likely received feedback about their teaching from colleagues and administrators at their school. However, the fact that participants had access to these resources throughout the entire research study, but typically did not demonstrate distinct level or trend changes on the target skills until they practiced the skill in TeachLivE, strengthens the case for a causal relationship between TeachLivE and participants’ observed improvement in performance.

Nevertheless, it is important for teacher educators to consider the extent to which variables interact to improve or maintain teachers’ proficiency with specific teaching competencies. District supervision was a particular concern in this study because participants received written and verbal feedback from their instructional coach about their delivery of the target skills in their classroom on at least three occasions during the semester. To assess the effect of direct supervision on participants’ acquisition of the
target skills in their classroom, the instructional coaches’ data forms and written feedback were examined, as well as the timing of their visit. Some participants received feedback on the target skills across baseline and intervention phases. However, we observed no corresponding changes in participants’ performance in the instructional groups they recorded. Other participants received coaching on target skills following intervention in TeachLivE, which may have contributed to their maintenance of the target skills. Only two participants received coaching visits that directly corresponded with the beginning of intervention on a new skill. Grayce received coaching immediately prior to intervention on specific praise (session 7), and Marie received coaching just prior to intervention on praise around (session 10). However, based on the written data for these sessions, Grayce received feedback on her overall praise rate, but it is not clear if she received feedback on specific praise. There is no indication that Marie received feedback on praise around or on related skills such as planned ignoring of problem behavior. Based on the available data, it is unlikely that district coaching had a substantial impact on participants’ acquisition of the target skills, although it could have facilitated participants’ maintenance of previously mastered skills.

Another variable that may have impacted participant performance is watching their weekly classroom recordings. Although this was not an expectation, Grayce, Deanna, and Marie indicated that they watched their videos every week or almost every week. In contrast, Lisa indicated that she never watched her classroom videos, which suggests it was not a necessary component for improvement on the target skills.
Future Research

There are several future research avenues for investigating the effectiveness of TeachLivE in teacher preparation programs. These include isolating the critical TeachLivE intervention components, targeting progressively more difficult skill sets in the virtual setting, evaluating a broader range of generalization from TeachLivE to authentic classrooms, and comparing TeachLivE to other training platforms available in the field.

First, the present study included several intervention components embedded into the TeachLivE sessions. Researchers should determine the most efficient and effective ways to utilize TeachLivE by comparing outcomes of different dosages of time in the simulation, group vs. individual intervention sessions, and various feedback protocols. Additionally, researchers could investigate the effectiveness of training protocols not utilized in the current study, such as visual performance feedback (Sweigart, Landrum, & Pennington, 2015), bug-in-ear technology (Elford, 2013), and teacher reflections on videos of their teaching in the simulator or classroom setting (Powell, 2016; Welsch & Devlin, 2007).

Second, researchers should investigate skill sets that extend from those targeted in the current study. Incorporating component skills into larger and more widely applicable units is a key recommendation for improving teacher preparation programs and related teacher and student outcomes (McLeskey & Brownell, 2015). In the academic domain, more intensive intervention structures could be practiced in TeachLivE to address errors for which basic error correction procedures are not sufficient to clear up student
misconceptions. For example, teachers could be taught to adapt the task demand of the test and delayed test to support student understanding and increase the probability of correct responding (Jones et al., 2007). Additionally, teachers who are already proficient with delivering OTRs and error correction could be taught to apply those skills within new instructional frameworks, such as leading a classroom discussion. In this context, they could increase the sophistication of their delivery of OTRs by scaffolding questions based on student responses, including strategically mixing low-level and high-level questions. They might also extend their error correction routine by utilizing prompts and modeling critical thinking skills, while still retaining foundation components of error correction, such as delivering timely tests and delayed tests and ending with students actively emitting a correct response (Lignugaris/Kraft & Harris, 2014). In the behavior domain, teachers could move beyond praise around to precision commands and individualized consequences (Bahadourian & Greer, 2005; Rhode et al., 1992; Yeager & McLaughlin, 1995) when the basic praise around technique is insufficient to produce the desired change in student behavior.

Third, in this study we explored generalization of target skills to a narrow range of instructional contexts in authentic classrooms. Isolating the critical intervention components and structures in TeachLivE and extending target skills to more sophisticated teaching practices as discussed above will allow researchers to explore the extent to which teachers generalize competencies from the virtual environment to more diverse instructional situations. For example, researchers might first evaluate to what extent individuals generalize target skills to a variety of language arts lesson structures. This
research might then be followed with investigations that explore generalized application of skills from TeachLivE to contexts that are increasingly different from those encountered in the virtual environment, such as different content areas, or older or younger students. These investigations are important because broader generalization of target skills will increase the usefulness of TeachLivE as a training tool in teacher preparation programs, especially because it is expensive and time-intensive to develop new TeachLivE avatars and classroom settings.

Finally, the field would benefit from controlled comparisons of the effectiveness of TeachLivE to other situated learning approaches, such as other classroom simulators or more traditional approaches like roleplaying. These data would contribute to a cost-benefit analysis of utilizing TeachLivE as a core component of teacher preparation, which would be valuable for teacher educators considering integrating the technology into their program.
REFERENCES


* Denotes inclusion in virtual simulation literature review.


Dawson, M. R., & Lignugaris/Kraft, B. (2015). *Specific praise and student outcomes: A research synthesis*. Unpublished manuscript., Utah State University, Department of Special Education Rehabilitation, Logan, UT.


Sutherland, K. S., & Wehby, J. H. (2001). Exploring the relationship between increased opportunities to respond to academic requests and the academic and behavioral outcomes of students with EBD. *Remedial & Special Education*, 22, 113-121.


Appendix A

TeachLivE Orientation Session Instructions
TeachLivE Orientation Session Instructions

The following script should serve as a guideline when introducing the TeachLivE Lab:

**Background**

“TeachLivE is a virtual classroom that was recently developed by the University of Central Florida. In 2009, Utah State University came on as the first partner. As a partner we have the rights to use the technology with Utah students to train teachers in Utah.

“Today you will be working with 5 middle school students. You will find that they represent *the range of personalities and abilities* you would encounter in a real classroom. As you can see the students are not physically present in the room, but you see a representation of what each student looks like on the screen. It is important to understand that you are interacting with *real people in real time*.

“The students’ names are on the podiums that represent their desks. In just a few minutes I will walk through the classroom and talk to each student, and then you will have a turn to walk through the classroom, so you become comfortable with the technology and start to learn the personalities of the students.”

**How the Technology Works**

**Wearing the microphone:**

“This is the microphone that you will wear when you are teaching the students in TeachLivE. The microphone allows the students to hear you better than other sounds in the room. However, we encourage those watching to observe quietly because the system picks up background noise as well.”

**Tracking/proximity using the Kinect system:**

“This Kinect system will track your movements and represent the proximity to students you would experience in a physical classroom. Before you begin each teaching turn it may take a few seconds for the Kinect to pick you up, and then it will track your movements as you move through the classroom. So, if I move closer to the desk of (student name) you see how the screen view changes to represent my proximity to that student.” (Continue to demonstrate as needed by moving through the students).

**The “Teacher Coordinator”:**

“It is important for you to know that there is a teacher coordinator in the other classroom with the students. He or she can see and hear us at all times. However, the students can only hear you when you are actively teaching them. In between turns you do not have to worry about the students hearing what you say. This is important for you to know, especially if you have something sensitive to share that would not be appropriate to say in front of a student.”
Classroom Pause/Classroom Come Back:
“At the beginning of each turn please say ‘classroom start’ or ‘classroom come back.’ This cues the teacher coordinator that we want the students to be able to actively listen again. When you are done with your teaching interaction, simply say, ‘classroom pause.’ The teaching scenario will stop right where you left it, the screen will show the door of the classroom instead of the students, and only the teacher coordinator will be able to hear you. You may also use this feature if you get stuck, need help, or want to try a section of your lesson again. When you ‘pause’ the classroom you can ask the mentor for help or address any questions you have. When you are ready to resume teaching simply say ‘classroom start’ or ‘classroom come back.’ The students will be able to hear you again and you can continue your turn.” (Demonstrate how this looks on the screen)

Classroom Walk-Through
“Now I will walk through the classroom and talk to the students so you can see how the technology works and start to become familiar with the students. Then it will be your turn to walk through the classroom.”

(During the walk-through demonstration be sure to approach each student’s desk and engage in a conversation. This could be a “getting to know you” question that you ask everyone, such as what is your favorite subject in school, or what did you do over the summer/winter break, or anything else that seems appropriate. Feel free to use the “classroom pause” and “classroom start” functions during the demonstration as well).

Now let each teacher walk through the lab and speak to the students for a few minutes.

*Note: It is imperative to give enough information about how the technology works, without revealing everything that happens behind the scenes. If teachers ask for more information or start guessing how the technology works, it is crucial not to divulge any information that does not appear in the script. A recommended response in this event is, “everything we have told you about the technology is true. Beyond that, you are free to come to your own conclusions about how TeachLivE works.” (Please do not confirm or deny their assumptions, just tell them they are free to wonder and come to their own conclusions).*
Appendix B

Letter of Information
LETTER OF INFORMATION

From TeachLivE™ to the Classroom: Building preservice special educators’ proficiency with essential teaching skills

Introduction/ Purpose Benjamin Lignugaris/Kraft and Melanie Rees Dawson in the Department of Special Education and Rehabilitation at Utah State University are conducting a research study to investigate the extent to which teachers develop proficiency on essential teaching skills in TeachLivE™ and then generalize their performance to their real classrooms. You have been asked to take part and provide consent because you are part of the Alternative Teacher Preparation (ATP) program at Utah State University. There will be approximately 4 participants at this site. There will be approximately 4 total participants in this research.

Procedures If you agree to be in this research study, the following will be expected:

1. Participants will participate in weekly training sessions in TeachLivE™
   - Prior to the first session on a new target skill, the participants will be required to view a training video and handout and then complete a short quiz. These materials will be available on the Canvas course site or via email 3-5 days prior to the session.
   - Approximately 4 participants and 1 instructor will be present at each training session. Participants will take turns participating in teaching scenarios, while the instructor and all other participants observe.
   - Participants will deliver vocabulary lessons to the TeachLivE™ students. The instructor will provide the content and structure for these lessons.
   - Participants will receive feedback from the instructor about the delivery of their lessons, will verbalize self-reflections about their performance, and will have the opportunity to provide feedback to their peers. All participants in the group will be present while feedback is given.
   - All training sessions will be video-recorded and will be viewed later by a research assistant to ensure that the instructor followed specific training protocols as outlined in the study.

Note: Viewing the videos, completing quizzes, and attending weekly training sessions are requirements of the program for practicum credit, and all ATP teachers will participate in these weekly trainings and activities regardless of if they are part of the research study or not. Those participating in the study will stay for a weekly assessment session where they will teach an additional lesson, as described below. Performance in the assessment session will in no way impact an individual’s grades in the ATP.
program.

2. Participants who are willing to participate in the research study will stay for an assessment session immediately following the training session each week.
   - Teachers will be given a break of approximately 5 minutes between the training and assessment session.
   - Teachers will be expected to stay an additional 5-30 minutes to complete the assessment session. These sessions will be private, with only 1 participant and the instructor present in the TeachLivE™ classroom. Each teacher will be assessed for a total of 5 minutes in the TeachLivE™ classroom. As soon as a teacher has completed their private assessment session, he/she is excused to leave. The order of participants will be rotated each week to ensure fairness.
   - Participants will deliver a different assessment lesson each week. The content and structure for these lessons will be provided by the instructors and distributed via email or on the Canvas course site 5-7 days before the session.
   - All assessment sessions will be video-recorded and the target skills will be scored later.

3. Observations will be conducted in the participants’ classrooms approximately once a week, via video-recording. The instructor will meet with each participant prior to classroom data collection to determine the best placement of the Flip Video Camera and to train the participant on submitting the weekly videos via USU’s big file transfer service. The participant will record their delivery of a lesson with the provided Flip Video Camera, and will submit the footage to the instructor. Additionally, data collected by the ATP professors and districts mentors during informal and formal observations may also be used.

4. Participants may be asked to participate in additional tiered supports if needed to improve performance on the essential target skills throughout the study. This could include additional feedback from the instructor on the assessment or classroom videos, visual performance feedback of performance, or reflections of one's teaching from the videos collected in TeachLivE™ or in the classroom.

5. Participants will be expected to participate in 15 weekly training sessions. It is anticipated that the intervention sessions will begin on August 31st and will conclude on December 7th (or Dec. 14th if a make-up session is needed).

6. Following all training and assessment sessions, the participants will complete a 10-20 minute survey about their experiences in the TeachLivE™ classroom.

7. Participants will be asked to record 2-4 additional lessons in their classroom after the training and assessment sessions have been discontinued. These videos will be scored to determine the extent to which participants maintained proficiency on target skills in the classroom. It is anticipated that these data will be collected by January 31, 2016.

**New Findings** During the course of this research study, you will be informed of any significant new findings (either good or bad), such as changes in the risks or benefits resulting from participation in the research, or new alternatives to participation that might cause you to change your mind about continuing in the study. If new information is obtained that is relevant or useful to you, or if the procedures and/or methods change at any time throughout this study, your consent to continue participating in this study will be obtained again.

**Risks** Participation in this research study may involve some added risks or discomforts. These include:
1. During the training sessions you will be expected to teach in front of your peers and receive feedback from the instructor in front of your peers. However, this is no different from the potential discomforts that will be experienced by other ATP teachers who are not participating in the study.

2. Participating in the study will require some extra time beyond the requirements of the ATP program, which will include staying after the weekly training session to participate in the private assessment session, preparing the assessment lessons on a weekly basis, and recording and submitting videos from the field placement classroom on a weekly basis.

**Benefits** There may or may not be any direct benefit to you from these procedures. It is anticipated that you will receive more practice and feedback on essential teaching skills, which may improve your teaching in your own classroom. The investigator expects to learn more about the effectiveness of TeachLivE™ and the extent to which teachers transfer their skills from the lab to the classroom. This information will help teacher trainers understand the best methods for training preservice special education teachers.

**Explanation & offer to answer questions** Melanie Dawson has explained this research study to you and answered your questions. If you have other questions or research-related problems, you may reach Melanie Dawson at 801-505-3290.

**Payment/Compensation** You will be paid a stipend of $300 for your participation in this study, to help compensate you for the extra time required to take part in this research. This stipend will be prorated in the amount of $20 per week of full participation in the study, which is defined as participating fully in the training and assessment sessions, and recording and submitting a lesson from the field placement classroom. The stipend will be paid following the last training and assessment sessions, and following completion of the final survey.

**Voluntary nature of participation and right to withdraw without consequence** Participation in research is entirely voluntary. You may refuse to participate or withdraw at any time without consequence or loss of benefits, with the exception of the forfeiture of the remainder of the prorated $300 stipend. You may be withdrawn from this study without your consent by the investigator due to poor attendance, disruptive behavior during training sessions, or failure to participate in the procedures as outlined in this document.

**Confidentiality** Research records will be kept confidential, consistent with federal and state regulations. Only Melanie Dawson, Benjamin Lignugaris/Kraft, TeachLivE™ instructors, research assistants, and ATP professors will have access to the data, which will be kept in a locked file cabinet in a locked room and/or on a password protected computer. To protect each participants’ privacy, a code will be given in the place of each individual’s name. Personal, identifiable information will be kept on file for one year in order to finish analyzing data. Melanie Dawson will keep videos of the training and assessment sessions to be used for training new research assistants on similar target skills for future studies. If you request further confidentiality, video footage will be edited so that your face is blurred and unrecognizable to those who view the video. No identifiable information will be attached to these videos. Additionally, you may request destruction of your videos after 1 year. Any future research using these data will be submitted to IRB for prior approval.

**IRB Approval Statement** The Institutional Review Board for the protection of human
participants at USU has approved this research study. If you have any pertinent questions or concerns about your rights or a research-related injury, you may contact the IRB Administrator at (435) 797-0567 or email irb@usu.edu. If you have a concern or complaint about the research and you would like to contact someone other than the research team, you may contact the IRB Administrator to obtain information or to offer input.

**Copy of consent** You have been given two copies of this Letter of Information. Please sign both copies and retain one copy for your files.

**Investigator Statement** “I certify that the research study has been explained to the individual, by me or my research staff, and that the individual understands the nature and purpose, the possible risks and benefits associated with taking part in this research study. Any questions that have been raised have been answered.”

Benjamin Lignugaris/Kraft, Principal Investigator  
(435) 797-2382; ben.lig@usu.edu

Melanie Rees Dawson, Student Researcher  
(801) 505-3290; melanie.dawson@aggiemail.usu.edu

**Participant:** By signing this document I agree to participate:

__________________________  __________________  Date
Appendix C

Sample Vocabulary Lesson Materials
## Sample Vocabulary Lesson Materials

**Training Lesson A**
Adapted from Harcourt: StoryTown, Bold Moves, Grade 6, lesson 1.

<table>
<thead>
<tr>
<th><strong>Vocabulary Words:</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>1) hysterical</td>
</tr>
<tr>
<td>2) crestfallen</td>
</tr>
<tr>
<td>3) incapacitated</td>
</tr>
<tr>
<td>4) perishable</td>
</tr>
<tr>
<td>5) lamented</td>
</tr>
<tr>
<td>6) ricocheted</td>
</tr>
<tr>
<td>7) ecstatic</td>
</tr>
<tr>
<td>8) mirth</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Student-Friendly Definitions:</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>1) <strong>hysterical</strong>- someone in a panic or very excited.</td>
</tr>
<tr>
<td>2) <strong>crestfallen</strong>- someone who is very disappointed and sad.</td>
</tr>
<tr>
<td>3) <strong>incapacitated</strong>- someone who is unable to work due to an injury.</td>
</tr>
<tr>
<td>4) <strong>perishable</strong>- food that goes bad if it is not kept cold.</td>
</tr>
<tr>
<td>5) <strong>lamented</strong>- expressed deep sadness because of something.</td>
</tr>
<tr>
<td>6) <strong>ricocheted</strong>- when something bounced off other objects.</td>
</tr>
<tr>
<td>7) <strong>ecstatic</strong>- extremely happy.</td>
</tr>
<tr>
<td>8) <strong>mirth</strong>- a feeling of amusement that causes you to laugh out loud.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Suggestions for Instruction:</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Examples/non-examples (Refer to Format 1)</td>
</tr>
<tr>
<td>Generate examples (Refer to Format 2)</td>
</tr>
<tr>
<td>Demonstration/Explanations (Refer to Format 3)</td>
</tr>
<tr>
<td>Sentence Substitution (Refer to Format 4)</td>
</tr>
<tr>
<td>Sentence Generation (Refer to Format 5)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Additional Resources:</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>• Picture cards for each word</td>
</tr>
<tr>
<td>• Reading passage that contains the target words</td>
</tr>
</tbody>
</table>
Formats for Teaching Vocabulary

Format 1: Example/Non-example
1. This word is “x”.
2. What word? (Group or individual)
3. This is an “x”/This is not an “x”.
4. Repeat step 3 with examples and non-examples.
5. Is this “x” or “not x”? (Group or individual)
6. Repeat step 5 for each example and non-example.

Format 2: Generate Example
1. This word is “x”.
2. What word? (Group or individual)
3. This is an “x”.
4. Repeat step 3 with examples and non-examples.
5. Name an “x” or give me an example of “x”. (Group or individual)
6. Repeat step 5 for multiple examples with multiple students.

Format 3: Demonstration
1. This word is “x”.
2. What word? (Group or individual)
3. “X” means (definition or synonym)
4. What does “x” mean? (Group or individual)
5. Show me how you “x” or explain how you “x”. (Group or individual)
6. Repeat step 5 for multiple examples with multiple students.

Format 4: Sentence Substitution
1. This word is “x”.
2. What word? (Group or individual)
3. “X” means (definition or synonym)
4. What does “x” mean? (Group or individual)
5. What’s another way of saying (sentence with definition or synonym)
6. Group or individual repeats sentence using target word for definition or synonym.
7. Repeat steps 5-6 for each word

Format 5: Sentence Generation
1. This word is “x”.
2. What word? (Group or individual)
3. “X” means (definition or synonym)
4. What does “x” mean (Group or individual)
5. Repeat step 4 until firm.
6. Use “x” in a sentence. (Group or individual)
7. Repeat step 6 for multiple students for each word
Lamented: you expressed deep sadness because of something.

"We missed the bus again," lamented the children.
Appendix D

Response Definitions and Coding Instructions:
OTR, Error Correction, Praise, Praise Around
Response Definitions and Coding Instruction: OTR

**Opportunities to Respond**: The teacher asks an academic question and directs it to a 
group or individual.

<table>
<thead>
<tr>
<th>OTR +</th>
<th>OTR -</th>
</tr>
</thead>
<tbody>
<tr>
<td>“Please read the first word aloud, Martha.”</td>
<td>The cue does not have a meaningful prompt</td>
</tr>
<tr>
<td>“What is 27 + 5, Everyone?”</td>
<td>The cue does not indicate if the response is directed to an individual or a group</td>
</tr>
<tr>
<td>“The prefix in the word reheat is ‘re.’ What is the prefix in the word reheat, row 1?”</td>
<td>The cue is stated as a question, instead of as a direction “Can you (or could you, would you, will you) spell father, Maria?”; instead of “spell father, Maria.” Also incorrect are “do you know,” or “how about you tell me the answer”</td>
</tr>
<tr>
<td>Poly means “many parts.” What does it mean? (This counts because the pronoun “it” follows the meaningful example directly. If “it” was more than once removed from the example it would not count).</td>
<td>“Raise your hand if you know what 27 + 5 is?” (not directed at a specific student or group)</td>
</tr>
<tr>
<td>“I would like everyone to read the list of words as I tap with my marker” (Each tap would count as a new cue in this example because “everyone” was specified at the beginning)</td>
<td>“Do you remember what we did yesterday, yes or no?” (There is no right or wrong answer to this- and no real academic content).</td>
</tr>
<tr>
<td>“Raise your hand when you know what 27 + 5 is?” followed by calling on a student who is not raising their hand, or on any student if all students are raising their hands (Raising hands is think time. Then the teacher is directing it towards a student of their choice instead of just calling on the kids who are willing to participate)</td>
<td></td>
</tr>
<tr>
<td>“Raise your hand if a paper clip is an example of a magnetic material” (raising hands is the answer to the question)</td>
<td></td>
</tr>
<tr>
<td>The question must be worded in a way to elicit an academic response (e.g., a yes no question with no right or wrong answer would NOT count)</td>
<td></td>
</tr>
<tr>
<td>If it is clearly implied who the question is directed at even if it is not explicitly stated (e.g., part of an ongoing string of questions to one student)</td>
<td></td>
</tr>
<tr>
<td>Note: it counts as an OTR even if a student does not respond or it is interrupted by behavior (as long as the OTR was completed before the interruption)</td>
<td></td>
</tr>
</tbody>
</table>
No Coding for OTRs

“Get out your notecards” (request must be academic, not behavioral)

“Point to problem number 5 on your math sheet”
(must require an active response that can be observed in the lab. The virtual students cannot accurately point to materials)

Kevin asks Maria what the Capital of Indiana is (must be teacher-directed, not student-directed)

Does that make sense?

Does anyone have a question?

<table>
<thead>
<tr>
<th>Question Type</th>
<th>Basic Knowledge</th>
<th>Application</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reading a word or repeating a word</td>
<td>Using a word in an original sentence</td>
<td></td>
</tr>
<tr>
<td>Reading a sentence</td>
<td>Giving an example or nonexample</td>
<td></td>
</tr>
<tr>
<td>Repeating a definition</td>
<td>Identifying an example or nonexample</td>
<td></td>
</tr>
<tr>
<td>Putting a definition in their own words</td>
<td>Sharing a personal experience related to a word</td>
<td></td>
</tr>
<tr>
<td></td>
<td>“Showing” or demonstrating what a word means (actions, gestures, etc.- <em>this is uncommon in the lab</em>)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Generating synonyms or antonyms</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Word replacement in a sentence</td>
<td></td>
</tr>
</tbody>
</table>
Response Definitions and Coding Instructions: Error Correction

**Error Correction**: When a student or makes an academic error the teacher models the correct answer, tests the student(s) on the initial question, and after one or more intervening responses delivers a retest. The complete cycle includes a model, test, and retest (each step in the cycle will be coded).

**Model**: Following an academic error by a student the teacher demonstrates the correct answer

<table>
<thead>
<tr>
<th>Model +</th>
<th>Model -</th>
<th>N/A</th>
</tr>
</thead>
<tbody>
<tr>
<td>“6+3 is 9”</td>
<td>Omitting the model</td>
<td>The model is n/a if the time runs out immediately following the student error</td>
</tr>
<tr>
<td>“This word is monkey”</td>
<td>Modeling an incorrect answer</td>
<td></td>
</tr>
<tr>
<td>“p-r-e says pre”</td>
<td>Treating a correct answer as an incorrect answer</td>
<td></td>
</tr>
<tr>
<td>“x stands for 2 in this equation”</td>
<td>Saying “The answer is____” without including the meaningful prompt, and it’s more than once removed from meaningful information.</td>
<td></td>
</tr>
<tr>
<td>“So close, (or nice try) but 6+3=9” (If the teacher uses very brief teacher talk, but none of it is negative, it may count as a Model +)</td>
<td>“No, the answer is _______ or No, 6 + 3 is 9” (must avoid saying “no,” “not quite,” etc.)</td>
<td></td>
</tr>
<tr>
<td>Saying “The answer is____” without including the meaningful prompt can count if it is only once removed from the meaningful information (e.g., if the teacher asked the question with the full prompt prior to the student’s incorrect response).</td>
<td>The word you are thinking of is quagmire. The word we are working on is quandary. (This is incorrect because the teacher starts with the non-example in their model instead of the example of the correct response).</td>
<td></td>
</tr>
<tr>
<td>“The word has an /l/ sound, instead of a /l/ sound, like this________ (drawing a comparison between the correct response and incorrect response can count if the teacher starts with the correct example instead of the non-example)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**Test:** Immediately following the model the teacher restates the question to the student or students who made the initial error.

<table>
<thead>
<tr>
<th>Test +</th>
<th>Test -</th>
<th>N/A</th>
</tr>
</thead>
<tbody>
<tr>
<td>“What is 6 + 3, Monique?”</td>
<td>Omitting the test</td>
<td>If the student immediately answers after the model the teacher usually doesn’t give a test. In this case mark it as n/a.</td>
</tr>
<tr>
<td>“Read word #3, everyone” (monkey)</td>
<td>If a behavior interrupts the teacher giving the test directly following the model, the teacher should repeat the model and then give the test <em>(some leniency allowed if it is a very brief interruption)</em></td>
<td>If the time runs out during or immediately following the model.</td>
</tr>
<tr>
<td>“p-r-e says what Francis?”</td>
<td>Tests that do not include a clear prompt (e.g. “what is the answer” instead of “What is 6 + 3?”) UNLESS the model was complete, and then it is considered only once removed.</td>
<td></td>
</tr>
<tr>
<td>Tests given without a model will still be coded (e.g., following an incorrect response the teacher asks the child to try it again).</td>
<td>If the Test is missing the components of an OTR it will be coded as Test- (i.e., asking it as a question, not specifying who it is directed at- unless it is clearly implied)</td>
<td></td>
</tr>
</tbody>
</table>

**NOTE:** *The test must also include the correct components of an OTR to be coded as Test +:*

- Not asked with can you, will you, would you, etc.
- No more than once removed from the meaningful information (if the model did not include the full question, the test must include the full prompt)
- Indicates who the question is directed to or it is clearly implied (teacher standing by desk, etc.)

If an individual student makes an error and the teacher tests the entire class (group response) it can still count as a Test +

When multiple tests are given right in a row just count the first test (or list the time range).

The test must match up with the initial error:

- If the student missed a definition they must be asked to give the same definition, etc.
- The only exception: If they were asked to give an example the teacher may ask for a sentence

Testing an individual or group who did not make the initial error. *(An exception to this rule in the lab: if the teacher cued the group and hears Monique’s voice, it will count as Test + if the teacher comes back only to Monique. Also, if one student makes a mistake the teacher can test the entire group).*

Testing the student(s) on the wrong question (e.g., asking for a definition when they made an error on giving a sentence).

**NOTE:** *If the teacher prompts the wrong question, or cues the wrong student or group, Test – will be coded, but the observer needs to also tally a correct OTR so it counts in the overall rate.*
**Delayed Test:** Following the test and one or more intervening response(s), the teacher returns to the initial question and directs it to the group or individual who made the initial error.

<table>
<thead>
<tr>
<th>Delayed Test +</th>
<th>Delayed Test -</th>
<th>N/A</th>
</tr>
</thead>
<tbody>
<tr>
<td>For an incorrect word in a reading passage, asking the student(s) to return to the beginning of the sentence or phrase can count as the intervening response, and when they encounter the target word it is the delayed test.</td>
<td>Omitting the delayed test</td>
<td>If the student incorrectly answers after the test, then mark the retest as n/a and list as a new error</td>
</tr>
<tr>
<td>After modeling and testing Marcus on math problem # 5 the teacher guides the group through problems #6 and #7. The teacher then returns to Marcus and re-asks questions 5.</td>
<td>The delayed test doesn’t include the clear prompt (e.g. “what is the answer,” instead of “what is 6 + 3”)</td>
<td></td>
</tr>
<tr>
<td>The delayed test must include the components of a correct OTR (as explained above)</td>
<td>The delayed test doesn’t indicate to who the question is directed at and it’s not clearly implied.</td>
<td></td>
</tr>
<tr>
<td>A delayed test can be coded as correct, even if it did not follow a correct model and/or test.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The delayed test must align with the initial error: • If the student missed a definition they must be asked to give the same definition, etc. • The only exception: If they were asked to give an example the teacher may ask for a sentence because these prompts are often used interchangeably.</td>
<td>The teacher models again before delivering the delayed test</td>
<td>If the teacher has less than 1 minute after the test and doesn’t deliver the delayed test, it is n/a</td>
</tr>
<tr>
<td>If an individual student makes an error and the teacher retests the entire class (group response) it can still count as a correct delayed test.</td>
<td>The delayed test doesn’t align with the initial question.</td>
<td></td>
</tr>
<tr>
<td>An intervening distraction can count as an “intervening response” if lasts 10 s or longer (e.g., dealing with a behavioral issue).</td>
<td>If the question asked right before the delayed test, even if to another student, is exactly the same as the delayed test (this functions as an additional model).</td>
<td></td>
</tr>
<tr>
<td>Note: Multiple delayed tests on the same question will be coded as extra OTRs, not as extra delayed tests.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Response Definitions and Coding Instructions: Praise

**Teacher Praise:** Positive teacher statements and gestures referring to student work or behavior.

**Specific Praise:** Positive teacher statements that make direct reference to an academic skill or behavior.

**General Praise:** Positive teacher statements and gestures that don’t reference a specific skill or behavior.

<table>
<thead>
<tr>
<th>Specific +</th>
<th>General +</th>
<th>Praise -</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Specific Academic Praise:</strong></td>
<td>&quot;Thank You&quot;</td>
<td>Statements that negatively state the behavior “Thank you for not tapping your pencil”</td>
</tr>
<tr>
<td>• Repeating the question and answer and pairing it with a praise statement (e.g., “you’re right 3 + 5 does equal 8”)</td>
<td>“Fantastic job”</td>
<td>Statements that directly follow a misbehavior or academic error:</td>
</tr>
<tr>
<td>• I’m so impressed with your accuracy on these math problems”</td>
<td>“Excellent”</td>
<td>o Thank you for raising your hand quietly (following a talk out)</td>
</tr>
<tr>
<td>• “Great improvement in reading rate, Marcus!”</td>
<td>Thumbs up or a high five</td>
<td>o Good job pronouncing that word (when a student mispronounces a word)</td>
</tr>
<tr>
<td>• “I loved how you sounded out the word and put it back together Francis.”</td>
<td>“Way to go”</td>
<td></td>
</tr>
<tr>
<td>• “Wonderful job adding the ones column first”</td>
<td>“Great example” <em>(doesn’t state what was great about it or connect it back to target word)</em></td>
<td></td>
</tr>
<tr>
<td>• “Wow! You lined up the decimal points perfectly”</td>
<td>“Correct, dog.” <em>(just repeating the word they read is general praise).</em></td>
<td></td>
</tr>
</tbody>
</table>

| Specific Behavioral Praise: | | |
| • “Nice work keeping our mouth and hands quiet during the math lesson.” | | |
| • ‘Thank you for raising your hand and waiting to be called on Monique” | | |
| • “Vince, fantastic reading volume!” | | |
| • “I like how table 3 is sitting quietly with their eyes forward.” | | |
| • “You are such a smooth reader Maria!” | | |
### No Coding for Praise

- Giving a reinforcer (points, marbles, tokens) will NOT count as praise unless it is paired with a praise statement.
- Negative or corrective statements
- Generic statements made with a neutral tone will not be coded as praise (e.g., a monotone “right” vs. “right!!” said with excitement and inflection)
- Statements that do not include a praise word or phrase will be considered feedback, and not praise, even if they are positive. For example, “you really thought through that answer,” will not count as specific praise unless it includes a praise word or phrase such as “good,” “thank you for....,” “I appreciate....,” etc.

### Number of Praise Statements: Consecutive statements count as one praise statement unless they occur more than 3 seconds apart, specify a behavior followed by a statement that specifies an academic response to different students (or vice versa), or are delivered to individual students while the teacher is moving from desk to desk.

<table>
<thead>
<tr>
<th>One Praise Statement</th>
<th>More than One Praise Statement</th>
</tr>
</thead>
<tbody>
<tr>
<td>o Consecutive praise statements with less than 3 seconds in between will count as one praise statement (see below):</td>
<td>* There are more than 3 seconds between consecutive statements</td>
</tr>
<tr>
<td>o Several general statements in a row will count as one general statement, unless*.</td>
<td>* The teacher is rotating to individual student desks while delivering the praise statements, or making direct contact with individual students if they are seated at a horseshoe table (e.g., such as high fives paired with each praise statement)</td>
</tr>
<tr>
<td>o Several specific statements in a row to different students (that specify the same OR different behavior) will count as one specific praise statement, unless*</td>
<td>* One statement clearly applies to an academic response of a student or group and the other statement clearly applies to behavior of a different student or group</td>
</tr>
<tr>
<td>o General and specific statements delivered consecutively will count as one specific statement, unless*</td>
<td>* A specific statement is directed at the targeted student in the praise around cycle (even if it follows a string of praise statements without 3 seconds in between)</td>
</tr>
</tbody>
</table>

*There are more than 3 seconds between consecutive statements
* The teacher is rotating to individual student desks while delivering the praise statements, or making direct contact with individual students if they are seated at a horseshoe table (e.g., such as high fives paired with each praise statement)
Response Definitions and Coding Instructions: Praise Around

**Praise Around Opportunity:** Disruptive student behavior that is observable by the teacher in the lab or classroom, for example talking out of turn, making fun of others, producing disruptive noises with mouth or objects.

<table>
<thead>
<tr>
<th>Examples</th>
<th>Nonexamples</th>
</tr>
</thead>
<tbody>
<tr>
<td>When a student exhibits off task behavior (as defined above) that lasts for five seconds or longer</td>
<td>Off-task behavior persists for less than 5 s and does not recur within 30 s or in relation to the same stimulus event.</td>
</tr>
<tr>
<td>If a student exhibits an off-task behavior for less than 5 s, but it recurs (from the same student) within 30 s and the total time of both incidences is more than 5 s, then the second time the behavior occurs would be a praise around opportunity (unless the teacher responds to the first incidence, in which case it would be coded as an opportunity at that point).</td>
<td>Behaviors that are strictly visual and go unnoticed by the teacher will not be coded (such as texting in class or sticking out a tongue to another student may not be observable by a teacher if s/he is writing on the board or looking down at a lesson plan), UNLESS the teacher sees it and begins a praise around cycle. IF the teacher starts the sequence for a behavior that is not verbal it counts as a PA opp.</td>
</tr>
<tr>
<td>If an off-task behavior coincides with the same stimulus event (e.g., talkouts during teacher-directed response opportunities, or makes fun of a student each time s/he gives a response) the second time the behavior occurs would be a praise around opportunity, even if it is more than 30 s later (unless the teacher responds to the first incidence, in which case it would be coded as an opportunity at that time)</td>
<td></td>
</tr>
<tr>
<td>The behavior must be the same class (or type) of behavior, not necessarily the exact same behavior (e.g., disruptive noises could include tapping, beat boxing, etc.)</td>
<td></td>
</tr>
</tbody>
</table>
**Praise Around Sequence:** In response to a praise around opportunity, the teacher should deliver 2 steps:

**Step 1:** The teacher gives a specific praise statement to another student who is exhibiting the desired behavior. The statement must reference the desired behavior that is *incompatible* with the problem behavior.

<table>
<thead>
<tr>
<th>Step 1+</th>
<th>Step 1-</th>
<th>N/A</th>
</tr>
</thead>
</table>
| To correctly deliver Step 1, the teacher must praise at least one student who is exhibiting the desired behavior.  
- Praise statements must be:  
  - **Specific**  
    (Directly states the behavior)  
  - **Stated positively**  
    (“Thank you for keeping your pencil flat” instead of “thank you for NOT tapping your pencil”)  
  - **Incompatible with the off-task behavior**  
    (“I appreciate you for having your cell phone put away and on silent.” This is *incompatible* with having a cell phone out and ringing because they can’t happen at the same time!)  
  
  *NOTE: These praise statements are ALSO coded as specific praise* | The teacher delivers a general praise statement to another student  
The teacher delivers a general or specific statement to another student who is also misbehaving (this would also be coded as praise -)  
The praise statement is negative instead of positive (“Thank you for not talking out while I’m teaching” instead of “thank you for sitting quietly.” Or “Thank you for not tapping your pen Francis, as opposed to “Thank you for keeping your materials quiet and flat on your desk, Francis”).  
The statement does not specify the behavior that is *incompatible* with the off-task behavior. (When a student is tapping their pencil the teacher says “Thank you for listening” instead of “thank you for keeping your materials flat and quiet.”)  
Praise statement is global instead of identifying at least one student who is engaged appropriately (e.g., “thank you to all who are....”)  
If the teacher doesn’t attempt a step 1 praise statement then it is coded as a – and “d.a.” (didn’t attempt) is written for the time | If the teacher responds by asking the misbehaving student an academic question, mark it as n/a academic engagement (n/a a.e.)  
If the time runs out immediately after the student exhibits the misbehavior |
**Step 2:** When the target student exhibits the target behavior the teacher delivers a specific praise statement to the student, which specifically references the desired behavior that is incompatible with the problem behavior.

<table>
<thead>
<tr>
<th><strong>Step 2 +</strong></th>
<th><strong>Step 2 -</strong></th>
<th><strong>N/A</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>To correctly deliver Step 2, the teacher must praise the target student once they exhibit the desired behavior.</td>
<td>Same rules for Step 1 - apply to Step 2 -.</td>
<td>If the teacher has completed step 1, but the student exhibits the same misbehavior again before the teacher has a chance to deliver step 2, then mark step 2 n/a and start a new PA opp.</td>
</tr>
<tr>
<td>- Praise statements must be:</td>
<td>- Statement is general</td>
<td>If the teacher completes step 1 with less than a minute left, then step 2 is n/a, unless the teacher delivers the step.</td>
</tr>
<tr>
<td>o <strong>Specific</strong> (Directly states the behavior)</td>
<td>o Statement is negative</td>
<td>If the teacher used academic engagement instead of step 1, then step 2 is also n/a (a.e.)</td>
</tr>
<tr>
<td>o <strong>Stated positively</strong> (“Thank you for keeping your pencil flat” instead of “thank you for NOT tapping your pencil”)</td>
<td>o Not incompatible with target behavior (or specific enough)</td>
<td></td>
</tr>
<tr>
<td>o <strong>Incompatible with the off-task behavior</strong> (“I appreciate you for having your cell phone put away and on silent.” This is incompatible with having a cell phone out and ringing because they can’t happen at the same time!)</td>
<td>Or, if the teacher doesn’t attempt the step</td>
<td></td>
</tr>
</tbody>
</table>

**Note:** These praise statements are ALSO coded as specific praise.
Appendix E

Video Handling Instructions
Video Handling Instructions

Capturing Video Data
1. Mount camera on tripod
2. Turn on camera
3. Point the camera so that the teacher (you) and the visual aids (words, pictures, etc.) are visible but NOT the students
4. Press the red button to begin recording
5. Record approximately 5-7 minutes of language arts instruction (preferably vocabulary instruction)
6. Press the red button to stop recording
7. Turn off camera

Uploading Video Data to Your Computer
1. Extend the USB connector from the left side of the camera
2. Plug the camera into a USB port on your computer
3. Click on the Flipvideo icon
   a) Open DCIM folder
   b) Open 100video folder
   c) Drag video to desktop to copy
4. Name the video Teacher #X (your assigned #), underscore (_) and the date (e.g., Sept17). So the complete title would look like this: #1_Sept17 (Your number is: )

Sending the Video to Melanie
1. Go to https://bft.usu.edu
2. Fill in the information as requested
   o For #4 you enter YOUR name and email address
   o For #5 you enter MY email address: melanie.dawson@aggiemail.usu.edu
3. Click “SUBMIT EMAIL” at the bottom of the page
4. Temporarily keep the video file in a folder on your computer as a back up (just in case the file transfer doesn’t work).
5. When you receive confirmation from Melanie that the file transfer was successful, DELETE the video from you camera and from your computer.

If you have questions please contact Melanie Dawson at melanie.dawson@aggiemail.usu.edu
Appendix F

Coding Sheets
Coding Sheet: OTR and Error Correction

<table>
<thead>
<tr>
<th>OTRs</th>
<th>+</th>
<th>-</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Application</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ERROR</th>
<th>Y or N</th>
<th>Y or N</th>
<th>Y or N</th>
<th>Y or N</th>
<th>Y or N</th>
<th>Y or N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EC Opp?</td>
<td>Y or N</td>
<td>Y or N</td>
<td>Y or N</td>
<td>Y or N</td>
<td>Y or N</td>
<td>Y or N</td>
</tr>
<tr>
<td>Model</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time or didn’t attempt</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time or didn’t attempt</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Retest</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time or didn’t attempt</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Coding Sheet: Praise and Praise Around

#### General Praise
(tally)

<table>
<thead>
<tr>
<th>Specific Praise</th>
<th>+ or -</th>
<th>Time</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Specific Praise</th>
<th>+ or -</th>
<th>Time</th>
</tr>
</thead>
</table>

### BEHAVIOR

<table>
<thead>
<tr>
<th>Student</th>
<th>Time range (5+s or recurs)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PA Opp?</strong></td>
<td>Y or N</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Step 1</strong></th>
<th>+ - n/a</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Time or didn’t attempt</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Step 2</strong></th>
<th>+ - n/a</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Time or didn’t attempt</td>
</tr>
</tbody>
</table>
Appendix G

Social Validity Survey
Social Validity Survey

Please answer the following questions as honestly as possible:

1) The TeachLivE classroom feels like a real classroom.
   Strongly Agree   Agree   Undecided   Disagree   Strongly Disagree

2) During my interaction with the TeachLivE students, I began to understand their different personalities.
   Strongly Agree   Agree   Undecided   Disagree   Strongly Disagree

3) During my interaction with the TeachLivE students I thought of them as real kids.
   Strongly Agree   Agree   Undecided   Disagree   Strongly Disagree

4) I was able to interact with the TeachLivE students like I would in a physical classroom.
   Strongly Agree   Agree   Undecided   Disagree   Strongly Disagree

5) My visual proximity to the TeachLivE students (zooming in when I approached a student; zooming out when I moved away) enhanced my interactions in the classroom.
   Strongly Agree   Agree   Undecided   Disagree   Strongly Disagree

6) The scenarios I encountered during the training sessions were similar to situations I encounter in my own classroom.
   Very Similar   Somewhat Similar   Undecided   Somewhat Dissimilar   Very Dissimilar

Why?

7) The scenarios I encountered during the assessment sessions were similar to situations I encounter in my own classroom.
   Very Similar   Somewhat Similar   Undecided   Somewhat Dissimilar   Very Dissimilar

Why?
Training Sessions

**Lab Preparation:**

8) The **training videos** were helpful in defining and describing each target skill.

<table>
<thead>
<tr>
<th>Strongly Agree</th>
<th>Agree</th>
<th>Undecided</th>
<th>Disagree</th>
<th>Strongly Disagree</th>
</tr>
</thead>
</table>

9) The **handouts** were helpful in defining and describing each target skill.

<table>
<thead>
<tr>
<th>Strongly Agree</th>
<th>Agree</th>
<th>Undecided</th>
<th>Disagree</th>
<th>Strongly Disagree</th>
</tr>
</thead>
</table>

10) The **quizzes** were helpful in solidifying my knowledge of each target skill.

<table>
<thead>
<tr>
<th>Strongly Agree</th>
<th>Agree</th>
<th>Undecided</th>
<th>Disagree</th>
<th>Strongly Disagree</th>
</tr>
</thead>
</table>

**Lab Sessions:**

11) **Hearing positive feedback from my peers** after each turn help me understand my strengths on the target skills.

<table>
<thead>
<tr>
<th>Strongly Agree</th>
<th>Agree</th>
<th>Undecided</th>
<th>Disagree</th>
<th>Strongly Disagree</th>
</tr>
</thead>
</table>

12) **Verbalizing my goals for improvement** after each teaching turn helped me improve my performance on the target skills.

<table>
<thead>
<tr>
<th>Strongly Agree</th>
<th>Agree</th>
<th>Undecided</th>
<th>Disagree</th>
<th>Strongly Disagree</th>
</tr>
</thead>
</table>

13) **Receiving verbal feedback from the instructor** after each teaching turn helped me improve my performance on the target skills.

<table>
<thead>
<tr>
<th>Strongly Agree</th>
<th>Agree</th>
<th>Undecided</th>
<th>Disagree</th>
<th>Strongly Disagree</th>
</tr>
</thead>
</table>

14) **Providing verbal feedback to my peers** after their teaching turns helped me improve my performance on the target skills.

<table>
<thead>
<tr>
<th>Strongly Agree</th>
<th>Agree</th>
<th>Undecided</th>
<th>Disagree</th>
<th>Strongly Disagree</th>
</tr>
</thead>
</table>
15) **Hearing the feedback** given to other teachers during the training session helped me improve my performance on the target skills.

<table>
<thead>
<tr>
<th>Strongly Agree</th>
<th>Agree</th>
<th>Undecided</th>
<th>Disagree</th>
<th>Strongly Disagree</th>
</tr>
</thead>
</table>

16) The **written data sheets** I received at the end of each session helped me improve my performance on the target skills.

<table>
<thead>
<tr>
<th>Strongly Agree</th>
<th>Agree</th>
<th>Undecided</th>
<th>Disagree</th>
<th>Strongly Disagree</th>
</tr>
</thead>
</table>

17) **Verbalizing the connections between the lab and my own classroom** helped me improve my teaching with my own students.

<table>
<thead>
<tr>
<th>Strongly Agree</th>
<th>Agree</th>
<th>Undecided</th>
<th>Disagree</th>
<th>Strongly Disagree</th>
</tr>
</thead>
</table>

**Lesson Preparation:**

18) On average, how much time did you spend each week preparing the **training lessons** (Lessons A, B, & C)?

- 0 mins
- 1-5 mins
- 6-10 mins
- 11-15 mins
- 15 mins or more

19) Did you work on your own or with others in preparing the **training lessons** (Lessons A, B, & C)?

(Circle one):  I prepared alone  I prepared with others

20) On average, how much time did you spend each week preparing the **assessment lessons** (Lessons 1-14)?

- 0 mins
- 1-5 mins
- 6-10 mins
- 11-15 mins
- 15 mins or more

21) Did you work on your own or with others in preparing the **assessment lessons** (Lessons 1-14)?

(Circle one):  I prepared alone  I prepared with others
22) The example teaching formats were helpful in preparing the training and assessment lesson plans.

   Strongly Agree   Agree   Undecided   Disagree   Strongly Disagree

23) The additional support materials (pictures and stories) were helpful in preparing the training and assessment lesson plans.

   Strongly Agree   Agree   Undecided   Disagree   Strongly Disagree

24) The lesson plans and example teaching formats were helpful in my own classroom.

   Strongly Agree   Agree   Undecided   Disagree   Strongly Disagree

   **Online Discussions**

25) The online discussions helped me make important connections between the skills I practiced in lab and my daily classroom teaching.

   Strongly Agree   Agree   Undecided   Disagree   Strongly Disagree

26) I enjoyed the interaction with my peers during the online discussions. This format facilitated an ongoing professional learning community with my colleagues.

   Strongly Agree   Agree   Undecided   Disagree   Strongly Disagree

   **Classroom Recordings**

27) The necessary materials and instructions were provided to make the classroom recordings as simple and straightforward as possible.

   Strongly Agree   Agree   Undecided   Disagree   Strongly Disagree

28) Submitting my video to Melanie was simple and took a reasonable amount of time

   Strongly Agree   Agree   Undecided   Disagree   Strongly Disagree
Final Question

29) If I offered you $50 for mastering a new teaching skill in your classroom within a week, would you choose to first practice the skill in TeachLivE and then apply the skill in your classroom, practice the skill with a peer and then apply the skill in your classroom, OR apply the skill in your classroom without practicing in TeachLivE or with a peer?

(circle one):  
- Practice in *TeachLivE*, then apply skill in classroom
- Practice *with a peer*, then apply skill in classroom
- Apply skill in classroom without practice in TeachLivE or with a peer

**Additional Comments** about your experience this semester:
Appendix H

Training Video URLs
Training Video URLs

Opportunities to Respond:
http://youtu.be/PDv2yJdwPLk

Error Correction:
http://www.youtube.com/watch?v=AGNhHzHeI-pI

Praise:
http://youtu.be/QiQGph5hvuE

Praise Around:
http://youtu.be/KPrCAIqKgVI
Appendix I

Instructional Handouts: OTR, Error Correction, Specific Praise, Praise Around
Instructional Handout: OTR

Target Skill: OPPORTUNITIES TO RESPOND

Definition of Opportunity to Respond (OTR): The teacher asks an academic question and indicates if it is directed to the group or to an individual student. (If it is a group response the teacher needs to include a signal as well).

Goal: 4 (or more) OTRs per minute

<table>
<thead>
<tr>
<th>Examples</th>
<th>Non-Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>“Please read the first word aloud, Martha.”</td>
<td>“Read this word.” (no indication if it is group or individual)</td>
</tr>
<tr>
<td>“What is 27 + 5, Everyone?”</td>
<td>“By raise of hand, who knows what 27 + 5 is?” (Again, this is incorrect because it is not directed at a specific student or group)</td>
</tr>
<tr>
<td>“What is the prefix in the word reheat, row 1?”</td>
<td>“Can you spell father, Francis.” Also, “will you ______,” “could you,” “do you know ______,” or “how about you answer question #5 ______.” (These are stated as questions, and open up the possibility of a student saying they can’t or won’t answer. Instead the teacher should directly state, “spell father, Francis”).</td>
</tr>
<tr>
<td>“Poly means ‘many parts.’ What does “poly” mean, Sam?”</td>
<td>Joe asks Brittany what the capital of Indiana is, and Brittany responds correctly (student-directed, not teacher-directed)</td>
</tr>
<tr>
<td>“Please read the next sentence for us, Jimmy.”</td>
<td>I’d like the entire class to line up at the door (This is behavioral, not academic)</td>
</tr>
<tr>
<td>“Think of a sentence using the word ‘persuade.’ (Think time). Sarah, please share your sentence with us.”</td>
<td>What did you do this weekend Sarah? (This is not an academic question)</td>
</tr>
<tr>
<td>“I would like everyone to read the list of words as I tap with my marker” (Each tap would count as a new response opportunity in this example because “everyone” was specified at the beginning of the sequence)</td>
<td>The question does not include a meaningful prompt or the task is not clear.</td>
</tr>
<tr>
<td>“Do you remember what we did yesterday class, yes or no?” (There is no right or wrong answer to this- and no real academic content)</td>
<td></td>
</tr>
</tbody>
</table>

**Question Type:** This semester we will focus on two essential question types: **Basic knowledge** questions (low level) and **application** questions (high level). Establishing basic knowledge supports student success on application questions. Therefore, we’d like you to attempt the following:

**Goal:**
- *FIRST,* establish **basic knowledge. THEN ask students to apply** their knowledge.
- Ask approximately 50% of each question type

The table below provides examples of each question type for a vocabulary lesson.

<table>
<thead>
<tr>
<th>Basic Knowledge (low level)</th>
<th>Application (high level)</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Repeating the word after the teacher</td>
<td>• Using the word in an original sentence</td>
</tr>
<tr>
<td>• Reading the word</td>
<td>• Giving an example or non-example of the word</td>
</tr>
<tr>
<td>• Reading a sentence with the word</td>
<td>• Identifying an example or non-example</td>
</tr>
<tr>
<td>• Repeating or reading the definition of the word</td>
<td>• Sharing a personal opinion or experience related to the word</td>
</tr>
<tr>
<td>• Rephrasing the definition in their own words</td>
<td>• “Showing”, or demonstrating what the word means (actions, gestures, etc.)</td>
</tr>
<tr>
<td></td>
<td>• Generating synonyms or antonyms for the word</td>
</tr>
<tr>
<td></td>
<td>• Replacing the word in a sentence</td>
</tr>
</tbody>
</table>
Instructional Handout: Error Correction

**Target Skill: ERROR CORRECTION**

**Definition of Error Correction:** When a student makes an academic error the teacher models the correct answer, tests the student on the initial question, and retests the student after a short delay. Each step is defined as follows:

- **Model:** In response to an academic error the teacher repeats the question and demonstrates the correct answer to the question.
- **Test:** Immediately following the model, the teacher restates the initial question and directs it to the group or individual who made the error.
- **Delayed Test:** Following the model/test, the teacher moves on to one or more intervening question(s), and then returns to the initial question and directs it to the group or individual who made the error.

**Goal:** To deliver all error correction steps in response to each academic error.

<table>
<thead>
<tr>
<th>Example</th>
<th>Non-Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>“6 plus 3 is 9.”&lt;br&gt;The capital of Indiana is Indianapolis.&lt;br&gt;“That word is monkey”&lt;br&gt;“p-r-e says pre.”&lt;br&gt;(Notice the teacher includes the full prompt as part of the model instead of saying “the answer is ______”)</td>
<td>“No. That’s wrong.” <em>(This statement is negative, and does not provide corrective feedback)</em>&lt;br&gt;“The answer is _____” <em>(Does not include the full prompt)</em>&lt;br&gt;Modeling an incorrect answer&lt;br&gt;Ignoring or not noticing an error&lt;br&gt;Moving straight to a test without a model, such as saying, “try it again.”</td>
</tr>
<tr>
<td>Test</td>
<td>Delayed Test</td>
</tr>
<tr>
<td>---------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------</td>
</tr>
</tbody>
</table>
| The following examples show how the test immediately follows the model. Examples of each test are in **boldface**):  

“6 plus 3 is 9.” **What is 6 plus 3, Veronica?**

The capital of Indiana is Indianapolis. **What is the capital of Indiana, Mark?**

“That word is monkey. **Everyone, the word is what?**” (signal for group response)

“p-r-e says pre.” **P-r-e says what, table 2?**” (signal for group response)

Note: If a teacher is interrupted after giving the model (e.g., by a behavioral issue), he/she must repeat the model again before delivering the test.

“6 plus 3 is 9. **What was the answer to that problem?**” (the original prompt must include meaningful information, not just a generic cue to answer the question).

Asking for a raise of hand on the test

Asking the question using “Can you,” “Will you,” “Would you,” or “Could you?” (The test must include all components of a correct OTR practiced previously)

Forgetting the test after the model

Asking the student a different question that was not the original prompt.

Testing a student or group of students who did not make the initial error.

---

<table>
<thead>
<tr>
<th>Delayed Test</th>
<th>Test</th>
</tr>
</thead>
</table>
| After a model and test with Veronica on 6+3=9, the teacher asks the class to answer another math fact, and then says, **What is 6 plus 3, Veronica?**

After a model and test on the word monkey, the teacher moves on to two more words on the word list and then points to **monkey** and says, **“Everyone, what is this word?”** (signal for group response).

After a model/test in a reading passage, the teacher prompts the student to go back to the beginning of the sentence or phrase (reading the words in the sentence prior to the target word counts as the intervening response)

““What was that answer?” (must include the meaningful prompt)

Asking for a raise of hand, or asking the question using “Can you,” “Will you,” “Would you,” or “Could you?”

Testing again immediately after the initial test (must be delayed)

Forgetting the delayed test

Providing a delayed test for a student or group who did not produce the initial error
Instructional Handout: Specific Praise

**Target Skill: SPECIFIC PRAISE**

**Definition of Praise:** Positive teacher statements and gestures referring to student work or behavior. Praise can be specific or general:

**Specific Praise:** Positive teacher statements that directly reference the academic skill or behavior being praised. (This is the type of praise we are primarily targeting because it is so much more powerful!)

**General Praise:** Positive teacher statements and gestures that don’t specifically reference student work or behavior.

**Goal:**
- 4 (or more) total praise statements per minute.
- More than half of your praise statements should be specific (aim for 60% or higher)
  - **OR** 2+ specific praise statements per minute

<table>
<thead>
<tr>
<th>Specific Academic Praise</th>
<th>Specific Praise</th>
<th>Non-Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>“Excellent! You’re right 3 + 5 does equal 8.” (repeating the question and answer and pairing it with a praise statement.)</td>
<td>“Good job,” “Way to go,” or “You’re on the ball.” (These statements do not directly reference an academic skill or behavior)</td>
<td>“Thank you for not tapping your pen Dominic.” (This is stated negatively, and draws attention to the undesirable behavior. Change it to “Thank you for keeping your materials quiet and flat on your desk.”)</td>
</tr>
<tr>
<td>“I’m so impressed with your accuracy on these math problems.”</td>
<td>“Great improvement in reading rate, Nathan!”</td>
<td>After Calvin left his seat, saying “I appreciate you staying in your seat Calvin” (Praise should follow the desired behavior, not misbehavior.)</td>
</tr>
<tr>
<td>“I loved how you sounded out the word and put it back together Fran.”</td>
<td>“Wonderful job adding the ones column first.”</td>
<td>Statements that do not include a praise word or phrase will be considered feedback, and not praise, even if they are positive. For example, “you really thought through that answer,” will not count as specific praise unless it is paired with a praise word or phrase such as “excellent,” “thank you for.....,” “I appreciate.....,” etc.</td>
</tr>
<tr>
<td>“Wow! You lined up the decimal points perfectly.”</td>
<td>“Fabulous example of the word ‘transparent.’” (This counts as specific praise because it connects the task and the target vocabulary)</td>
<td>Distributing points or tickets without stating what the student did to earn them.</td>
</tr>
</tbody>
</table>

**Specific Behavioral Praise:**
- “Nice work keeping your mouth
and hands quiet during the math lesson.”

- “Thank you for raising your hand and waiting to be called on Mandy.”
- “Joseph, fantastic reading volume!”
- “I like how table 3 is sitting quietly with their eyes forward.”
- Distributing points or tickets and pairing them with a specific praise statement

“Keep your hands to yourself.”
(This is a corrective comment, not praise).

<table>
<thead>
<tr>
<th>Examples</th>
<th>Non-Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>“Thank You”</td>
<td>“Please sit down.” (This is a corrective comment, not praise.)</td>
</tr>
<tr>
<td>“Fantastic job!”</td>
<td>“Way to go” (said sarcastically)</td>
</tr>
<tr>
<td>“Excellent”</td>
<td>“Fantastic job figuring out the remainder to that division problem.” (This is not general because it references a specific skill.)</td>
</tr>
<tr>
<td>“Great example.” (This statement does not connect back to the target word and/or point out what is good about the example).</td>
<td></td>
</tr>
<tr>
<td>Thumbs up or high five (positive gestures are general praise)</td>
<td></td>
</tr>
<tr>
<td>“Yes, luxury.” (just repeating the word the child read is considered general praise).</td>
<td></td>
</tr>
</tbody>
</table>
Instructional Handout: Praise Around

Target Skill: PRAISE AROUND

Definition: When a student misbehaves the teacher praises another student for exhibiting the desired behavior. Then, when the target student exhibits the desired behavior, the teacher praises the target student. Each step in the praise around sequence is defined in further detail below.

First, a teacher must identify an opportunity to use praise around. There is an opportunity to praise around any time a student exhibits persistent and/or recurring misbehavior and another student is engaged appropriately. The teacher should then deliver the following steps:

Step 1: Praising another student for the exhibiting the desired behavior. The teacher ignores the misbehavior and delivers a specific praise statement to another student who is exhibiting the desired behavior. The praise statement must specifically reference the desired behavior that is incompatible with the problem behavior.

Step 2: Praising the target student for exhibiting the desired behavior. Once the target student engages in the desired behavior, the teacher delivers a specific praise statement to that student within approximately 10-20 seconds. The praise statement must specifically reference the desired behavior that is incompatible with the problem behavior.

Goal: To deliver all praise around steps in response to each persistent or recurring misbehavior.

<table>
<thead>
<tr>
<th>Praise Around Opportunity</th>
<th>Examples</th>
<th>Non-Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>When a student misbehaves for 5 seconds or longer</td>
<td>When all students are on task (no need for praise around)</td>
</tr>
<tr>
<td></td>
<td>When a student misbehaves for less than 5 seconds, but it happens again. <em>(The teacher may choose to ignore the first occurrence of a mild misbehavior, but should address the problem if it recurs).</em></td>
<td>When the entire class is exhibiting the same problem behavior at the same time (if so, there is no opportunity to praise the target behavior. This is rare!)</td>
</tr>
</tbody>
</table>
|                           | Typical problem behaviors to watch for:  
|                           | • Talking out, engaging in side conversations  
|                           | • Making fun of others  
|                           | • Tapping pen/pencil  
|                           | • Making disruptive noises with mouth, objects, or hands  
<p>|                           | • Playing with objects, such as a cell |</p>
<table>
<thead>
<tr>
<th>top 1: Praising another student for the desired behavior</th>
<th>Examples</th>
<th>Non-Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>When Timothy is tapping his pens on the desk the teacher says, “Thank you Veronica for keeping your materials quiet.”</td>
<td>When Timothy is tapping his pens on the desk the teacher says, “I need you to keep your pen flat and quiet Timothy.” <em>The teacher delivered a corrective statement instead of ignoring the behavior and praising another student.</em></td>
<td></td>
</tr>
<tr>
<td>When Katie and Sarah are whispering to one another about a math assignment the teacher says, “Fantastic job following my directions to work quietly on your own, Betty.”</td>
<td>When Katie and Sarah are whispering to one another about a math assignment the teacher says “Nice job Betty.” <em>This statement does not specify the target behavior. It is general praise.</em></td>
<td></td>
</tr>
<tr>
<td>When Samuel is texting in class the teacher says, “Thank you Row 1 for having all personal items and electronics put away in your backpack.”</td>
<td>When Samuel is texting in class the teacher says, “Thank you Row 1 for NOT having your cell phones out during class.” <em>This is stated negatively and draws attention to the undesirable behavior.</em></td>
<td></td>
</tr>
<tr>
<td>When Missy is talking out the teacher says, “I appreciate how Greg is listening quietly to the vocabulary examples.”</td>
<td>When Missy is talking out the teacher says, “Thanks for having your eyes forward Greg.” <em>Eyes forward is not really the target behavior. Also, it is possible for a student to have their eyes forward AND talk out at the same time, therefore the behaviors are not incompatible.</em></td>
<td></td>
</tr>
</tbody>
</table>
| Step 2: Praising the target student for the desired behavior | Timothy stops tapping his pencil after the teacher praised Veronica. The teacher turns to Timothy and says, “Wow Timothy, I appreciate you putting your pencil down and sitting with your materials flat and quiet.”

Katie and Sarah stop whispering about the math assignment, and Katie raises her hand, when the teacher praised Sally. The teacher immediately says, “Thank you Katie and Sarah for beginning to work on your own and for raising your hand when you have a question about the work.”

Samuel puts his phone away after the teacher praised Row 1. The teacher says, “Thank you Samuel for putting all personal supplies in your backpack during class so you can pay attention to the lesson.”

Missy stops talking when the teacher praises Greg for listening quietly. The teacher immediately walks to Missy’s desk and says, “I really appreciate your quiet attention while I’m teaching Missy. It helps you and others learn our vocabulary words.”

When Timothy stops tapping his pencil the teachers says “thank you for NOT tapping your pencil.” *(this statement emphasizes the negative behavior)*

When Katie and Sarah stop whispering the teacher says, “Thank you Katie and Sarah. Good job.” *(must be specific praise, not general)*

Samuel puts his phone away, and the teacher says, “Thank you for following directions Samuel.” *(Following directions is too broad and does not clearly specify the desired behavior).*

When Missy stops talking the teacher ignores Missy, or gets back to her a long time afterwards. *(The teacher must praise the target student as soon as it is clear that he/she is behaving, otherwise the student will likely return to the problem behavior. Exact times will vary, but during lab practice we’d like you to get back to the student within 10-20 seconds.)*

Other Non-examples:
- The statement doesn’t specify the behavior that is *incompatible* with the target behavior.
- The statement is sarcastic, negative, or corrective.
Appendix J

Quizzes
1) According to the video and the handout, what is the definition of a teacher-directed opportunity to respond? (1 point)

2) In the first section of the video, Leanne states, “the more opportunities students have to ______________ in learning the more quickly they will ______________ the material.” (2 points- 1 per blank)

3) What are the two types of response opportunities (as defined in the video)? Give an original example of each (4 points- 1 per box)

<table>
<thead>
<tr>
<th>Type of Response Opportunity</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4) Why is it important to ask the question first, and then indicate who you want to answer? (1 point)

5) Why is it a problem to ask “who knows the answer” to a question or to wait for a raise of hand? (1 point)

6) Please provide two original non-examples of response opportunities below. Explain why they are non-examples. (2 points)

7) What are the two things a teacher must do regardless of if it is an individual response opportunity or a group response opportunity? (2 points)

8) What is the one thing the teacher must add when asking for a group response? (1 point)

9) Give two examples of an effective signal. (2 points)
10) Why is **unison** responding important? (1 point)

11) How many response opportunities should a teacher deliver per minute? (1 point)
   In general you should aim for:
   For very short/concise answers:

12) According to the video, why is it important to maintain a high rate of response opportunities in the classroom? (1 point)

13) The **handout** outlines two question types. Please provide the **name** of each question type and give **at least two examples** of each question type. (4 points - 1 point per box)

<table>
<thead>
<tr>
<th>Question Type</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low Level, or</td>
<td>1)</td>
</tr>
<tr>
<td>_______________________</td>
<td>2)</td>
</tr>
<tr>
<td>High level or</td>
<td>1)</td>
</tr>
<tr>
<td>_______________________</td>
<td>2)</td>
</tr>
</tbody>
</table>

14) Using the information from the handout, outline the goals for strategically delivering **both** types of questions (2 points):

   Approximately what percentage of each question type should you aim to ask?

   In what order should you ask the questions?

   Why?
Error Correction Quiz

1) According to the handout, what is the basic definition of error correction? (You do not need to include the definition of each step for this question). (1 point)

2) In the first section of the video, Leanne states that “practice can be ________________ if students are practicing ________________ without being ________________.” (3 points- 1 point per blank).

3) List each step of the error correction sequence and then give the definition for each step. (6 points- 1 point per box)

<table>
<thead>
<tr>
<th>Step</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 1:</td>
<td></td>
</tr>
<tr>
<td>Step 2:</td>
<td></td>
</tr>
<tr>
<td>Step 3:</td>
<td></td>
</tr>
</tbody>
</table>

4) Please give an original example of a model. Please include information about the student error and then demonstrate how the teacher would correctly model. (2 points)

5) Please give two original non-examples of a model. Please explain why they are non-examples (2 points)

6) Please give an original example of a test. In your answer please demonstrate how the test will immediately follow the model. (2 points)

7) Please give two original non-examples of a test. Please explain why they are non-examples (2 points)

8) What is the purpose of the delayed test? (1 point)

9) Please give an original example of a delayed test. Please demonstrate how the delayed test will follow one or more intervening response(s). (2 points)
10) Please give two original non-examples of a delayed test. Please explain why they are non-examples. (2 points)

11) How often should you correct errors in the classroom? (1 point)

12) Throughout the video Melanie and Leanne discuss WHY error correction is so important in the classroom. Please explain at least one reason that error correction is such a crucial teacher skill in the classroom. (1 point)
Praise Quiz

1) According to the video and the handout, what is the definition of teacher praise? (1 point)

2) In the first section of the video, Melanie states that praise “will show your students that you will give ___________________ ___________________ for the ___________________ things that they do.” (3 points- 1 point per blank)

3) List, define, and give an example of the two basic kinds of praise. (6 points- 1 point per box)

<table>
<thead>
<tr>
<th>Kind of Praise</th>
<th>Definition</th>
<th>Example (please provide original examples, not those on the video and handout)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4) Specific praise can be academic or behavioral. Please give an original example of each. (2 points)

Specific Academic:

Specific Behavioral:

5) The video explains three different situations/comments that are non-examples for praise. Using this information, write three original non-examples of praise (one of each type) and explain why they are non-examples. (3 points)

6) Paraphrase the “cake analogy” used in the video (1 point). What are the three main concepts taught through this analogy (3 points)? Why do you think these are important things to remember when delivering praise in the classroom (1 point)? (5 points total)

7) How many praise statements are recommended per minute during teacher-directed instruction? (1 point)
8) What percentage of specific praise is recommended? Why is specific praise more powerful than general praise? (2 points)

9) Throughout the video Melanie and Leanne discuss WHY praise is so important in the classroom. Provide at least two reasons that praise is such a crucial skill in the classroom. (2 points)
Praise Around Quiz

1) According to the handout, what is the basic definition of praise around? (You do not need to include the definition of each step on this question). (1 point)

2) In the first section of the video, Melanie states that “praise around is a non-confrontational technique for increasing and ___________________ desired ________________ and decreasing ________________ behaviors in your classroom. (1 point for all 3 blanks correct).

3) List each step of the praise around cycle and then give the definition for each step. (6 points- 1 point per box)

<table>
<thead>
<tr>
<th>Step</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>First, a teacher must identify if there is an ________________ to use ________________ (from handout)</td>
<td></td>
</tr>
<tr>
<td>Step 1:</td>
<td></td>
</tr>
<tr>
<td>Step 2:</td>
<td></td>
</tr>
</tbody>
</table>

4) Some behaviors are so mild or brief that a teacher doesn’t need to praise around. What is the “rule of thumb” Melanie describes in the video to determine if you should praise around? (It is also listed on the handout). (2 points)

5) List three typical problem behaviors you see in your own classroom that would be opportunities to use the praise around strategy. (3 points)

6) Using one of the problem behaviors you listed above, please demonstrate how you would praise another student for exhibiting the desired behavior (Step 1). (1 point)
7) Using the scenario you chose in question #6, please give two original non-examples of praising another student for exhibiting the desired behavior (Step 1). Explain why they are non-examples (2 points)

8) Using a different misbehavior than you used in Question #6, please demonstrate how you would praise the target student once he/she stopped misbehaving and starting exhibiting the desired behavior (Step 2). (1 point)

9) Please give two original non-examples of praising the target student for exhibiting the desired behavior (Step 2). Explain why they are non-examples (2 points)

10) What is the recommendation for how quickly should you get back to the target student to praise him/her once he/she is exhibiting the target behavior? And, WHY is this important? (2 point)

11) What does it mean to praise a student for exhibiting the behavior that is incompatible with the problem behavior (Step 1 and Step 2)? WHY is this important? (2 points)

12) Why do you think praise around should be the first strategy for dealing with typical problem behaviors in the classroom? (2 point)
Appendix K

TeachLivE Session Objectives Form
TeachLivE Session Objectives Form

**TLE TeachLivE™ Session Objectives**

**SESSION PLANNING TEMPLATE**

<table>
<thead>
<tr>
<th>REQUESTER INFORMATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name:</td>
</tr>
<tr>
<td>University:</td>
</tr>
<tr>
<td>Department:</td>
</tr>
<tr>
<td>Class:</td>
</tr>
<tr>
<td>Date:</td>
</tr>
<tr>
<td>Duration Period in EST: to</td>
</tr>
<tr>
<td>Facilitator Name:</td>
</tr>
<tr>
<td>Phone Number:</td>
</tr>
<tr>
<td>Skype Address:</td>
</tr>
</tbody>
</table>

**DESCRIPTION OF THE SESSION**

Highlight Your Chosen Focus-

**Avatars to be utilized in session:**

Middle School High School

**The following avatars require a ONE month notice prior to scheduling:**

ELL Adult Avatar

Is this session a demonstration? YES NO

Will session be recorded? YES, for research (See form below)
YES, for media (Please contact TeachLivE for approval)
YES, for student feedback (See from below)
NO

Session will focus on: CONTENT or PEDAGOGY or BOTH

Session Details: _________________________________________________________

Number of Participants: ______

Scheduled activities for participants in session:

__________________________________________

*** LESSON PLANS AND RUNNING ORDER (IF APPLICABLE) MUST BE SENT ONE WEEK BEFORE THE SCHEDULED SESSION***
<table>
<thead>
<tr>
<th>BEHAVIOR LEVEL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Choose the Preferred Behavior Escalation Level <strong>0</strong></td>
</tr>
</tbody>
</table>

Behavior Level: 0----------1--------2--------3--------4--------5

0 = no classroom misbehavior

1 = mild misbehavior -> distraction, fidgeting, inattention at *low* frequency

2 = mild/moderate misbehavior -> distraction, fidgeting, inattention, mild resistance at *low* frequency

3 = moderate misbehavior -> distraction, fidgeting, inattention, resistance at *medium* frequency

4 = moderate / intense misbehavior -> distraction, fidgeting, inattention, resistance, bullying behavior at *medium* frequency

5 = intense misbehavior -> distraction, fidgeting, inattention, resistance, bullying behavior at *high* frequency including personal attacks towards teacher and students

<table>
<thead>
<tr>
<th>1ST GOAL/OBJECTIVE FOR PARTICIPANTS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Description:</strong></td>
</tr>
<tr>
<td><strong>Measurement:</strong></td>
</tr>
<tr>
<td>Importance:      x Essential □ Important □ Desirable</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>2ND GOAL/OBJECTIVE FOR PARTICIPANTS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Description:</strong></td>
</tr>
<tr>
<td><strong>Measurement:</strong></td>
</tr>
<tr>
<td>Importance: □ Essential □ Important □ Desirable</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>3RD GOAL/OBJECTIVE FOR PARTICIPANTS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Description:</strong></td>
</tr>
<tr>
<td><strong>Measurement:</strong></td>
</tr>
<tr>
<td>Importance: □ Essential □ Important □ Desirable</td>
</tr>
</tbody>
</table>

Please send this form and attach any additional required materials -at least a week prior to your session.
Appendix L

TeachLivE Session Instructions
TeachLivE Session Instructions

Welcome/Introduction

Skill Focus:
- “Tonight is (state the date) and we will be practicing (state the target skill).”
- Distribute the handout for the skill(s), or verify that the teachers brought their own.
- “The definition of (target skill) is (read from handout)”
- “The goal is ___________”
  - Opportunities to Respond: at least 4 opportunities per minute (approximately half basic and half application questions)
  - Error Correction: to complete each step of the error correction sequence for each academic error
  - Specific Praise: at least 4 overall praise statements per minute with more than half (2+ or more than 50%) specific
  - Praise Around: to complete each step of the praise around cycle each time there is a persistent or recurring misbehavior

Skill Maintenance: (If applicable):
- “We will also maintain proficiency with (state prior target skill or skills).”
- “As a reminder, the goal for this skill is (refer to list above)”

Practice Turns

Round 1:
- Each participant will teach for one 2-minute turn
- Rotate participants until everyone has had a turn
- During this round prompting IS allowed

Round 2:
- Each participant will teach for one 2-minute turn
- Rotate participants until everyone has had a turn
- During this round prompting IS allowed

Round 3:
- Each participant will teach for one 3-minute turn
- Rotate participants until everyone has had a turn
- During this round prompting IS NOT allowed

Note: If for any reason a session must be shortened (e.g., extensive tech problems, power outage, etc.) do everything in your power to give all teachers an equal number of turns.

Verbal Feedback

Following each teacher turn:
- Ask the group: “What is one thing (participant name) did well on (target skill)?”
• Share the collected data on the target skill. You may also reiterate or clarify any feedback you heard from the participants.
• (If applicable, quickly share data of the skill being maintained, but focus feedback on the new target skill).
• Ask the teacher: “What is one thing you would like to improve on (target skill)?”
• **During the THIRD round** add a question in addition to or instead of the above question that requires the teacher to make a connection to their classroom, such as:
  o “What did you practice tonight on (target skill) that you will take to your students tomorrow?”
  o “What have you practiced or observed tonight that would be beneficial to your students?”
  o “What do you plan to implement with (target skill) in your own classroom that will benefit your students the most?”

Please Note:
• Be sure to isolate feedback to the target skill(s) for that session. Gently deflect comments or questions about future target skills.
• Keep feedback concise and focused and move on to the next teacher’s turn.
• You may answer teacher questions during this time, but again, it must be concise and focused on the target skill for that session.

**Written Feedback**
• Collect data on teacher performance for each turn.
• Keep all written feedback focused on the target skills for that session.
• Distribute the data forms at the end of the session
• (Make copies of the feedback forms from the session so we can keep a hard copy).

**Conclusion**
• During the last **3-5 minutes**:
  o Ask for any remaining questions about the target skill
  o Summarize the participants’ successes with the target skill, and if necessary point out what they can focus on improving next time
  o Conclude with any final remarks about the skill or session
  o Distribute the participation self-assessment forms
Appendix M

Feedback Forms
Feedback Form: OTR

**Target Skill: Opportunities to Respond (OTR)**

Goal: 4 OTRs/min, approximately 50% each question type

### Turn 1

<table>
<thead>
<tr>
<th>Basic Knowledge</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Application</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>% Basic Knowledge</th>
<th>Basic OTRs/total OTRs</th>
<th>Total OTR Rate (# of OTRs/mins)</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Application</td>
<td>Application OTRs/total OTRs</td>
<td></td>
</tr>
</tbody>
</table>

/ = correct and complete, 0 = attempted but incorrect, p = prompted

### Turn 2

<table>
<thead>
<tr>
<th>Basic Knowledge</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Application</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>% Basic Knowledge</th>
<th>Basic OTRs/total OTRs</th>
<th>Total OTR Rate (# of OTRs/mins)</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Application</td>
<td>Application OTRs/total OTRs</td>
<td></td>
</tr>
</tbody>
</table>

/ = correct and complete, 0 = attempted but incorrect, p = prompted

### Turn 3

<table>
<thead>
<tr>
<th>Basic Knowledge</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Application</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>% Basic Knowledge</th>
<th>Basic OTRs/total OTRs</th>
<th>Total OTR Rate (# of OTRs/mins)</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Application</td>
<td>Application OTRs/total OTRs</td>
<td></td>
</tr>
</tbody>
</table>

/ = correct and complete, 0 = attempted but incorrect
Feedback Form: Error Correction

**Target Skill: Error Correction**
Goal: Deliver all error correction steps in response to each academic error

### Turn 1

<table>
<thead>
<tr>
<th>Academic Error (Student and description)</th>
<th>Model</th>
<th>Test</th>
<th>Delayed Test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+ -</td>
<td>+ -</td>
<td>+ -</td>
</tr>
<tr>
<td></td>
<td>+ -</td>
<td>+ -</td>
<td>+ -</td>
</tr>
<tr>
<td></td>
<td>+ -</td>
<td>+ -</td>
<td>+ -</td>
</tr>
</tbody>
</table>

+ = correct and complete, - = incorrect or didn’t attempt, p = prompted

<table>
<thead>
<tr>
<th>Basic Knowledge OTRs:</th>
<th>Application OTRs:</th>
<th>Total OTR Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Turn 2

<table>
<thead>
<tr>
<th>Academic Error (Student and description)</th>
<th>Model</th>
<th>Test</th>
<th>Delayed Test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+ -</td>
<td>+ -</td>
<td>+ -</td>
</tr>
<tr>
<td></td>
<td>+ -</td>
<td>+ -</td>
<td>+ -</td>
</tr>
<tr>
<td></td>
<td>+ -</td>
<td>+ -</td>
<td>+ -</td>
</tr>
</tbody>
</table>

+ = correct and complete, - = incorrect or didn’t attempt, p = prompted

<table>
<thead>
<tr>
<th>Basic Knowledge OTRs:</th>
<th>Application OTRs:</th>
<th>Total OTR Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Turn 3

<table>
<thead>
<tr>
<th>Academic Error (Student and description)</th>
<th>Model</th>
<th>Test</th>
<th>Delayed Test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+ -</td>
<td>+ -</td>
<td>+ -</td>
</tr>
<tr>
<td></td>
<td>+ -</td>
<td>+ -</td>
<td>+ -</td>
</tr>
<tr>
<td></td>
<td>+ -</td>
<td>+ -</td>
<td>+ -</td>
</tr>
</tbody>
</table>

+ = correct and complete, - = incorrect or didn’t attempt

<table>
<thead>
<tr>
<th>Basic Knowledge OTRs:</th>
<th>Application OTRs:</th>
<th>Total OTR Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Feedback Form: Specific Praise

**Target Skill: Specific Praise**

Goal: 4 praise/min, with 2+ specific/min (or >50% specific)

### Turn 1

<table>
<thead>
<tr>
<th>Specific Praise</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>General Praise</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Academic Error</th>
<th>Model</th>
<th>Test</th>
<th>Delayed Test</th>
<th>% EC steps</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+ -</td>
<td>+ -</td>
<td>+ -</td>
<td></td>
</tr>
<tr>
<td></td>
<td>+ -</td>
<td>+ -</td>
<td>+ -</td>
<td></td>
</tr>
</tbody>
</table>

### Turn 2

<table>
<thead>
<tr>
<th>Specific Praise</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>General Praise</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Academic Error</th>
<th>Model</th>
<th>Test</th>
<th>Delayed Test</th>
<th>% EC steps</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+ -</td>
<td>+ -</td>
<td>+ -</td>
<td></td>
</tr>
<tr>
<td></td>
<td>+ -</td>
<td>+ -</td>
<td>+ -</td>
<td></td>
</tr>
</tbody>
</table>

### Turn 3

<table>
<thead>
<tr>
<th>Specific Praise</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>General Praise</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Academic Error</th>
<th>Model</th>
<th>Test</th>
<th>Delayed Test</th>
<th>% EC steps</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+ -</td>
<td>+ -</td>
<td>+ -</td>
<td></td>
</tr>
<tr>
<td></td>
<td>+ -</td>
<td>+ -</td>
<td>+ -</td>
<td></td>
</tr>
</tbody>
</table>

*i = correct and complete, 0 = attempted but incorrect, p = prompted*
Feedback Form: Praise Around

**Target Skill: Praise Around**
Goal: Deliver all praise around steps in response to each disruptive student behavior

**Turn 1**

<table>
<thead>
<tr>
<th>Behavior (Student and description)</th>
<th>PA Step 1</th>
<th>PA Step 2</th>
<th>% PA steps (# steps/# possible)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+ -</td>
<td>+ -</td>
<td></td>
</tr>
<tr>
<td></td>
<td>+ -</td>
<td>+ -</td>
<td></td>
</tr>
<tr>
<td></td>
<td>+ -</td>
<td>+ -</td>
<td></td>
</tr>
</tbody>
</table>

+ = correct and complete, - = incorrect or didn’t attempt, p = prompted

<table>
<thead>
<tr>
<th>Specific Praise:</th>
<th>General Praise:</th>
<th>Total Rate</th>
<th>Specific Rate</th>
<th>% Specific</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Academic Error</th>
<th>Model</th>
<th>Test</th>
<th>Delayed Test</th>
<th>% EC steps</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+ -</td>
<td>+ -</td>
<td>+ -</td>
<td></td>
</tr>
<tr>
<td></td>
<td>+ -</td>
<td>+ -</td>
<td>+ -</td>
<td></td>
</tr>
</tbody>
</table>

**Turn 2**

<table>
<thead>
<tr>
<th>Behavior (Student and description)</th>
<th>PA Step 1</th>
<th>PA Step 2</th>
<th>% PA steps (# steps/# possible)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+ -</td>
<td>+ -</td>
<td></td>
</tr>
<tr>
<td></td>
<td>+ -</td>
<td>+ -</td>
<td></td>
</tr>
<tr>
<td></td>
<td>+ -</td>
<td>+ -</td>
<td></td>
</tr>
</tbody>
</table>

+ = correct and complete, - = incorrect or didn’t attempt, p = prompted

<table>
<thead>
<tr>
<th>Specific Praise:</th>
<th>General Praise:</th>
<th>Total Rate</th>
<th>Specific Rate</th>
<th>% Specific</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Academic Error</th>
<th>Model</th>
<th>Test</th>
<th>Delayed Test</th>
<th>% EC steps</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+ -</td>
<td>+ -</td>
<td>+ -</td>
<td></td>
</tr>
<tr>
<td></td>
<td>+ -</td>
<td>+ -</td>
<td>+ -</td>
<td></td>
</tr>
</tbody>
</table>
## Turn 3

<table>
<thead>
<tr>
<th>Behavior (Student and description)</th>
<th>PA Step 1</th>
<th>PA Step 2</th>
<th>% PA steps (# steps/# possible)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+ -</td>
<td>+ -</td>
<td></td>
</tr>
<tr>
<td></td>
<td>+ -</td>
<td>+ -</td>
<td></td>
</tr>
<tr>
<td></td>
<td>+ -</td>
<td>+ -</td>
<td></td>
</tr>
<tr>
<td></td>
<td>+ -</td>
<td>+ -</td>
<td></td>
</tr>
</tbody>
</table>

+ = correct and complete, - = incorrect or didn’t attempt, p = prompted

<table>
<thead>
<tr>
<th>Specific Praise:</th>
<th>General Praise:</th>
<th>Total Rate</th>
<th>Specific Rate</th>
<th>% Specific</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Academic Error</th>
<th>Model</th>
<th>Test</th>
<th>Delayed Test</th>
<th>% EC steps</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+ -</td>
<td>+ -</td>
<td>+ -</td>
<td></td>
</tr>
<tr>
<td></td>
<td>+ -</td>
<td>+ -</td>
<td>+ -</td>
<td></td>
</tr>
<tr>
<td></td>
<td>+ -</td>
<td>+ -</td>
<td>+ -</td>
<td></td>
</tr>
</tbody>
</table>
Appendix N

Participation Self-Assessment Form
Participation Self-Assessment Form

<table>
<thead>
<tr>
<th>Name: __________________________</th>
<th>Date: __________________________</th>
</tr>
</thead>
</table>

**TeachLivE Participation Points**

*Directions:* Circle the point value you feel reflects your performance in the following areas:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. I arrived on time and stayed the entire session</td>
<td></td>
<td>0</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. I brought all necessary materials</td>
<td></td>
<td>0</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. I participated in all teaching turns with my best effort</td>
<td></td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>4. I actively observed and provided feedback for my peers</td>
<td></td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>5. I demonstrated a professional attitude</td>
<td></td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**TOTAL POINTS**

_____ /15
Appendix O

Treatment Fidelity Checklist
Treatement Fidelity Checklist

Observer: ___________________ Session Date: ___________ Target Skill: ___________________

<table>
<thead>
<tr>
<th>Intervention Components</th>
<th>No (0)</th>
<th>Yes (1)</th>
<th>Total</th>
<th>Points Possible</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Instructor introduces the session by reminding the teachers of the <strong>target skill</strong>, the <strong>definition</strong> of the target skill, and the <strong>goal</strong> of the target skill.</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Instructor provides a <strong>handout</strong> (or verifies that participants brought their own) with the definition, examples, and non-examples of target skill</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Each participant is given three turns to teach (If the turns vary from 2 min, 3 min, 3 min, make a note in the comments)</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. <strong>Interactor Fidelity</strong>: The interactor delivers the session as prescribed, depending on the target skill: <strong>OTR</strong>: NO errors and NO misbehavior <strong>EC</strong>: Only errors. No misbehavior <strong>PRAISE</strong>: Errors (EC maint.). No misbehavior <strong>PA</strong>: Mostly misbehavior. Occasional errors (EC maint.)</td>
<td>Round 1</td>
<td>Round 2</td>
<td>Round 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1st</td>
<td>2nd</td>
<td>3rd</td>
<td>4th</td>
<td>1st</td>
<td>2nd</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>5. <strong>Soliciting Group Feedback</strong>: After each turn the instructor asks the observers to verbalize what the teacher did well on the target skill.</td>
<td>Round 1</td>
<td>Round 2</td>
<td>Round 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1st</td>
<td>2nd</td>
<td>3rd</td>
<td>4th</td>
<td>1st</td>
<td>2nd</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6. <strong>Mentor Feedback</strong>: The instructor shares the teacher’s data on the target skill, and reiterates or clarifies any feedback the other teachers provided.</td>
<td>Round 1</td>
<td>Round 2</td>
<td>Round 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1st</td>
<td>2nd</td>
<td>3rd</td>
<td>4th</td>
<td>1st</td>
<td>2nd</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7. <strong>Soliciting Participant Feedback</strong>: The instructor asks the teacher to reflect on their performance, especially areas in which they can still improve on the target skill. (After the 3rd turn the mentor asks for a <strong>classroom connection</strong>).</td>
<td>Round 1</td>
<td>Round 2</td>
<td>Round 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1st</td>
<td>2nd</td>
<td>3rd</td>
<td>4th</td>
<td>1st</td>
<td>2nd</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>8. The instructor may prompt the teacher during</td>
<td>Round 1</td>
<td>Round 2</td>
<td>Round 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>
the first two turns, but does not provide any prompting during the 3rd turn.

<table>
<thead>
<tr>
<th></th>
<th>1st</th>
<th>2nd</th>
<th>3rd</th>
<th>4th</th>
<th>1st</th>
<th>2nd</th>
<th>3rd</th>
<th>4th</th>
<th>1st</th>
<th>2nd</th>
<th>3rd</th>
<th>4th</th>
</tr>
</thead>
<tbody>
<tr>
<td>Round 1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Round 2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Round 3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

9. The instructor provides feedback on the target skill(s) for that session only. Deduct points if skills for FUTURE sessions are addressed. (Order of skills: OTRs, EC, Praise, PA)

10. The instructor provides a data sheet for each participant with feedback from each teacher turn. (If not possible to view from the video, view as a permanent product).

11. The data sheet includes feedback on the target skills for that session only, and not for target skills addressed later in the study (viewed as a permanent product).

**Fidelity of Implementation Percentage for Session** (Total points attained/total points possible):
CURRICULUM VITAE

MELANIE REES DAWSON
melanie.dawson@live.com

EDUCATION

Ph.D. (2016) Disability Disciplines
Specialization: Special Education
Utah State University, Logan, Utah
Advisor: Dr. Benjamin Lignugaris/Kraft

M.Ed. (2008) Elementary Education
Utah State University, Logan, Utah
Advisor: Dr. Sylvia Read

B.S. (2002) Special Education & Elementary Education
Emphasis: Mild/Moderate Special Education
Utah State University, Logan, Utah
Honors: Magna Cum Laude Graduate

Emphasis: Education
Utah State University, Logan, Utah
Honors: Summa Cum Laude Graduate

CERTIFICATIONS/ENDORSEMENTS

Administrative Supervision
Special Education, mild/moderate disabilities, K-12
Elementary Education, 1-8
Theater Arts, 6-12
Teaching English as a Second Language
Reading Endorsements, Level I and Level II

RESEARCH AND EDUCATIONAL INTERESTS

Evidence-based instruction and behavior strategies for students with mild/moderate disabilities
Effective use of virtual classrooms, including TeachLivE, in teacher preparation
Implementation of Multi-tiered Systems of Support (MTSS)
Administrative leadership
Applied Behavior Analysis
Single subject research design
RESEARCH EXPERIENCE:

Current  **Dissertation:** *From TeachLivE to the Classroom: Building Preservice Special Educators’ Proficiency with Essential Teaching Skills*
Providing educators with practice opportunities on essential teaching skills in TeachLivE, and investigating the extent to which they transfer to the classroom

2014-Current  **Co-Investigator:** *Critical Incidents in the Statewide Implementation of Multi-Tiered Systems of Support (MTSS)*
Interviewing project directors of state scale up efforts to identify the critical incidences that facilitate or hinder statewide implementation of MTSS

2014  **Research Assistant:** *Questions in the Classroom: Training Teachers to use a Strategic Approach for Asking Reading Comprehension Questions*
Double-coded teachers’ delivery of various question types, and provided technical assistance for all practice sessions conducted in TeachLivE

2013  **Principal Investigator:** *Meaningful Practice: Generalizing Essential Teaching Skills from TeachLivE to the Special Education Classroom*
Evaluated teachers’ proficiency with target skills in the virtual environment and piloted data collection in their classroom settings

2012  **Principal Investigator:** *TeachLivE vs. Roleplay: Comparative Effects on Preservice Special Educators’ Acquisition of Basic Teaching Skills*
Examined teachers’ proficiency with counterbalanced target skills practiced in TeachLivE or roleplay to determine the comparative effectiveness of each setting

UNIVERSITY TEACHING

Fall 2011-Current  **Main Instructor/Lab Mentor**
*SPED 5790: TeachLivE Lab Practicum (Part I)*
Utah State University, Salt Lake Center
Target skills: Opportunities to respond, error correction, praise, praise around

Spring 2012-Current  **Main Instructor/Lab Mentor**
*SPED 5790: TeachLivE Lab Practicum (Part II)*
Utah State University, Salt Lake Center
Target Skills: Precision requests, consequences, instructional scaffolds

Spring 2014  **Main Instructor**
*SPED 5360: Teaching Students with Mild/Moderate Disabilities (Part II)*
Utah State University, Salt Lake Center
Professor: Marilyn Likins
Topics: Positive and reductive behavioral strategies, self-monitoring, single subject research designs, prompting/fading, generalization/maintenance, working effectively with paraeducators, teaming with parents

Spring 2013  **Teaching Assistant, Guest Instructor**
*SPED 5360: Teaching Students with Mild/Moderate Disabilities (Part II)*
Utah State University, Salt Lake Center
Professor: Marilyn Likins

Fall 2012-Current  **Guest Instructor** on classroom data collection systems
*SPED 5350: Teaching Students with Mild/Moderate Disabilities (Part I)*
Utah State University, Salt Lake Center
Professor: Marilyn Likins

**ADMINISTRATIVE INTERNSHIPS**

2011-Current  ATP Program, USU
Mentors: Marilyn Likins, Randy Schelble, Ben Lignugaris/Kraft
Teacher training and supervision, didactic instruction, research, grant writing

Jan-June 2014  Lincoln Elementary, Salt Lake City School District
Mentors: Peggy Paterson and Samantha Salazar
Grade level PLCs, behavior management, instructional supervision, family nights

May-Dec 2013  Kearns High School, Granite School District
Mentors: Maile Loo and Heather Sonne
Teacher and student supervision, discipline, extracurricular activities

**TEACHER SUPERVISION**

Current  Student Teaching Supervisor (elementary), USU Distance Program
Conducting regular visits and providing feedback for small-group reading and math instruction

2013-2014  Administrative Supervision, Kearns High School and Lincoln Elementary
Observed and provided feedback to teachers across departments in instruction and behavior strategies
2012-2013   Placement Supervisor (secondary), USU ATP Program
Conducted regular formal and informal observations and provided feedback in whole-class reading and math instruction

PROFESSIONAL TEACHING EXPERIENCE

Elementary Education
2010-2011   5th Grade General Education Teacher
Stansbury Elementary, West Valley City, Utah
Tier I and Tier II instruction in reading, language arts, and math

2006-2010   Reading Specialist, Grades K-4
Stansbury Elementary, West Valley City, Utah
Tier II small group reading interventions

Special Education
2004-2006   Self-Contained Reading and Math Inclusion Teacher
Jackson Creek Middle School, Bloomington, Indiana
Intensive reading remediation and math inclusion accommodations

2004-2005   Self-Contained Reading, Math, and Science Teacher
Bachelor Middle School, Bloomington Indiana
Focused remediation in core subjects, aligned with general curriculum

2003-2004   Resource Language Arts Teacher
Kearns Junior High, Kearns, Utah
Reading and language arts remediation, behavior systems

Theater Education
2013-2014   Summer Theater Camp Instructor, Ages 6-13
Children’s Theater of Salt Lake, Salt Lake City, Utah
Creative dramatics, script creation, and performance

2005-2006   Creative Dramatics Teacher
Jackson Creek Middle School, Bloomington, Indiana
Theatre history, acting, playwriting, technical theatre, and drama club

2003-2004   Program Coordinator and Instructor, preK-12
Children’s Theater of Salt Lake, Salt Lake City, Utah
Musical theater, improvisation, and audition training
RELATED EMPLOYMENT

2010-present TeachLivE Lab Coordinator and Mentor
Utah State University, Salt Lake Center
Implementation of the TeachLivE Simulated classroom

2007-2012 Special Education Recruitment Coordinator for the Wasatch Front Area,
Utah State University, Logan, Utah
SPED 1000 concurrent enrollment supervisor and paraeducator
recruitment

2001-2004, Drama Specialist, Opera By Children Program
2006-2011 Utah Festival Opera, Logan, Utah
Drama consultant/trainer for original student operas

PUBLICATIONS

Manuscripts In Press:

essential teaching skills from TeachLivE to the special education classroom.
Teacher Education Special Education.

Manuscripts in Preparation:

Dawson, M. R., Hayes, A., & Lignugaris/Kraft, B. (in progress). TeachLivE vs. roleplay:
Comparative effects on special educators’ acquisition of basic teaching skills.

Dawson, M. R., & Lignugaris/Kraft, B. (in progress). Specific praise and student
outcomes: A research synthesis.

Conference Proceedings:

effects on special educators acquisition of basic teaching skills. In A. Hayes, S.
Hardin, L. Dieker, C. Hughes, M. Hynes, & C. Straub. Conference Proceedings
for First National TeachLivE Conference. Paper presented at First National
TeachLivE Conference: Orlando, FL, University of Central Florida.

Non-Referred Articles:

Morgan, R. L., Rees, M., & Lyman, J. (December, 2009). Recruiting the next generation
of special education teachers. The Utah Special Educator, 32 (2), 44-46.
Morgan, R. L., & Rees, M. (2008). From tutor to teacher: For some peer tutors, the experience can lead to a teaching career. The Utah Special Educator, 28 (4), 42-43.

PRESENTATIONS


GRANT EXPERIENCE

2014 Special Education Personnel Preparation Grant from the Utah State Office of Education. Funded $171,234 ($85,617 per year for two years).

Role: Grant Writing Assistant. My main responsibilities included gathering and reporting ATP teacher data, and assisting in writing and editing the narrative.

2016 Special Education Personnel Preparation Grant from the Utah State Office of Education. Funded $224,812 (Year 1: $122,402, Year 2: $102,410)
Role: Principal Investigator. My responsibilities included writing and editing the grant narrative, reporting data from past ATP projects, and assisting Dr. Marilyn Likins in creating the budget.

PROFESSIONAL SERVICE AND MEMBERSHIPS

Junior Achievement volunteer
Member, Council for Exception Children
Member, International Reading Association
Reviewer for Council for Exceptional Children Student Award and Scholarships
Guest Reviewer for Education and Treatment of Children (2012; 2014)
Critical Friends Group, professional learning community facilitator, Bloomington, IN
Monroe County Youth Shelter workshop leader, Bloomington, Indiana
Utah Theater Association, workshop presenter
Publicity and event volunteer, Special Olympics, Utah State University
College of Education, council member, Utah State University

AWARDS AND HONORS

Presidential Fellowship, School of Graduate Studies (2011-2012)
University Club Scholar, 4-year undergraduate scholarship (1998-2002)
Glenda Thornley Theater Scholarship (2002)
Education Achievement Award, 1st place, Golden Key International Honour Society, (2003)
Performing Arts Showcase Scholarship, 1st place, Golden Key International Honour Society, (2002)
Theater Arts Sterling Scholar, state runner up (1998)