1-1-2019

Measures of music teacher effectiveness in three early career levels

Alicia L. Canterbury

MEASURES OF MUSIC TEACHER EFFECTIVENESS IN THREE EARLY CAREER LEVELS

A Dissertation
presented in partial fulfillment
for the degree of Doctor of Philosophy
in the Department of Music
at the University of Mississippi

by

ALICIA L. CANTERBURY

August 2019

ABSTRACT

The purpose of this study was to observe the teaching of five novice music teachers in three early career levels to determine if effective teaching behavior increases over time. A secondary purpose was to find if any specific teacher behaviors associated with effectiveness were more prominent when teaching their peers in pre-service, another educator's students in student teaching, and their own students in novice teaching. Eight video recordings were transcribed for the purpose of identifying rehearsal frames and their instructional targets. Certain rehearsal frames were selected for further analysis. Additionally, five expert teachers were selected to review and evaluate the participants' videos using a summative evaluation form based on six items related to teacher effectiveness.

Corroborative findings across instructional targets, rehearsal frames, and the summative evaluations indicated a general lack of specificity across all levels which improved somewhat during student teaching and peer teaching. Unidentified targets and nonspecific positive feedback were observed less frequently in student teaching and novice teaching. No growth was found across levels in specific positive and negative feedback.

Directives were the most frequently observed teacher verbal category in selected rehearsal frames and information and demonstrations were consistently the highest ranked items across all levels in the summative evaluation. However, instructional directive scores were

inconsistent. Results indicate that new teachers of all levels may have appropriate pedagogical knowledge, but have difficulty explaining it to students.

Teacher modeling doubled during novice teaching. Positive modeling was the second most observed category of teacher behaviors in peer teaching and third in novice teaching. Concurrent performance model was the second most observed category during student teaching and novice teaching. Many of the evaluators commented on the overuse of rote teaching and performing with students and lack of competency-based education techniques during student teaching and novice teaching.

Three participants had low rates of rehearsal frames with multiple performance trials (one every 50 minutes). These participants' trends on their summative evaluations indicated decreased scores in novice teaching or flatlining across levels. Specifically, these participants had low ratings in feedback, flow, and musical model across all levels. Implications for higher education training are discussed.

DEDICATION

This work is dedicated to the strong women in my life past and present who have modeled tenacity, grace under pressure, determination, kindness, and humility. Your influence is unmeasurable.

ACKNOWLEDGMENTS

I would like to express sincere gratitude to my dissertation advisor, Dr. Michael Worthy, for his guidance, feedback, humor, and excitement for teaching during my time at the University of Mississippi. His dedication and passion for music education has been a great inspiration to me. I would also like to thank Dr. Alan Spurgeon, Dr. Andrew Paney, and Mr. Greg Johnson who have congenially shared their knowledge during the dissertation process. I would also like to acknowledge the novice teachers who participated in the study as well as the expert teachers who assisted me in evaluating teacher behaviors. A final word of thanks is extended to my parents, Mr. and Mrs. Jim Canterbury, whose constant support and belief in me has meant everything. I also give thanks to my God who has been faithful to me in every season.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Promoting growth in early career teachers has been a goal in teacher training since its inception. Experiences and courses taken in pre-service and in student teaching ideally place novice teachers on a trajectory of self-reflection and growth in effective practices. However, research in teacher effectiveness has raised as many follow-up questions for each answer it has provided. Researchers who have explored this multi-faceted topic create inconclusive and challenging questions for future researchers to consider. Further, many researchers believe that a richly descriptive summary of effective teaching is elusive and that most studies focus on classroom management rather than pedagogical knowledge of the student (Berliner, 1986; Madsen, Standley, Byo, & Cassidy, 1992; Duke, 1999/2000; Shulman, 1987; Siedentop & Eldar, 1989).

As with most any other profession, expert teachers should serve as models for those entering the field. In this regard, Berliner (1986) and Rosenshine (2008) indicate that effective master teachers are consistent in systematic instruction whereas less effective or novice teachers may be purposeful, but not always in a sequential way. Rosenshine (2008) denoted that one of the main differences between expert and novice teachers is that expert teachers have a larger frame of reference to work from and can make transfers between knowledge sets. Novice teachers may have a similar set of skills and pedagogical knowledge, but the delineation between

experts and novices is their ability to use those skills consistently (Berliner, 1986; Shulman, 1987; Siedentop & Eldar, 1989). Additionally, Siedentop and Eldar (1989) "do not believe that expertise in teaching exists and that effectiveness and experience are necessary but not sufficient conditions for demonstrating it" (p. 257). This suggests that many novice teachers can be effective and may develop at a quicker pace than expected, perhaps due to prior training in pre-service conditions which encourages automaticity in teaching skills (Berliner, 1986; Siedentop & Eldar, 1989; Bloom, 1986; Rosenshine, 2008). Automaticity is encouraged in processing what Rosenshine described as the role of seven, in that only seven new items can be processed at one time before the teacher becomes overloaded. Retention of new material is also assisted when an individual is encouraged to explain a new concept to a peer or colleague.

And so, it is a small wonder that pre-service teachers can be given all the necessary tools for them to be immediately effective in their classroom in the first years of teaching. This may be due to the short amount of time spent in preservice teaching (Shulman, 1987) in comparison to the hopefully numerous years in the field. In this brief time period, Shulman (1987) argued the purpose of teacher training is not to indoctrinate, but rather to provide pre-service teachers with the tools for critical thinking and develop a solid understanding of their teaching philosophy. Helping pre-service teachers to understand that teaching is not solely a means to an end (band contests, marching shows, concerts, etc.), but is "concerned with both means and end" is important.

Opinions about the process of novice teachers' growth in expertise are mixed (Murnane & Phillips, 1981; Rosenshine, 2008; Shulman, 1987). First impressions would indicate that experience solely provides the novice with the greatest gains, in that expertise is found in "error, success, and refinement" (Shulman, 1987, p.4). Levin, Hammer, and Coffey (2009) believe that

novice teachers can be effective if they have the predisposed ability to frame learning in real time. Siedentop & Eldar (1989) observed seven elementary physical education teachers of varying degrees of experience and discovered that first-year teachers and expert teachers were both effective in that they provided similar instruction and their students were equally successful. This, in turn, gave the researchers hope that novice teachers could indeed become effective at an early period in their educational careers.

Berliner (1986) believed that the highest rates of growth occur in the first five years of teaching. Many researchers seem to agree with Berliner's theory. Rivkin, Hanusheck, and Kain (2005) stated in their report regarding effectiveness of elementary classroom teachers in Texas that "experience is not significantly related to achievement following the initial years in the profession" (p. 419). Their data tended to suggest that novice teachers have the highest gains in effective measures in the first years of teaching and begin to flatline in their effectiveness level later in their careers. This would seem to indicate that experience does not always equate with a high level of skill in teaching. Chingos and Peterson (2010) found similar results in their analysis of late elementary school teachers in Florida and their effectiveness on math and reading scores. Similarly, Marsh (2007) found in his review of university professors' student evaluation forms tracking effectiveness over 13 years that professors showed very similar effectiveness levels mid-career as they did as early-career professors.

Since growth in effectiveness seems to occur in the first five years of teaching and then begins to flatline (Berliner, 1986; Chingos & Peterson, 2010; Marsh, 2007; Rivkin, et al., 2005), the amount of growth within those years may be due in part to rigorous pre-service training and student teaching. Therefore, attempting to track early career levels of effectiveness might shed more light on understanding this process. An influential theory of expertise development was

conceived by Dreyfus and Dreyfus (1980, 1986) which included a five-stage process of acquiring expertise. These stages include novice, advanced beginner, competent, proficient, and expert. Berliner (1988) takes Dreyfus and Dreyfus' theory and transfers it to teaching expertise. The novice teacher uses rule-based methods and is unable to transfer knowledge in context. The advanced beginner understands contexts more but is unable to transfer key content to students. The competent teacher is able to transfer knowledge to new situations and is able to create objectives appropriate for the situation at hand but lacks intuitiveness and may be sluggish to respond. An individual considered to be proficient has increased in discernment and can see the larger picture of reasons why an event with students has been successful or not. Proficient teachers can predict student behavior, competency levels, and pitfalls in objectives given to students. At the expert stage, some behaviors are ingrained to the point of automaticity. To an outside observer, an expert may so seamlessly engage with students that management may go unnoticed.

Most studies define an expert as one with more than five years of teaching expertise, recommendation by other experts in the field, or professional achievement evidenced through performance accolades (Anderson-Nichols, 1997; Berliner, 1986; Duke & Simmons, 2006; Goolsby, 1996, 1997, 1999). According to studies that attempt to define expert teacher behaviors, expert teachers can organize and plan instruction with specific targets in mind and are relentless in pursuit of their goals for their students and have high expectations. However, they are flexible and are able to change and/or deconstruct the trajectory of the original lesson target if a student is having difficulty with a given concept. Master teachers review work from previous lessons and consider prior knowledge before introducing new concepts. They also model appropriate playing or singing techniques. Experts can motivate students to achieve specific

learning targets with enthusiasm and energy. They have excellent interpersonal skills which furthers good rapport with students using proactive strategies. They are good communicators and provide a variety of feedback styles including specific and technical feedback. Additionally, students are successful because effective teachers know what their students can realistically accomplish due to their in-depth knowledge of pedagogy and the skill level of the student. They can accomplish a great deal in a given time period due to their focus on quick pacing which deters off-task student behavior. Master teachers are skilled at classroom management and can anticipate student outcomes.

Novice teacher definitions vary widely across studies and can include pre-service teachers, student teachers, and in-service teachers with less than five years of experience. However, studies that have compared micro-experience levels tend to have similar findings between the three levels, with a slight decrease in effective behaviors during in-service novice teaching. Novice teachers spend more time talking than do expert teachers. Novices tend to focus on less nuanced issues such as music fundamentals, tuning individual instruments, and fixing wrong pitches. Time management can be challenging—novices tend to inadvertently allow student off-task behavior due to poor pacing or because of time spent on one task. They tend to not provide feedback for performance trials and ask their students to repeat a task with no stated purpose which can result in frustration. Novices have difficulty knowing how to adjust their lessons when faced with impromptu student questions, responses, or poor performances that fail to meet instructional targets. They also tend to lack understanding of specific student needs. However, novices may have deep pedagogical knowledge and can provide for student success if the proper tools of scaffolding are given in teacher training (Siedentop & Eldar, 1989).

Authentic Context Learning

Novice teachers may be placed on a higher trajectory toward success if they become knowledgeable of master teacher characteristics during teacher training programs. Several studies have been performed to investigate the effectiveness of authentic context learning (ACL) settings for long-term skill development. Students seem to increase in effective teaching behaviors when provided with ACL experiences such as peer teaching and field teaching (Blackwell & Roseth, 2018; Bowers, 1990; Butler, 2001; Byo, 1990; Cassidy, 1990; Collins, 1978; Colwell, 1995; Fant, 1996; Goodrich, Bacura & Stauffer, 2017; Haston & Russell, 2012; Killian, Dye, & Buckner, 2008; Nápoles, 2016; Paul, 1998; Paul et al., 2001; Powell, 2014; Powell, Weaver, & Henson, 2018; Worthy, 2005). Findings indicated overall positive outcomes in ACL experiences. Peer teaching provided technical assistance in lesson planning and methodology, but lacked authenticity (Powell, 2014; Scmidt, 2010). Preservice field teaching provide authenticity, but preservice teachers had difficulty responding intuitively to student performance (Powell, 2014). Some studies indicated that participants were comfortable teaching their peers while others indicated the opposite. The type of ACL experience was important in that ACL experiences were positive if they included specific feedback from instructors or experts. Additionally, giving preservice teachers structured predetermined goals in peer teaching yielded more positive results. ACL experiences also provided more awareness of teaching behaviors which provided further scaffolding for success when participants moved on to teaching in the field (Fant, 1996; Paul, 1998). Further, efficacy levels increased when students were tracked longitudinally within a series of micro-teaching and peer teaching segments (Bowers, 1990; Collins, 1978; Cassidy, 1990). Researchers who have tracked student progress within a

series of peer teachings have found that students usually grow in effective behaviors over the given time span (Blackwell & Roseth, 2018; Butler, 2001; Colwell, 1995; Killian et al., 2008; Worthy, 2005).

Purpose of the Study and Research Questions

The purpose of this study is to observe the teaching of five novice music teachers in three early career teaching levels (peer teaching, student teaching, and novice teaching) to determine if effective teaching behavior increases over time. The study's secondary purpose is to find if any specific teacher behaviors associated with effectiveness are more prominent when teaching their peers in pre-service, another educator's students in student teaching, and their own students in novice teaching.

Video excerpts were analyzed using two methods. An expert teacher observed one participant's videos in sequential order of career levels and completed the Teacher Effectiveness Evaluation Form (TEEF) after watching each video. The summative evaluation form asked observers to rate flow (pacing), feedback, modeling, verbal instruction, and teaching style. The videos were analyzed to identify rehearsal frames and their instructional targets, and to measure the frequency and duration of specific teacher and student behaviors within selected rehearsal frames.

Most researchers have observed heterogenous groups of participants to identify novice and expert behaviors. However, longitudinal studies utilizing video recordings from three points in time over three early career levels are unknown. While micro-growth has been studied in pre-service teaching and student teaching, micro-growth in novice teaching has not been studied.

Further, the use of rehearsal frames and summative evaluation forms as tools for analysis are vetted as useful observational, quantitative, and descriptive tools. Therefore, the research questions to be answered in this study include:

1. What are the frequencies of instructional targets in rehearsal frames of the three early career levels?
2. What are the frequency and duration of specified teacher and student behaviors observed in selected rehearsal frames?
3. What are the ratings of teaching effectiveness for each participant at each level of teaching?

Limitations

The study focuses on the specific results for a multiple case study with a small number of participants (N = 5). Divergent results could be found with different participants, a larger sample size, or those with different training experiences in the three levels of teaching depending on where they received teacher training. Similarly, results could be affected by the school in which the participant did their student teaching and where they are currently teaching. Observation of these participants at different times could provide alternative outcomes to those in the current study. Participants were responsible for video recording their own student teaching and a wide range of video qualities may have unwittingly altered the researcher's perception of teaching behaviors and student responses. The use of a video camera to observe behavior, regardless of the setting, may unintentionally alter behaviors, as may the presence of an outside observer. All video recordings focused on the participant and not on the students. Video recording both

students and the teacher may have provided a broader spectrum of behaviors and different results.

Definition of Terms

Pre-service teacher—A teacher who is currently attending a college/university in preparation for in-service teaching. Teaching experiences offered through the college/university may include peer teaching and field teaching.

Student teacher—A teacher who is currently attending a college/university in preparation for in-service teaching and is currently teaching in the field under a supervising teacher. Their capacity and involvement may be limited, depending on the situation.

Novice teacher—An in-service teacher who has graduated from a college/university and is state-certified to teach music. A novice teacher has less than five years of teaching experience.

Rehearsal Frames—A sequence in instruction that begins when a teacher explicitly or implicitly identifies an instructional target that is followed by isolated or decontextualized performances of the musical segment and may include verbal instruction, modeling, approximations and repetitions. A rehearsal frame concludes with a recontextualized performance of the musical segment.

Instructional Targets—An objective which guides learning.

CHAPTER II

REVIEW OF LITERATURE

Teacher effectiveness is a worldwide concern in music and other subject areas. Efficacy has been studied in locations such as China, Japan, Malaysia, Iran, and Italy (Biasutti & Concina, 2017; Gau & Liu, 2013; Hamid, Hassan, & Ismail, 2012; Liu & Meng, 2009; Mehrpour & Mirsanjari, 2016; Price, Ogawa, & Arizumi, 1997). Moreover, a broad array of disciplines have been interested in efficacy as well. These range from a United States Air Force's squadron officer college, university psychology and physics departments, English teachers in Iran, high school science, and school districts (Chingos & Peterson, 2011; Keller, 2011; McCall, 2008; Mehrpour & Mirsanjari, 2016; Murray, Rushton, & Paunonen, 1990; O'Meara, 2007). In the field of music education, there have been numerous studies covering a vast range of expertise from pre-service to veteran teachers concerning elementary to adult education. Most researchers are concerned about a variety of teaching behaviors related to effectiveness. However, there are some who are concerned with specific behaviors such as eye contact (Browning, 2007; Frederickson, 1992), nonverbal communication (Heath-Reynolds, 2014), and a large interest area of social skills, social intelligence, and personality in relation to expert teaching (Biasutti & Concina, 2017; Gau & Liu, 2013; Hamann, Lineburgh, & Paul 1998; Juchniewicz, 2008; Murray et al., 1990; Running, 2011; Westbrook, 2004).

The purpose of this literature review is to explore the elusive profile of the effective "master teacher" and to explore how novice and pre-service teachers compare in those qualities. The review will also investigate how effectiveness has been measured and the effectiveness of actual context learning in pre-service teaching as well as in student teaching. Longitudinal studies which have explored individualized pathways to effectiveness and implications found within the body of literature will be surveyed as well.

Experience and Expertise

Many researchers in the past believed that a richly descriptive summary of effective teaching is perhaps nonexistent in that most studies focus on classroom management versus pedagogical knowledge of the student (Berliner, 1986; Madsen, Standley, Byo, & Cassidy, 1992; Duke, 1999/2000; Shulman, 1987; Siedentop & Eldar, 1989). An influential theory of expertise development was conceived by Dreyfus and Dreyfus (1980, 1986) in which individuals may go through a five-stage process of expertise. These include novice, advanced beginner, competent, proficient, and expert. Berliner (1988) takes Dreyfus and Dreyfus' theory and transfers it to teaching expertise. The novice teacher uses rule-based methods and is unable to transfer knowledge within context. The advanced beginner understands contexts more but is unable to transfer key content to students. The competent teacher can transfer knowledge to new situations and is able to create objectives appropriate for the situation at hand but lacks intuitiveness and may be sluggish to respond. An individual considered to be proficient has increased in discernment and can see the larger picture of reasons why an event with students has been successful or not. Proficient teachers can predict student behavior, competency levels, and

identify potential pitfalls in objectives given to students. At the expert stage, some behaviors are ingrained to the point of automaticity. To an outside observer, an expert may so seamlessly engage with students that management may remain unnoticed. According to Berliner, the most significant growth period is the first five years of teaching. Therefore, it could ideally take five years for a teacher to shift from a novice toward expertise in this period. However, it is also possible for teachers to never reach the expert level and to even regress during their first five years (Anderson-Nichols, 1997; Berliner, 1988). There have been many studies performed which seem to indicate truth in Berliner's theory of growth, as will be discussed later. Table 1 indicates Anderson-Nichols' summary of Berliner's continuum (1997).

Opinions about how novice teachers grow in expertise are mixed (Murnane & Phillips, 1981; Rosenshine, 2008; Shulman, 1987). First impressions would indicate that experience solely provides the novice with more opportunities for expertise. Experience is found in "error, success, and refinement" (Shulman, 1987, p.4). Berliner (1986) indicates that it may take approximately five years for novice teachers to provide evidence of more effective behaviors. However, classroom experience may not be the sole means of efficacy in that some success may be attributed to alternate experience outside of the field of traditional education (Murnane & Phillips, 1981; Siedentop & Eldar, 1989). There are many researchers and authorities who question if expertise and effectiveness should be equated with experience (Chingos & Peterson, 2011; Kini & Poldosky, 2016; Marsh, 2007; Rivkin, et al., 2005; Siedentop & Eldar, 1989). Levin, Hammer, and Coffey (2009) believe that novice teachers can be effective but must have the predisposed ability to frame learning in real time. Siedentop & Eldar (1989) observed seven elementary physical education teachers of varying degrees of experience and

Table 1

Anderson-Nichols' Summary of Berliner's Continuum of Teacher Expertise

Stage	Novice	Advanced Beginner	Competent	Proficient	Expert
Years of Experience	Pre-Service to 2	1 to 4	4 to 5	5 to ?	5 to ?
Behaviors	Deliberate	Insightful	Rational	Intuitive	Arational
Environment	Context-free	Context-guided	Chooses	Situational or Context-free	Situational or Context-free
Cognitive Process	Rational	Rational	Rational	Analytical or Intuitive	Arational or Intuitive
Actions	Pawn	Pawn	Personal agency	Personal agency	Personal agency
Knowledge	Propositional	Transitional	Procedural episodic case	Procedural episodic case	Procedural episodic case

Note. In Anderson-Nichols' notes, teacher expertise, though influenced by experience, cannot be solely determined by the number of years of teaching. Many other factors, as illustrated by this chart, contribute to expertise development. The span of years is intended to serve as a point of reference, not an established fact.

discovered that first-year teachers and expert teachers were both effective in that they provided similar instruction and students in their classes were equally successful. This, in turn, gave the researchers hope that novice teachers could indeed become effective at an early period in their educational careers. Using several statistical formulas, Rivkin et al. (2005) stated in their report

tracking mathematics and English scores in the state of Texas that "experience is not significantly related to achievement following the initial years in the profession" (p. 419). Further, teachers with advanced degrees do not statistically raise the quality of teaching. Beginning teachers performed the worst in the results of statistical equations; however, their data tended to suggest that novice teachers have the highest gains in effective measures in the first three years of teaching, but then begin to flatline in their effectiveness level later in their career. This would seem to indicate that experience does not always equate with a high level of skill in teaching.

Chingos and Peterson (2011) found similar results in their analysis of upper elementary school teachers in Florida and their effectiveness on math and reading scores over a seven-year period. The study looked at the teacher's certification methods as well as where the teacher graduated from school. Results indicated that advanced degrees and attendance at elite universities in undergraduate or graduate levels bore no significance in test scores. Further, early experience provided the greatest gains and experience past that either created a worsening in effectiveness or a plateau.

Marsh (2007) came to similar conclusions in his review of 195 university professors' student evaluation forms tracking effectiveness over 13 years. Professors showed very similar effectiveness levels mid-career as they did as beginning professors in undergraduate and graduate courses. Therefore, a poor instructor remained poor throughout the 13-year period, as did those with excellent student ratings. The professors were from a diverse population across the university studied; therefore, it would seem that their findings would indicate a global trend in student evaluations and other means of evaluation may be necessary to receive a true snapshot of a given instructor's effectiveness.

Kini and Poldolsky's literature review (2016) presented similar findings but the authors are more positive in their outlook regarding growth in effectiveness later in careers. Additionally, Kini and Poldosky's review stated that teachers of all levels increased in effectiveness when the climate of the school was supportive and collegial. Further, their review indicates that expert teachers should be involved in the development of novices and that schools should be spread equitably with both levels for optimal success in student achievement. Their review seems to indicate that there is hope for novice teachers in producing highly effective teaching behaviors and will continue to do so past the early years.

The Profile of a "Master Teacher" and a "Novice Teacher"

There are several studies which have primarily focused on finding qualities of "master teachers" or "effective teachers." Researchers have explored evidence of effective behaviors by analyzing video recordings (Babb, 2010; Beebe, 2007; Cavitt, 2003; Colprit, 2000; Duke & Buckner, 2009; Duke & Chapman, 2011; Duke & Simmons, 2006; Evans, 2012; Hendel, 1995; Kim, 2016; Millican, 2017; Patterson, 2009; Price et al. 1997; Roesler, 2013; Sheldon et al., 2008; Silvey & Baumgartner, 2016; Singletary, 2016; Taylor, 2009; Whitaker, 2017; Williams, 2016; Worthy, 2003, 2006; Worthy & Thompson 2009), surveys (Anderson-Nichols, 1997; Hamid, Hassan, & Ismail, 2012; Liu & Meng, 2009; Miksza, Roeder, & Biggs, 2010; Millican, 2009; Montemayor, 2014; Rohwer & Henry, 2004; Sheldon et al., 2008; Silvey and Baumgartner, 2016; Singletary, 2016; Taylor, 2009; Whitaker, 2017), short answer questions (Gao & Liu, 2013), audio recordings (Montemayor, 2014; Silvey, and Baumgartner, 2016; Yarbrough, Price, & Hendel, 1994), interviews (Hendel, 1995; Patterson, 2009; Whitaker, 2008),

field notes (Worthy & Thompson, 2009), and personal opinion from experience (Polk, 2009). Additionally, expert participants represented a variety of levels across a broad spectrum of ensembles as well as studio teaching. Some studies observed the same teacher with different levels of ensembles as well (Beebe, 2007; Worthy, 2003).

According to extant literature, effective teachers are able to organize, plan, and have a specific target in mind during a given lesson or period (Cavitt, 2003; Derby, 2001; Duke & Simmons, 2006; Millican, 2009; Millican & Forrester, 2018; Rohwer & Henry, 2004; Taylor, 2009; Worthy, 2003, 2006; Worthy & Thompson, 2009). Experts are relentless in their goals for their students and have consistently high expectations (Beebe, 2007; Cavitt, 2003; Duke & Simmons, 2006; Evans, 2012; Gau & Liu, 2013; Liu & Meng, 2009; Miksza et al., 2010; Roesler, 2013; Worthy, 2006). However, they are flexible and are able to change and/or deconstruct the trajectory of the original lesson target if a student is having difficulty with a given concept (Anderson-Nichols, 1997; Duke & Simmons, 2006; Gao & Liu, 2013; Hendel, 1995; Millican & Forrester, 2018; Singletary, 2016). Master teachers review work from previous lessons and consider prior knowledge before beginning new concepts (Berliner, 1986; Duke & Simmons, 2006; Rosenshine, Froehlich, & Fakhouri, 2002). Teachers also model appropriate playing or singing techniques (Babb, 2010; Beebe, 2007; Cavitt, 2003; Derby, 2001; Millican, 2017; Millican & Forrester, 2018; Sheldon et al., 2008; Singletary, 2016; Taylor, 2009; Worthy & Thompson, 2009). Effective teachers are also able to motivate students towards given learning targets (Derby, 2001; Evans, 2012; Miksza et al., 2010; Rohwer & Henry, 2004; Roesler, 2013) through enthusiasm and energy (Gao & Liu, 2013; Miksza et al., 2010) as well as excellent interpersonal skills (Gao & Liu, 2013; Hamid et al., 2012; Millican & Forrester, 2018; Polk, 2009) which furthers good rapport with students (Duke & Simmons, 2006; Millican, 2009) using

proactive strategies (Anderson-Nichols, 1997; Berliner, 1986; Worthy and Thompson, 2009). Expert teachers are good communicators and provide a variety of feedback styles (Cavitt, 2003; Derby, 2001; Duke & Simmons, 2006; Hendel, 1995; Price et al., 1997; Worthy and Thompson, 2009). Feedback styles which provide the most student success include specific and technical feedback (Berkley, 2011; Cavitt, 2003; Duke & Simmons, 2006; Hendel, 1995; Price et al., 1997; Worthy, 2006; Worthy & Thompson, 2009). In Cavitt's (2003) and Duke and Simmons' (2006) studies, negative feedback was used more frequently than positive feedback and provided specificity and were not meant to be considered as sarcastic or rude. Although positive feedback was used infrequently, episodes of positive feedback were lengthy. Additionally, students are successful because effective teachers know what their students can realistically accomplish (Duke & Simmons, 2006; Polk, 2009; Roesler, 2013) due to their in-depth knowledge of pedagogy and skill level of the student (Duke & Simmons, 2006; Millican & Forrester, 2018; Patterson, 2009; Polk, 2009; Worthy, 2006). They are able to accomplish a great deal in a given time period due to their focus on quick pacing (Anderson-Nichols, 1997; Cavitt, 2003; Colprit, 2000; Derby, 2001; Duke & Buckner, 2009; Duke & Chapman, 2011; Duke & Simmons, 2006; Patterson, 2009; Price et al., 1997; Rohwer & Henry, 2004; Taylor, 2009; Worthy, 2003, 2006; Worthy & Thompson, 2009) which allows for less moments of off-task behavior of students. This means that master teachers are also skilled at classroom management (Anderson-Nichols, 1997; Millican, 2009; Rohwer & Henry, 2004; Worthy & Thompson, 2009) and know what their students will do before they do it (Anderson-Nichols, 1997; Hendel, 1995). Additionally, effective teachers have a high capacity for learning and were most likely very good students themselves (Hamid et al., 2012; Liu & Meng, 2009; Polk, 2009).

Master teachers also knew which objectives were most appropriate for the grade level and/or experience of the ensemble taught. In Worthy's study (2003) of a master conductor who conducted the same piece with a high school honor band and a college ensemble, multiple targets were prevalent with the college ensemble while the same performance issues were addressed as single targets with the high school ensemble. The high school ensemble had a slightly quicker pace with more frequent teaching talk episodes than in the college rehearsal. Similarly, Singletary's (2016) study of five middle school band directors rehearsing a beginning and an advanced ensemble found that modeling and individualized instruction were observed more in the beginning ensemble. More opportunities for student questions and responses were found in the beginning ensemble as well. Lengthier performing times were noted in the advanced ensemble. Patterson's (2009) study of three renowned choir directors directing four types of ensembles (middle school, high school, collegiate, community church choir) found that directors focused more on developmental skills with beginning ensembles and on expressive elements with advanced groups. In an opposing vein, Derby's (2001) study of expert elementary, middle, and high school choral directors indicated that all directors used quick pacing and brief but frequent teacher talk and students were more successful versus unsuccessful in performance trials, regardless of the level. Similarly, Montemayor (2014) focused on before and after performance trials of 29 beginning band ensembles over a two-day period of the same piece. Findings indicated that short-term improvements were difficult to measure, although there was a positive correlation between rehearsal effectiveness and performance. This would seem to indicate short-term observations and evaluations of even the most gifted teachers may not provide an accurate portrayal of an ensemble or the teacher's efficacy levels. However, it would

seem from these studies that master teachers have deep pedagogical knowledge and can quickly adapt instruction to fit the needs of a variety of ensembles.

There are also skills which seem to be more important in the eyes of current educators. In Rohwer and Henry's survey (2004) of several levels of university music professors, participants selected classroom management, clear instructions, pacing, ability to motivate, and positive attitude as the top five qualities out of 14 skills and personality characteristics. Additionally, all participants selected expressive musicianship, error diagnosis, and sight-reading as the top three out of nine musical skills effective teachers encompass. Millican and Forrester's (2018) survey sent to expert college professors and public-school educators revealed that developmental knowledge and relationships with students were most important indicators in effective teachers. Polk (2009) opined that master teachers were creative in delivery, modeled well, and stayed current with the use of their instrument. Liu & Meng's questionnaire (2009) to teachers, students, and parents indicated that true teacher effectiveness was represented in high levels of student performance.

It is relevant to depict the qualities of a master teacher so that pre-service and novice teachers can comprehend and begin to approximate those behaviors early in development. However, it is also important to give a depiction of the novice teacher so that those who are teaching in higher education and those in the field as mentors can consider what pre-service and novice teachers require. It has already been stated that novice teachers can be effective and can demonstrate efficacy. Nevertheless, several studies have compared master teachers to novice teachers to better understand novice teachers.

There have been a small group of studies which compare novice teachers to expert teachers which concern music teachers specifically (Anderson-Nichols, 1997; Bergee, 2005;

Goolsby, 1996, 1997; 1999; Pike, 2014; Wagner and Strul, 1979) as well as various educational levels (Gatbonton, 2008; Kavanagh & Rainey, 2017; Livingston & Borko, 1989; Westerman, 1991). These studies utilized a small range of teaching episodes in relation to expertise levels. Apart from a smaller subset in the Goolsby (1997) study which longitudinally tracked differences in verbal instruction with undergraduate music education majors, most of these studies were not longitudinal and therefore compared heterogenous groups.

Anderson-Nichols (1997) compared novice and experienced elementary general music teachers. A survey was given to participants to rank important teacher practices and priorities. The researcher also analyzed video recordings, compiled field notes, and conducted interviews. Findings indicated that experts were more selective in their use of information regarding the classroom environment, had better management of classrooms, were more adept and intuitive at altering plans based on student performance, and had increased problem-solving skills. Novices had the same skills as experts in relation to lesson planning and prescriptive skills and were comfortable with diagnostic skills.

Bergee (2005) compared thought processes and rehearsal procedures of two undergraduate student directors, one masters student director, and one expert orchestral director. They conducted the same university symphony orchestra and rehearsed the same piece without stopping. The participants wore a microphone and attempted to verbalize their thought processes while conducting. Their verbalizations were recorded and their conducting was videoed. Additionally, participants were interviewed after conducting and an expert conductor critiqued their videos. Findings indicated that novices were overwhelmed with the activity and were unsure as to how to respond and found that multiple tasks were difficult to accomplish at once. The intermediate conductor provided no verbal remarks and was frustrated by the inability to

perform the task. He stated that he could easily identify mistakes of any performance if he were sitting and observing, but the difficulty was identifying them in front of a group. The expert was adept at not only conducting, but also teaching the ensemble. Comments made to the group related less to surface areas of musicianship but focused more on interpretation.

Goolsby (1996, 1997, 1999) conducted a series of studies related to how different levels used instructional time, teacher verbalization types, and rehearsal procedures. Goolsby (1996) observed 30 band directors (10 student teachers, 10 novice teachers, and 10 experienced teachers) teaching in varying socioeconomic populations over a period of three consecutive rehearsals. Data was used during the last two rehearsals to provide information which may have allowed for a higher comfort level. Rehearsals were measured in real time and were recorded in relation to preparation time, initial teacher talk, warm-ups, time for each musical selection, breaks, closure, and dismissal. Times were converted to percentages of the overall rehearsal. Results indicated that student teachers talked the most and students played the least. Novice teachers indicated lower verbal instruction, extended breaks, and less time on the first selection than the student teachers which indicated better pacing than student teachers. However, many percentages dipped in novice teaching with higher percentages in student teaching and expert teaching. Expert teachers provided the most break time, used instructional time equally, gave more student performance time, used non-verbal communication while students were more on-task and demonstrated the least amount of talking.

Goolsby's second study (1997) used the same data set from the previous study, but also included a second part in which pre-service teachers were videoed during their junior and senior years. Additionally, training was given to participants which included 15-minute videos of expert teachers in which students had to answer 10 open-ended questions. Both data sets were analyzed

using a form for time sampling and addressed one of 15 performance variables. Further, 10 rehearsal variables and comments were recorded using frequency distribution. Additionally, videos were analyzed for sequences of instruction found in each rehearsal. Findings indicated that student teachers asked vague questions and repeated sections of a piece with no stated instructional target and lacked feedback. Novice teachers displayed behaviors very similar to student teachers but had fewer vague questions. Novices had the lowest number of sequences of instruction in the study. Experienced teachers stopped more frequently and had more instructional targets. Most of their comments were in relation to rhythm and tempo. However, they also addressed expressive elements and tone quality more than their less-experienced colleagues. Experts asked specific questions to guide students during instruction. In a similar realm, the pre-service teachers saw a positive increase in desired behaviors. This may be because participants had guided exposure to expert conductors and knew what to emulate.

In the third series of Goolsby's studies (1999), 10 novice and 10 experienced band directors taught the same unfamiliar piece to their own ensembles over an academic quarter. Rehearsals were video recorded and analyzed in real time using a series of stopwatches. Frequency counts of 15 types of instructional targets and teacher verbalizations were collected as well as the number of sequential units of instruction. Additionally, the amount of rehearsals and the amount of time in which a piece was taught was analyzed. Results indicated that ensembles under experienced directors learned the piece using fewer rehearsals. Directors used more performance time and focused on expressive elements. Experts also provided more specific positive feedback and more sequential units of instruction. Novice teachers used more verbal instruction and had higher amounts of overall teacher talk. Unguided listening for ensembles was

higher and non-purposeful repetition of sections was prominent. Vague questioning was also observed in novice teaching.

Pike (2014) performed a case study in which one novice and one expert teacher were observed teaching a children's group piano course at a performing arts academy. No specific guidelines were given to participants regarding structure of the lessons or the study itself. Data was triangulated by using in-class observation, teacher questionnaires, lesson plans, teacher interviews, and member checks. The researcher used a constant comparison method of analysis to saturate the data collected. Constant comparison is a method used by qualitative researchers in order to receive a full scope of a grounded theory. A grounded theory provides in-depth information about a given phenomenon. In this study, the novice teacher had no prior experience with group piano while the expert had taught the course for several years. Findings indicated that the novice gave less time to the students (a six-week course which fit their schedule) with no focus on practice outside of the group piano time and used one instructional book. Students in the novice group played 18% of the time and spent 36% of the time with music theory and 27% with score preparation. Participants gave a recital at the end of the six weeks which suggested low prep-time and many mistakes in performance. The expert teacher created a course which used twice as much time (12-week course) with expectations that the students in the group would practice outside of class time and complete a practice log. Playing time for students in the expert's group was 59% with a small focus on improvisation and games. The expert teacher gave students more opportunities for rapport among the group than the novice who treated their participants more like students in a private studio. In the final performance, the expert utilized background tracks with individual students and duets with the teacher to cover up any possible mistakes as well as providing a further level of comfort. While two of the participants in the

novice teacher's group dropped out before the six-weeks were over, none of the expert teacher's participants left the course.

Wagner and Strul (1979) observed pre-service teachers, student teachers, and experienced teachers twice during the scope of 15 minutes and used Moore's observation form (MTRA) to collect data. The form gave grids for time sampling and frequency counting. Results indicated that academic instruction increased and verbal directions decreased over the three levels of expertise. Singing and movement activities were used more prominently with experts. Preservice teachers lost control of students the most and episodes of non-teaching were the least in expert teachers. Experts used negative specific feedback more and decreased in approving social mistakes and academic mistakes. All of this indicates that expert teachers in this study had deeper pedagogical knowledge and greater skill in classroom management.

Schleuter (1991) studied three student teachers curricular thinking in teaching elementary general music using a case study model. Findings indicated that participants were more concerned about students' enjoyment levels than curricular understanding and the participants had difficulty transferring skills learned in pre-service teaching to the student teaching setting. Further, participants were locked in a framework of objectives and planning lessons around the objectives but were not utilized in a sequential fashion. Participants also focused on large-group instruction over individualized growth and differentiation and had lower expectations in student performances than did their cooperating teachers.

Gatbonton (2008) studied four novice English as a Secondary Language (ESL) teachers and reviewed their data in relation to a previous study using expert teachers. Both groups taught a 4-week ESL module to a classroom of adult learners. Participants taped their recollections after every lesson regarding the lesson. The tapes were coded and frequency counts for each code

were created. Expert teachers were focused on language management, student progress, and level checks of students. Novice teachers, however, were more concerned about how well they knew the students and the students' behaviors. Additionally, self-critique was an issue discussed more in novice teacher recollections. Expert teachers focused on providing input and anticipating future tasks as well as monitoring each student, whereas novice teachers were concerned about their explanations and students' negative reactions.

Kavanagh and Rainey (2017) focused on how instructional scaffolding within an alternative education program can assist novice teachers in learning to teach effectively. Participants were 12 novice teachers with no educational background enrolled in an alternative education summer program in which participants team taught in groups of four and received methodology training with a master teacher. Video recordingsof teaching in methodology classes as well as team teaching lessons in regular classrooms were analyzed. Analysis indicated that participants were successful when given small scaffolding steps and indicated higher signs of effectiveness in the team-teaching model after being given those steps.

Livingston and Borko (1989) studied three elementary or secondary math student teachers and their cooperating teachers over a one-week period. Participants were interviewed before and after each lesson. Analysis of recorded interviews and field notes provided the appropriate data. Experts used no written lesson plans but utilized mental plans instead. Most of the planning occurred outside of the classroom. Field notes indicated that experts kept their plans on track and used students' thoughts and questions as springboards into other ideas. Improvisational teaching was apparent throughout each lesson and was based on student needs. Reflections on post-teaching were concise. On the other hand, novice teachers had very detailed written plans and had rehearsed their plans prior to teaching them. Field notes indicated that they

were not as successful as their cooperating teachers in that they were unsure as to what to do with student questions. Questions were usually relegated to the next lesson. Reflections were also less concise and had a wide scope of thoughts.

Similarly, Westerman (1991) studied five elementary student teachers and their cooperating teachers with a focus on decision making within teaching. Participants were video recorded during teaching and interviewed before, after, and several months later. Participants watched their video immediately after teaching and again several months later. Printed materials were also analyzed. Coding of behaviors and frequency counts were the data points. Findings indicated that experts knew their students and had prior knowledge of global student reactions to problematic issues, were proactive in their planning, and were flexible in their lesson plans. Novice teachers focused on curricular objectives and followed them literally whereas experts used them as a guide. Further, novice teachers taught lessons procedurally and wanted to finish their lesson plans. Student behavior for novices provided difficulty and stopped the flow of lessons while experts fixed behavior issues without hindering student outcomes.

All of these studies indicate that student teachers and novice teachers spend more time using teacher talk than expert teachers (Goolsby, 1999; Pike, 2014; Wagner & Strul, 1979). Bergee (2005), Pike (2014), and Schleutter's (1991) novice teachers focused on music fundamentals over other performance aspects and lacked high expectations. Schleutter's (1991) student teachers were more focused on student enjoyment than cognitive process. Novice teachers in Goolsby's 1997 study spent most of the time tuning individual instruments whereas student teachers spent much of their time focused on fixing wrong pitches. Similarly, Pike's (2014) novice teacher lacked time management skills and had 5% more wasted time in rehearsal than the expert. In Goolsby's 1999 study in which novice and expert teachers taught the same

piece, novices started and stopped more often without providing feedback while experts spent most of their time in performance and addressing interpretative elements such as balance, style, tone, and intonation. Similarly, in the 1997 study, directors also focused on the use of modeling and guided listening. In the 1996 study, experts gave equal time to all parts of the rehearsal and gave lengthier break times. Of particular interest is Goolsby's finding in the 1997 study in which the greatest gains were found in student teachers in that their lessons moved from the focus of one-dimensional knowledge of wrong notes to working more with students on interpretative areas as well. Additionally, the amount of sequential instruction tripled in student teachers over time. Similarly, the intermediate conductor in Bergee's (2005) study focused less on superficial learning and conducted more artistically but did not verbalize. However, Schleutter's study (1991) indicated the opposite in transition of skills from pre-service to student teaching. Therefore, some studies seem to indicate that student teachers can increase effective teaching behaviors and may strengthen the concept that effective teaching improvements begin early in many educators' careers (Chingos & Peterson, 2011; Kini & Poldosky, 2016; Marsh, 2007; Rivkin et al., 2005; Siedentop & Eldar, 1989; Rosenshine, 2008).

Similar findings were seen in other educational settings as well. Westerman (1991) and Livingston and Borko (1989) found in their case studies that elementary and secondary student teachers focused on specific lesson objectives and were process-related in their teaching. Both studies indicated difficulty in knowing what to do with student questions and responses and how to adjust teaching. Westerman's study (1991) indicated that student teachers lacked understanding of specific student needs. In a more hopeful vein, Kavanagh and Rainey (2017) and Gatbonton (2008) who explored language-arts secondary and ESL novices found that novices have deep pedagogical knowledge but give less detail than experts. However, they can

encourage student success if they have been given the tools of scaffolding in teacher training. Livingston and Borko (1989) also found that novices were able to modify pacing based on student needs; however, the difficulty positioned itself in knowing what to do after pacing had been adjusted.

In general, it would seem from these studies that novice teachers tend to rely heavily on teacher talk to get points across during lessons, are process oriented, lack depth of understanding in feedback, and at times may not know how to respond to individual student needs. Data regarding time management is mixed, as some were able to adjust pacing accordingly while others tended to waste time on specific details which could be adjusted with appropriate pedagogical techniques.

Authentic Context Learning

Novice teachers may be coached to imitate the behavioral characteristics and utilize the strategies of "master teachers," placing them on a higher trajectory of effectiveness during teacher training programs. Several studies have been performed to determine if certain teaching methods are effective in improving skill for long-term affect (Blackwell & Roseth, 2018; Bowers, 1990; Butler, 2001; Byo, 1990; Cassidy, 1990; Collins, 1978; Colwell, 1995; Fant, 1996; Goodrich, et al., 2017; Haston & Russell, 2012; Killian, Dye, & Buckner, 2008; Nápoles, 2016; Paul, 1998; Paul et al., 2001; Powell, 2014; Powell, et al., 2018; Worthy, 2005). While an exhaustive study of researched techniques and course pedagogies is a worthwhile endeavor, studies discussed are ones which encourage self-reflection as the researcher's focus is authentic context learning (ACL) such as peer teaching, the use of video reflection in correlation with

teacher effectiveness, and studies which focus on modifying typical novice teacher behavior. These behaviors include extended teacher talk (Nápoles, 2016; Worthy, 2005), conducting style (Byo, 1990), and use of sequential training and/or rehearsal frames (Bowers, 1990; Butler, 2001; Cassidy, 1990; Collins, 1978). Students seem to increase in effective teaching behaviors when provided with ACL experiences such as peer teaching as well as field teaching.

Paul et al. (2001) investigated 30 undergraduate instrumental music education majors from four large universities and studied their relationship with four types of ACL experiences—field teaching, peer teaching, watching videos of one's teaching, and watching the student videos with an advisor. University professors reported the number of times each participant participated with the four experiences. Participants submitted a 10-minute teaching episode with a large-group ensemble which represented their best teaching. The episode was recorded during the first three to four weeks of student teaching in order to look for the effect of pre-service ACL experiences versus other effects if video recorded later in the semester. Hamann et al.'s (1989) Survey of Teaching Effectiveness (STE) form was utilized by three adjudicators and their average score was used for assessment. The STE score was formulated with the scores from the four ACL experiences. Findings indicated that effectiveness increased with the use of early field teaching, peer-teaching, and watching teacher videos. Watching videos with an instructor did not produce higher ratings than the other three experiences which contradicts Goolsby's (1997) and Livingston and Borko's (1989) studies which used guided watching of videos in lessons with positive results in participants. Haston and Russell (2011) studied five preservice teachers' journals and interviews from a university lab school's volunteer band and strings program and found that preservice teachers reported greater identity and increased confidence when given the opportunity to teach in the field for one to two years. Performance and education majors were

required to teach two years in both groups while the education majors were required to teach for one year, but also with both groups. Similarly, Baughman and Baumgartner (2018) studied four undergraduate music education majors and their involvement in an adult community choral ensemble and reviewed participant questionnaires, video-logs, and interviews. Results indicated increased confidence in teaching skills and a better grasp with specific pedagogical content than before this experience.

Powell (2014) measured student's concerns in relation to two types of ACL. Four senior-level preservice teachers were surveyed regarding their initial experiences in an instrumental course which required four peer teachings and two field teachings. The study focused on impressions in the first peer teaching and field teaching in a middle school band setting. The same lesson was taught in the initial field teaching as the first peer teaching. Observation forms, field notes taken by the researcher, interviews with the participants, and interviews with the cooperating teachers in the field provided data collection points. Participants felt that initial peer teaching assisted with technical issues but that it lacked authenticity, as students in the course had different abilities than students in the field. Participants perceived peer feedback as positive and that it assisted with delivery. Field teaching gave participants the authenticity which was lacking in peer teaching. However, participants were unclear as to how to respond to problematic issues such as playing technique and fingerings. Frustrations continued in that middle-school students were confused about vague statements and directives made by participants. Nervousness was also an issue reported in field teaching and participants lacked a real identity as a teacher.

Paul (1998) investigated the effectiveness of peer teaching in a similar case study by focusing on three participants who were currently student teaching or finishing student teaching. Participants had four semesters of peer teaching which included two semesters of conducting and

two semesters of instrumental methods courses. Peer teaching segments were video recorded and audio recorded and were critiqued by participants to further planning of the next lesson. Video recordings were also critiqued with instructor guidance. The instructor also provided modeling, small group, and individual work in the courses with peer teaching. Participants watched their videos again during the student teaching semester and were interviewed. The participants indicated that they matured in their self-image throughout the two semesters and felt that technical skills had improved since peer teaching. They felt that the technical skills in planning assisted them greatly in student teaching, as teaching skills became more intuitive and they were able to identify problematic passages for students. However, peer teaching was stressful to two out of three participants who felt that they were being judged by their peers for their musicianship as well as their pedagogical knowledge. One participant felt that more focus in pre-service needed to be in management and overall teaching techniques.

Schmidt's (2010) study about learned experiences in relation to John Dewey's theory of experience and teacher learning focused on six graduates of an undergraduate instrumental music education program. The organization of pre-service experiences were peer teachings and field teachings over a four-semester sequence. Data collection during this period included video recordings, written self-assessments, instructor assessments, and other written assessments. During student teaching, participants attended four seminars which assisted students in decompressing information gleaned from student teaching. Those episodes were also videoed by the researcher as data collection points and were coded. There was a positive interaction in peer teaching in relation to sequencing and planning. Five out of the six participants felt that university-selected field teachings were not helpful in that most of the time period was spent in passive observation but felt that field teachings selected by participants were more helpful.

Student teaching was problematic for some participants in that some were distracted by university observers or they fell back to bad habits which had been curtailed in preservice teaching.

Fant's (1996) study was similarly modeled after Paul et al.'s (2001) study in that 40 student teachers from 11 universities were asked to submit a video recording of themselves early in student teaching and then were interviewed regarding their preservice experiences in curricular ACL and noncurricular ACL experiences. Curricular ACL experiences included observation, microteaching, and tutoring. Noncurricular ACL experiences included participating in private lessons, being an ensemble director or section leader, and non-music teaching. An evaluator watched participants' videos and used Hamman and Baker's STE form (1996) and Bergee's Rehearsal Effectiveness Scale. The scores were averaged with the scores from ACL experiences. Results indicated that early field experiences with feedback and peer microteaching provided the highest scores. Early field teaching, regardless of feedback, noncurricular experiences, and musical background of participants, were not related to student teacher performance. However, early field experiences without feedback were negatively correlated to student teaching evaluations. Participants felt that courses like methods labs were good places for peer teaching.

Goodrich et al. (2018) examined the effectiveness of peer mentoring, a different type of ACL not usually examined in research. Twenty-six students in one elementary music methods course using four types of peer mentoring were examined. Participants were interviewed mid-semester and late semester, peer mentoring segments were videoed, and post-course questionnaires were given. Data was coded and researchers used a constant comparative analysis to saturate data. Similar to Pike's (2014) qualitative study, constant comparative analysis was

used to assist in creating a grounded theory for this study on peer mentoring. The mentoring segments included feedback in student-led warmups, small group instruction in guitar and recorder, and a small group microteaching. Results indicated positive remarks regarding feedback in student-led warmups. However, there were mixed findings in relation to small group work in guitar and recorder in that some felt they would have learned better in an individualized or large group setting. Participants were the most positive regarding microteaching in which five students took different roles—one teacher, three students, and one mentor who provided written feedback to the teacher.

Findings from these studies indicated overall positive experience in ACL settings. Peer teaching provided technical assistance in lesson planning and methodology but lacked authenticity. Some studies indicated that participants were comfortable teaching their peers while others indicated the opposite. The type of ACL experience was significant in that ACL experiences were positive if proper scaffolding was given which included feedback. ACL experiences also provided more awareness of teaching behaviors which provided further scaffolding for success when participants moved on to teaching in the field.

Many studies have reviewed the effectiveness of specific learning objectives within the peer learning context. Bowers (1990) investigated the effectiveness of two different types of training in sequential patterns of instruction. Participants were part of three sections of a music methodology course for elementary teachers. The three sections were given treatments in relation to five peer teachings in the semester. One class received no treatment, the second class received passive training in systematic instruction, and the third class received active training in systematic instruction. Passive training included written activities in relation to sequencing and active training included guided watching of a video using sequential patterns and then providing

mini-teaching activities to learn how to use sequencing. Evaluators watched the final peer teaching video and evaluated using a summative form. Results indicated no significant difference between the groups; however, the class with the active sequencing treatment received the highest scores. Additionally, time was used differently between the treatment groups. Participants with sequence training provided more specific feedback than those without.

Similar results were found in Collins' (1978) study in which two groups of 10 were given treatment in enthusiasm training. One group received no treatment and the other group was given a two-week intensive training period with peer teaching and microteaching. Both groups were video recorded prior to any training and were recorded again after the training period. A second post-test video was performed several weeks after the initial post-test video. Video recordings were time-sampled and coded using an evaluation form. Findings indicated that participants with enthusiasm training increased significantly in enthusiasm in both post-tests whereas the control group's scores remained similar throughout all three video recordings. Cassidy (1990) provided similar enthusiasm training to elementary majors in a music course and provided students with peer teaching experiences as well as field experience at a preschool. While all participants' efficacy levels increased in the field teaching experience, participants with enthusiasm training provided more interactive lessons than those without the training.

Blackwell and Roseth (2018), Butler (2001), Colwell (1995), Killian et al. (2008), Nápoles (2016), and Worthy's (2005) studies do not structurally provide for a control group. Rather, the purpose of their studies was to track progress throughout a given time period in peer teaching episodes. Blackwell and Roseth (2018) concentrated on the use of problem-based learning in a secondary instrumental methods course. Assessment of teaching videos, scenario discussion, and peer teaching were utilized as tools in problem-based learning style. Interviews

and open-ended writing throughout the semester indicated that participants were more comfortable in pedagogical content and appreciated peer-teaching episodes. Butler (2001) concentrated on the usefulness of concept mapping in relation to microteaching experiences and found that the two methods in conjunction may provide for higher efficacy levels in the future. Colwell (1995) provided a treatment for a music course for elementary teachers in which groups taught mini-lessons to their peers or to a class of kindergartners with general or specific self-evaluation tools. Participants in both mini-lesson treatments increased in effectiveness. Killian et al., (2008) found in the study of four unrelated peer teaching episodes of the same students that student interaction improved and teacher talk decreased throughout. Similar findings were seen in Worthy's study (2005) of students' self-reported levels of teacher talk and student performance in four related peer teaching episodes by using Duke's (2019) Simple Computer Recording Interface for Behavioral Evaluation also known as SCRIBE. SCRIBE is a behavioral observation software in which large-scale behaviors including teacher talk, teacher modeling, and student performance (Beebe, 2007; Cavitt, 2003; Colprit, 2000; Singletary, 2016; Taylor, 2009; Worthy, 2003, 2006; Worthy & Thompson, 2009) are initially analyzed for frequency and duration, percentage of total time, mean, and rate per minute. Students reported an increase in playing and singing time and a decrease in teacher talk throughout the peer teaching episodes. In relation to teacher talk, Nápoles (2016) utilized a "Rule of Seven" in which students had to give instructions to participants in peer teaching in less than seven words. While at times this device seemed constrictive to students, it did provide the opportunity for participants to be self-aware in an area where many novice teachers struggle.

How Effectiveness is Measured

A tool frequently used in measuring effectiveness is video and/or live observation of teachers of varying degrees of expertise. While some studies use the researcher solely as the observer, other studies measure large populations of varying levels to indicate observational trends. Results of studies with large observation populations will not be discussed, as perceptions of large populations is not the focus of the study. However, the methodology of these studies can assist in comprehensive triangulation of data sets. Studies which have a large population as observers tend to utilize summative evaluations (Nápoles & MacLeod, 2011, 2013, 2014, 2015; Whitaker, 2008; Silveira, 2014; Byo, 1990; Madsen, 1999; Redding, 2011; Yarbrough et al., 1994), time sampling observation sheets (Babb, 2010; Byo, 1990; Hughes, 1992; Madsen, Standley, & Cassidy,1989; Nápoles, 2006; Yarbrough & Price, 1981), and frequency count forms (Goolsby, 1997; Nápoles, 2006; Whitaker, 2008; Yarbrough & Price, 1981) for specific behaviors. Other researchers use a mixed methods format which include field notes (Bergee, 2005; MacLeod, 2017; Pike, 2014), interviews (Bergee, 2005; Pike, 2014; Whitaker, 2008), short answers (MacLeod, 2017; Madsen, 1999; Silveira, 2014), continuous response digital interface (CRDI) (Silveira, 2014; Colwell, 1995), and statistical tests to triangulate data with summative evaluation forms.

Nápoles and MacLeod performed a group of studies (2011, 2013, 2014, 2015) in which summative evaluation forms were used immediately after viewing expert teaching segments of varying lengths for focus areas such as feedback, teacher high/low delivery, student progress, student engagement, and teacher high/low engagement. Most of their forms utilized a 5-point Likert scale with anchor terms of agree/disagree or very low/very high. Due to the specific

behavior focuses in each study, lessons were scripted and were not part of a regular teaching day. Behaviors observed included teacher demeanor, competency, feedback, modeling, positivity levels, delivery, student progress, musicianship level, and overall effectiveness of the lesson. Evaluators in the study usually gave higher scores to teachers with high enthusiasm, regardless as to whether information and feedback was accurate. Whitaker (2008) asked participants in their study (band directors and their students) to complete evaluation forms that included a 10-point Likert scale and identified effectiveness levels in relation to pacing, information presentation, musical information, feedback, verbal clarity, use of voice, conducting gesture, enthusiasm, and overall effectiveness. Silveira (2014) utilized a summative form as well as a continuous response digital interface (CRDI) but did not provide detail as to behaviors observed on the form. The CRDI is a device used to continuously rate on a semantic differential scale in real-time. College students observed a scripted lesson with poor pacing with four different foci while observing—ongoing teacher effectiveness, teacher intensity, teacher pacing, or general perceptions. All participants responded similarly and provided negative answers to pacing issues. Other studies only asked participants to give an overall intensity or effectiveness rating over a 10-point Likert scale after observing a lesson or a short conducting segment (Byo, 1990; Madsen, 1999; Redding, 2011). These studies also used time sampling information sheets or a requested comment page as well to encourage more rigorous data collection. In a similar vein, Yarbrough et al. (1991) had participants evaluate 20 short segments of a scripted choral rehearsal and gave a numerical grade for each segment and then a letter grade to match (A+, A, A-, etc.).

Another frequently used collection device for large populations included checklists, frequency counts, and time sampling sheets. Babb (2010), Byo (1990), Hughes (1992), Madsen, Standley, and Cassidy (1989), Nápoles (2006), and Yarbrough and Price (1981) used time

sampling sheets in a variety of ways. Most of the studies mentioned above used time sampling in 10- or 15-second intervals for participants to mark specified behaviors including off- and on-task student behavior (Hughes, 1992; Nápoles, 2006), high or low intensity conducting gesture (Byo, 1990), or high or low teacher intensity (Madsen et al., 1989; Yarbrough & Price, 1981). However, Babb used 5-minute teaching segments to code three categories of voice building techniques (2010) and used SCRIBE for data collection. Similarly, Goolsby's series of studies (1996, 1997, 1999) used a stopwatch to document timing of particular behaviors. Coding behaviors is another data collection device frequently used (Goolsby, 1997; Nápoles, 2006; Whitaker, 2008; Yarbrough & Price, 1981) in combination with other means of collection. Whitaker (2008) coded conductor behaviors such as nonverbal agreements and disagreements, "catch phrases," enthusiasm, and overall impressions in interview questions with band directors and their students. Yarbrough and Price (1981) coded locations of eye contact (group, individual, music, or other) and parts of learning sequences while Nápoles (2006) coded types of teacher talk including academic, reinforcement, directives, and questions. Similarly, Goolsby (1997) coded types of teacher behaviors in varying levels of band directors such as modeling, feedback, and questions as well as instructional targets.

Several studies triangulate data and use various data collection types. Interviews (Bergee, 2005; Pike, 2014; Whitaker, 2008), short answers (MacLeod, 2017; Madsen, 1999; Silveira, 2014), field notes (Bergee, 2005; MacLeod, 2017; Pike, 2014), questionnaires (Whitaker, 2008; Pike, 2014), and lesson plans (Pike, 2014) are commonly utilized to saturate data. Of interest is Silveira's use of the CRDI to measure teaching effectiveness in relation to a scripted lesson with purposeful pacing lapses. Silveira labeled both sides of the device with good/bad anchors.

Colwell (1995) also used the device in her study along with behavioral checklists as a means of student self-evaluation during an elementary methods course.

In relation to the current study, another popular way to analyze teaching is by analyzing data using rehearsal frames to identify segments for analysis. The rehearsal frame was introduced by Duke (1994) and grew out of the idea that response to student performance and achievement had been underutilized as a measurement instrument. Systematic observation and direct instruction as discussed by Yarbrough and Price (1981, 1989) had also been used as an assessment instrument. Direct instruction includes three sections—task presentation, student response, and teacher feedback (Yarbrough & Price, 1981, 1989). Systematic observation has been quite effective in identification of day to day music making routines in interaction with teacher and student. However, rehearsal frames take systematic observation a step further and view interactions as dependent measurement which focuses on achievement (Rosenshine, et al., 2002). Duke's purpose in the framework is to "focus attention on the process of effecting positive change in the performance of a musical work. The nexus of the model's structure is the musical outcome—the achievement of tangible musical goals. (1994, p. 83)" Furthermore, teachers who desire to change an aspect of performance usually need to do so by involving "a more elaborate instructional process in which the conductor leads the ensemble musicians through a sequence of performance episodes in an effort to improve some aspect of the performance of a piece. (Duke, 1994, p. 84)" Therefore, in a rehearsal frame, a teacher identifies and verbalizes the problem or target which needs improvement. The students provide performance trials in relation to the target. In response, the teacher may choose to reduce the complexity of the task and identify individuals having difficulty. If the target has not been met, they may choose to decontextualize or remediate. The group then demonstrates the target without

assistance. If the target is performed correctly, the teacher will recontextualize the problem and insist on appropriate performance of the target (Duke, 1994, 1999/2000). Sometimes, targets can be accomplished through one rehearsal trial; however, two rehearsal trials seem to be a good indicator of moments when student behavior has been adjusted (Worthy, 2003).

Several studies have been performed using rehearsal frames to sample longer teaching episodes to identify master and novice teacher behaviors (Berkley, 2011; Beebe, 2007; Cavitt, 2003; Chapman, 2014; Colprit, 2000; Duke & Buckner, 2009; Duke & Chapman, 2011; Henninger, 2002; Kim, 2016; Millican, 2017; Montemayor, 2014; Patterson, 2009; Roesler, 2013; Singletary, 2016; Taylor, 2009; Whitaker, 2017; Williams, 2016; Worthy, 2003, 2006; Worthy & Thompson, 2009). The majority of studies examined band directors; however, rehearsal frames can be used to observe choral directors (Derby, 2001; Patterson, 2009), orchestra directors (Williams, 2016) elementary Orff ensembles (Taylor, 2009), non-traditional ensembles (Kim, 2016), and studio lessons (Colprit, 2000; Duke & Buckner, 2009; Duke & Chapman, 2011).

Analyses of rehearsal frames have usually included identifying and categorizing instructional targets. Instructional targets mentioned in studies have included vocal production, articulation, diction/pronunciation, dynamics, intonation/tone, pitch accuracy, rhythmic accuracy, technical facility, text emphasis/word stress, tempo, editorial, balance/blend, notes, phrasing, tone, imitation, multiple targets, and unidentified target (Beebe, 2007; Berkeley, 2011; Cavitt, 2003; Patterson, 2009; Worthy, 2003, 2006; Worthy & Thompson, 2009). Analysis typically uses SCRIBE, a behavioral observation software in which large-scale behaviors including teacher talk, teacher modeling, and student performance (Beebe, 2007; Cavitt, 2003; Colprit, 2000; Singletary, 2016; Taylor, 2009; Worthy, 2003, 2006; Worthy & Thompson, 2009)

are initially analyzed for frequency and duration, percentage of total time, mean, and rate per minute. A second viewing categorizes verbal behaviors including information statements, directive statements, nonspecific positive feedback, nonspecific negative feedback, specific positive feedback, specific negative feedback, questions, and off-task statements (Beebe, 2007; Berkely 2011; Cavitt, 2003; Patterson, 2009; Taylor, 2009, Worthy, 2003, 2006; Worthy & Thompson, 2009). Teacher modeling behaviors observed are positive modeling, negative modeling, and teacher performance with the ensemble (Beebe, 2007; Berkely 2011; Cavitt, 2003; Patterson, 2009; Taylor, 2009, Worthy, 2003, 2006; Worthy & Thompson, 2009). Student performance behaviors include whole group, small group, individual, performance approximations, and student talk (Beebe, 2007; Berkely 2011; Cavitt, 2003; Patterson, 2009; Taylor, 2009, Worthy, 2003, 2006; Worthy & Thompson, 2009).

Researchers who use rehearsal frames as the unit of analysis have reported very similar results to previous studies which use summative evaluation forms, coding of participant responses, and field notes. It can be inferred from that expert educators know and understand their populations and will use methods which fit the ensemble (Cavitt, 2003; Singletary, 2016; Patterson, 2009; Worthy, 2003), provide a quick pace (Kim, 2016; Taylor, 2009; Worthy, 2003, 2006; Worthy & Thompson, 2009), model well (Beebe, 2007; Kim, 2016; Worthy & Thompson, 2009), and ask fewer questions (Berkely, 2011; Derby, 2001).

It seems that each type of observational system (summative evaluation, questionnaire, coding, time sampling) will most likely provide slightly different types of information. Therefore, perhaps a way to saturate data from video recordings may be to use a combination of two observational systems. Therefore, it is the goal in the current study to utilize a variety of data collection methods.

Longitudinal Studies for Effectiveness

The use of longitudinal studies to track effective teaching behavior in early career levels using video are unknown to the researcher currently. Only one study's (Goolsby, 1997) partial purpose was to view types of verbalizations in video recorded rehearsals ranging in junior to senior year rehearsals. Researchers using a longitudinal method tend to employ qualitative measurements. There are several studies which reflect teachers' self-reported efficacy levels and transfer of learning longitudinally from preservice to student teaching and/or student teaching to novice teaching (Bartolome, 2017; Clark, Byrnes, & Sudweeks, 2014; Dabback, 2018; Fant, 1996; Hoy & Spero, 2005; Killian & Dye, 2009; Miksza & Berg, 2013; Powell, 2018). Goldhaber, Krieg, and Theobald (2017) used data from six Washington state area teacher education programs to discover efficacy levels from student teaching to where teachers are presently. The data indicated that participants were more successful if student teaching demographics were similar to where they were currently teaching. There is a large body of studies which primarily concerns itself with self-efficacy levels and concerns of pre-service and novice teachers. However, that is not the purpose of the study.

Bartolome (2017) tracked a group of teachers for two and a half years using semi-structured interviews with their experiences in service learning, practicum, student teaching, and transfer in the first year of teaching. Participants felt that ACL experiences in service and practicum learning benefitted them greatly in that it provided experience and opportunities to learn how to plan lessons. Additionally, ACL experiences provided pre-service teachers with more comfort going into student teaching as well as into the first year of teaching. Student teaching was also beneficial in that students learned the day-to-day workings of school such as

meetings, paperwork, and knowledge of individualized education programs (IEP). Transfers included comfort level with observations and the ability to self-reflect. However, certain deficits were found in first-year teacher's reflections in that there was no preparation for understanding the isolation of the job and improvement of people skills to work with difficult faculty and administration.

In opposition, Powell (2018) performed a case study tracking student teaching experiences to first-year experiences with band directors using observations and interviews. The qualitative data indicated lack of transfer and control in student teaching from experiences in preservice courses. Student teachers expressed their cooperating teacher's distrust in student teachers' skills and did not allow them to work with certain ensembles resulting in lack of agency. Transfer of these thoughts occurred in the first years of teaching as well, specifically with those who became assistant band directors. Further, reflective thought indicated similar views regarding lack of agency, indicating that the pattern of lack of independence may continue in further student teaching episodes.

Dabback (2018) used a case study design and tracked three student teachers to their fourth year in teaching. He used interviews and journals to track identity and transfer of knowledge. All three expressed lack of preparation in preservice courses and student teaching regarding classroom management and the ability to realize the border between healthy versus unhealthy student/teacher relationships. These non-musical skills would certainly create an environment for lack of efficacy.

However, transfer of effectiveness occurred from student teaching to novice teaching in Goldhaber, Krieg, and Theobald's study (2017) in which math scores were measured in relation to elementary teacher's student teaching placement. Their findings indicated that teachers who

taught at schools with similar demographics as the ones in which they completed student teaching were more effective than those with different demographics. This would seem to imply that teacher placement is essential in providing opportunity to practice teaching skills for transfer.

Clark et al. (2014) focused on differences between one-year internships and 15-week student teaching episodes in preservice and student teaching. While it would seem that a lengthier time period would be more effective in adjusting teaching behavior, student teaching provided more modeling and verbal support and also gave participants a higher sense of efficacy. However, this perception in student teaching may be inflated, as Hoy and Spero (2005) and Killian and Dye (2009) found in their quantitative and qualitative measures. Hoy and Spero tracked participants in a master's degree program in early education and found consistently through four quantitative measures that efficacy levels increased in student teaching, but significantly decreased in the first year of teaching. Although Killian and Dye focused on the journey from peer teaching, field teaching, and student teaching, their qualitative data indicated that with experience, participants were becoming more realistic in their perceptions of their own effectiveness. In their journey, students also became more aware of the importance of lesson planning and sticking to their plan versus veering away from the objectives. Similar to Fant's findings (1996), participants wanted more feedback from their instructors. Additionally, participants felt that their journey assisted them greatly in skill development which would obviously increase effective teaching behaviors moving into the novice years. In a related study, Miksza and Berg (2013) found that skill development changed perceptions from being concerned with competence only to being concerned about techniques with more nuance. Additionally, students adjusted from a self-focused perspective to one on student impact.

Longitudinal studies indicate mixed findings in relation to the ability to transfer knowledge from methods classes to student teaching and then to in-service teaching. In-service teachers felt that more courses needed to be focused on administrative parts of the job, more feedback from instructors was necessary, and additional ACL experiences would be beneficial. Of separate interest is how/where students teach in placements and how this may affect their preparation as an in-service teacher.

Need for the Study

Further exploration in understanding the development of teacher effectiveness in the novice years of teaching is needed in that the highest gains of effectiveness occur in the first five years of in-service teaching (Anderson-Nichols, 1997; Berliner, 1986, 1988). There have been no longitudinal studies performed that investigate effectiveness in early career levels. There is a great divide in teaching skills which differentiate expert and novice teachers. However, novice teachers can be as effective as those with more experience if the appropriate skills have been developed in their preservice courses and student teaching. Novice teachers have identified that ACL experiences may increase effectiveness and student teaching may further ground appropriate behaviors encouraged in preservice. However, it has also been reported that novice teachers and student teachers may have difficulty transferring knowledge from the different early career levels.

Longitudinal studies utilizing video as the main source of analysis for measurement of effectiveness are unknown currently. Using videos to assess teacher behavior is well-documented and can provide more information about novice teachers' pathway toward becoming experts.

Using a multi-methods approach to collect data may solidify previous data relating to behaviors of teachers in peer-teaching, student teaching, and novice teaching. Specific behaviors have been identified in prior research which are problematic for early career teachers. These include pacing (flow), feedback, teacher talk, and modeling. Therefore, the study will attempt to identify trends in these behaviors across the early career levels.

The questions that will guide the study are:

1. What are the frequencies of instructional target categories in rehearsal frames of the three early career levels?
2. What are the frequencies and durations of specified teacher and student behaviors observed in selected rehearsal frames?
3. What are the ratings of teaching effectiveness at each level of teaching and individually?

CHAPTER III

METHODOLOGY

The purpose of this study was to observe the teaching of five novice music teachers in three early career teaching levels to determine if effective teaching behavior increases over time. Siedentop and Eldar (1989) and Berliner (1986) indicate that expertise can be acquired at a quicker pace regardless of experience and may be developed more quickly if teachers are given strategies to affect student behavior at an early stage (Levin, Hammer, & Coffey 2009). Numerous studies indicate that the greatest gains in skills occur in novice teaching (Berliner, 1986; Chingos & Peterson, 2011; Marsh, 2007; Rivkin et al., 2005). In the present study, early career levels have been selected for further analysis in the hopes of better understanding the acquisition of teaching skills. To do so, a multiple case study design is used as well as two modes of data collection.

Participants

The Institutional Review Board of the University of Mississippi, which maintains and governs procedures dealing with human subjects for the purposes of research, reviewed this study for approval. The researcher provided both the research purpose and procedures for their

review. Additionally, the board requires that human subjects must voluntarily participate in the study. Upon board approval, permissions were gathered from the participants. The participants received a recruitment letter and a consent form. The recruitment letter and the consent form can be found in Appendices A and B. Other rules and regulations, as stated by the Institutional Review Board were followed to the fullest extent. Each participant of this study was coded using letters (A, B, C, etc.) to ensure anonymity.

Participants in this study were five novice teachers who received an undergraduate degree in music education from the University of Mississippi between 2016 and 2017. For the purpose of this study, a novice teacher is defined as a teacher with one to five years of experience. Participants were selected based on their proximity to the researcher, the number of years they had taught, longevity at the same school, and by recommendations of university faculty who were familiar with the participants' work.

Participant A is a white female in her twenties. She had an elementary focus in her preservice training and student taught at a rural upper elementary school (Grades 5-6). After graduation, she was hired to teach at a rural primary elementary school (Grades K-2) and has taught for two years. Participant B is a white female in her twenties. She had an elementary focus in her preservice training and student taught at a rural middle-grades elementary school (Grades 3-4). After graduation, she was hired to teach at a suburban elementary school (Grades K-5) and has taught for two years. Participant C is a white male in his twenties. He had a secondary band focus in his preservice training and student taught at a rural middle school (Grades 6-8). After graduation, he was hired as an assistant band director at a suburban middle school. He has been teaching there for two years. Participant D is a white male in his twenties. He had a secondary band focus in his preservice training and student taught at a suburban middle school (Grades 6-

8). He was hired as the head band director at the same middle school where he student taught with a focus on percussion. He has been teaching there for two years. Participants C and D teach at the same middle school. Participant E is a Mexican female in her mid-twenties. She had a choral focus in her preservice training and student taught at a suburban middle school (Grades 6-8). After graduation, she was hired to teach at a suburban middle school and has taught for three years.

Materials

As a part of their degree requirements, all participants took a capstone course prior to the semester of student teaching. One segment of the course was a three- to four-week peer teaching mini-unit that was video recorded. Instrumental preservice teachers taught a three-part recorder piece for soprano recorder and choral/elementary teachers taught an SAB or SATB selection. All pieces were selected by the participants. All teaching segments were approximately ten minutes in length. The videos were recorded using a Canon ZR500 camcorder and connected to a MacBook Pro via a digital to video firewire cable. The purpose of the mini-unit was for students to plan, execute, and self-evaluate their rehearsal segments. The instructor met individually with students to set teacher behavior goals and student performance goals and to adjust these goals during each subsequent rehearsal. The videos observed for the study were the first and last in the mini-unit. Teachers A, B, and E taught a choral selection and teachers C and D taught a recorder selection.

In student teaching, participants were required to video record themselves in a variety of teaching situations as a part of their student teaching portfolio. The videos were further observed

by their university supervisors and were self-evaluated to set teacher behavior goals and student performance goals. Since the videos were recorded independently, there were a variety of video camera positions in relation to the student teacher. Therefore, some videos did not show teachers' faces or were placed in a distant location making some teacher behaviors difficult to decipher (i.e. eye contact or facial expression). Additionally, lengths were not standardized and ranged from 10 minutes to an hour. Further, video quality varied, as participants did not utilize the same equipment. Participants' first video early in the semester and their last video late in the semester were used for analysis.

Participants' student teaching videos were obtained from their student teaching portfolio which were on file at the university. Participant A's first video was of a fifth-grade general music class and second video was of a sixth-grade general music class. Participant B's first video was a fourth-grade general music class while the other was a third-grade general music class. Participant C's first student teaching segment was of a seventh-grade low brass warm-up and second segment was of a seventh-grade wind ensemble rehearsal. Both of Participant D's videos were of two different seventh- and eighth-grade wind ensemble sectional rehearsals. Participant E's first video was of an advanced girls' choir and second video was an advanced girls' small ensemble.

I visited all participants twice during their second or third year of employment as a teacher and video recorded four entire class periods. I asked participants if I could record the grade level, class, or ensemble they felt the most comfortable teaching. Class periods ranged from 30 minutes to an hour. The recordings primarily focused on the teacher. I used a Zoom Q2n Handy Video Recorder to record all teaching episodes and then transferred the footage onto an external hard drive for ease of storage and retrieval. The first observation occurred in late August

of 2018. Participants A and B and were video recorded teaching kindergarten general music, Participant C was video recorded teaching a seventh-grade snare drum sectional, Participant D was video recorded teaching a seventh-grade marimba sectional, and Participant E was video recorded teaching a beginner sixth grade chorus. The second set of observations occurred between November of 2018 and January of 2019. I recorded the participants over a three-day period.

Due to nature of their jobs, I observed Participant A and B teaching three different grade levels during the same time span, as they were elementary general music specialists. I observed Participant A teaching general music to a kindergarten, first grade, and a self-contained class for mildly intellectually disabled children. I observed Participant B teaching general music to a fourth grade, third grade, and first grade class. Participants C, D, and E taught at secondary schools and I video recorded them teaching the same grade level over a three-day consecutive period. Participant C was video recorded teaching a sixth-grade auxiliary ensemble. Participant D was video recorded teaching the same ensemble as in August (seventh-grade marimba sectional). Participant E was video recorded teaching an advanced seventh- and eighth-grade women's ensemble.

Data Collection Tools

Analysis of the video recordings included a summative evaluation form called the Teacher Effectiveness Evaluation Form (TEEF). I created a modified version of Hamann and Baker's Survey of Teaching Effectiveness (1996), in that the STE form was meant to observe instrumental teachers and a wider range of behaviors than were part of the current study. I

created an initial survey sheet and piloted it using video recordings unrelated to the study. After some adjustments were made to the items and the number of responses available in the Likert scale weight, observers responded to six items using a five-point Likert scale with one marked as poor and five marked as excellent with anchor definitions. Those items were information and demonstrations, musical model, flow (related to pacing), instructional directives, feedback, and teaching style. Table 2 indicates the anchor definitions for the TEEF Items.

Table 2

Anchor Definitions for TEEF Items

TEEF Items	Excellent	Poor
Information and Demonstrations	Presented correct information; accurate demonstrations	Presented incorrect, contradictory, or misleading information; did not or could not accurately demonstrate (i.e., clapped or sang incorrect rhythms; did not demonstrate or provide information)
Musical Model	Expressive and accurate	Non-expressive, incorrect or inappropriate modeling; no modeling evidenced
Flow	Appropriate balance between teacher directives/explanations and student participation; one activity led to another without interruptions or breaks	Teacher talked too much; too much time spent going from one activity to another; long, disruptive breaks between and within activities
Instructional Directives	Specific directives identifying tasks to be accomplished	Non-specific directives with no specific tasks to accomplished
Feedback	Specific positive or negative feedback provided; utilized student ideas and comments when/where applicable	No feedback or non-specific feedback provided
Teaching Style	Secure, animated; captured students' attention and interest	Sluggish, lethargic, insecure; students were bored or disinterested

Additionally, the observers gave an overall rating of the lesson. These elements were selected for analysis because previous studies have found these aspects to be problematic for early career teachers (Anderson-Nichols, 1997; Bergee, 2005; Goolsby, 1996, 1997; 1999; Pike, 2014; Wagner and Strul, 1979). Observers completed the TEEF immediately after watching each video. The TEEF can be found in Appendix C.

Five observers were recruited to evaluate the participant's videos. Each observer was responsible for evaluating one participant's videos (N = 8) and one video from each level of another participant's videos for reliability purposes. The five observers were expert veteran teachers with experience in the participant's area. Therefore, the observers for Participants A, B, and E had elementary/choral teaching experience and observers for Participants C and D had band teaching experience. Observers were given training prior to watching videos and using the TEEF form. Inter-rater reliability (agreements / [agreements + disagreements]) was 83%.

Secondly, the researcher analyzed the teaching segments to identify rehearsal frames and their instructional targets. Rehearsal frames were identified using the outlines discussed in Duke's (1994; 1999/2000) articles. The researcher transcribed all videos (N = 40) and identified 316 rehearsal frames. The instructional target(s) for each rehearsal frame were identified and categorized. Instructional target categories were adapted for elementary settings from observational studies (Hendel, 1995; Orman, 2002; Price, 1990), as no studies have been performed using rehearsal frames as a tool for analysis in elementary general music classrooms. As an aside, while no general music studies have been performed using rehearsal frames as a tool for analysis, Taylor's study (2005) of upper-elementary Orff ensembles utilized rehearsal frames with success. Additionally, rehearsal frames are appropriate for analysis in elementary general music classes in that the refinement of musical skills and performance which is evidenced by

using rehearsal frames, can and should be demonstrated in all classrooms, regardless of levels. While the wording of "rehearsal frame" and "performance trial" may be questionable in an elementary general music setting, this verbiage will continue to be used in the study with the understanding that a rehearsal frame is simply a refined teaching sequence. However, certain classes or ensembles of any level may have more or less rehearsal frames depending on the objectives to be accomplished in the given time period.

Instructional target categories were adapted from Derby (2001) and Patterson's (2009) studies for choral settings. Instructional target categories were adapted from Cavitt's (2003) and Worthy and Thompson (2009) studies for instrumental settings. Instructional target categories and definitions can be found in Appendix B. Rehearsal frames with two or more performance trials were selected for further analysis (N = 60).

The software program SCRIBE 5 beta version for observation and assessment (Duke, 2019) was used during the analysis of rehearsal frames with two or more performance trials. SCRIBE was designed for use in observational research and is available as a downloadable software application at the Center for Music Learning at the University of Texas at Austin website. SCRIBE allows the user to input behavior categories to observe and presents results in a chronological record of event timings and summary tables that provide frequency and duration data collected during the observation interval. While watching the video recordings, the researcher clicked on-screen buttons that were labeled with specific behavioral categories. The program summarized frequency and duration data, including rate, proportion of time, and standard deviation for the behaviors and target categories. Figure 1 depicts the screen created for observations in SCRIBE.

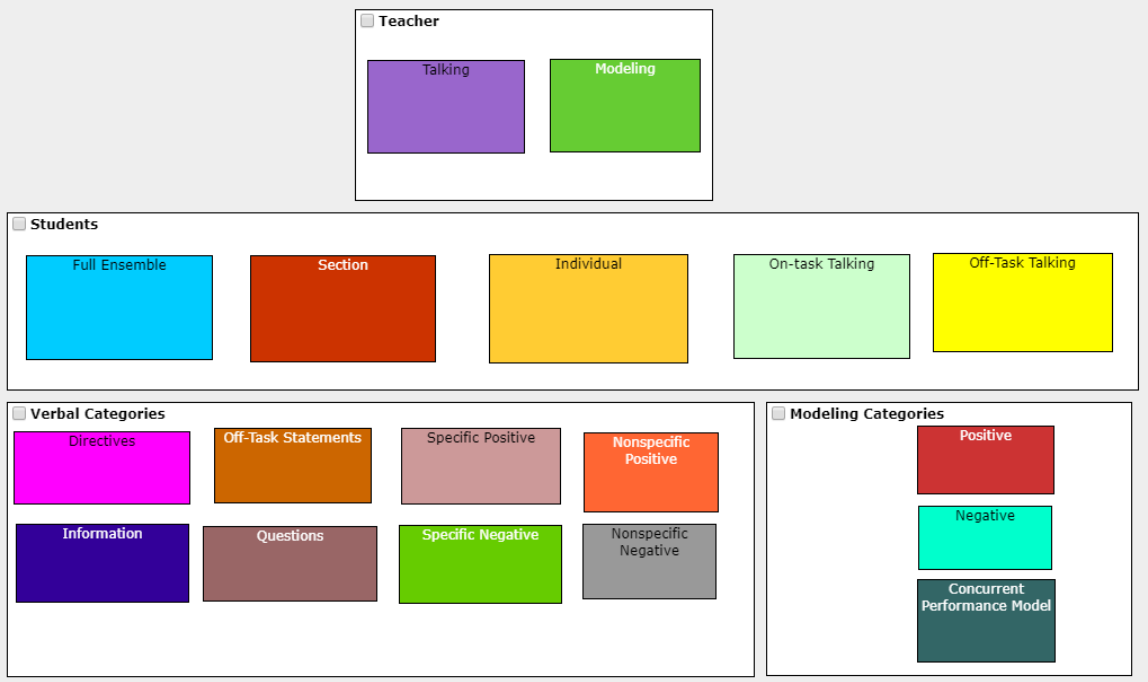

Figure 1. Depiction of SCRIBE observational categories for analysis of rehearsal frames with two or more performance trials.

Rehearsal frames with two or more student performance trials were analyzed by observing teacher and student behaviors specified in previous research (Beebe, 2007; Cavitt, 2003; Colprit, 2000; Derby, 2001; Singletary, 2016; Taylor, 2009). The primary teacher behaviors identified for duration measures were Teacher Talk and Modeling. Verbal behaviors were categorized using directive, information, questions, specific positive feedback, specific feedback, negative feedback, general positive feedback, general negative feedback, and off-task talking. Teacher modeling categories included positive modeling, negative modeling, and concurrent performance model. Student behaviors included whole group performance, small group performance, individual performance, on-task student talk, and off-task student talk. Table 3 provides operational definitions for each category and subcategory. A second researcher provided reliability on approximately 25% of the rehearsal frames selected for further analysis

by identifying and categorizing the instructional targets in rehearsal frames with two or more student performance trials. Inter-rater reliability (agreements / [agreements + disagreements]) was 80%.

Table 3

Operational Definitions for Student and Teacher Behaviors Recorded Using SCRIBE

Teacher Modeling Category (TM): Teacher performance in which the teacher plays, fingers, sings, claps, and/or counts the correct performance of a passage or the incorrect performance of a passage as a means of demonstration.
• Positive Modeling: Teacher behavior occurring between student performance trials in which the teacher sings, chants, plays an instrument, or mimics playing an instrument to demonstrate the correct performance of a passage or technique. • Negative Modeling: Teacher behavior occurring between student performance trials in which the teacher sings, chants, plays an instrument, or mimics playing an instrument to demonstrate the incorrect performance of a passage or technique. • Concurrent Performance Model: Teacher performance occurring simultaneously with student performance, including singing, chanting, patting, snapping, playing an instrument, or movement that mimics playing an instrument.
Teacher Talk Category (TT): Any verbalization by the teacher not related to modeling.
• Information Statements: Teacher verbalizations that convey information about the subject matter but does not direct the student to perform any specific action (e.g., "Lower pitches on the recorder are harder to play because you need slower air," "Singing on the ah vowel on higher pitches is difficult because you need more air to support the sound.") • Directives: "Do it" statements. Procedural or musical instruction given to student between and during performance trials. Procedural directives include where to begin in the music and who plays or sings. Musical directives also refer to aspects of musical expression (e.g., "Tenors, sing more quietly.") • Question: "On-task" question posed by the teacher related to the subject matter or rehearsal and to which the teacher expects a student response (e.g., "How many times do we play this motive?") • Specific Positive Feedback: Positive evaluations of preceding trials that describe one or more specific aspects of performance (e.g., "Your vowels were very pure in this phrase.") • Specific Negative Feedback: Negative evaluations of preceding trials that describe one or more specific aspects of performance (e.g., "You are rushing the phrase because you are not watching me.") • Nonspecific Positive Feedback: Positive evaluations of preceding trials that do not describe any specific aspects of performance (e.g., "Nice.") • Nonspecific Negative Feedback: Negative evaluations of preceding trials that do not describe any specific aspects of performance (e.g., "That was terrible.") • Off-task Statements: Any verbalization that does not pertain to the task at hand. This category may include comments made during interruptions or off-task comments initiated by the teacher.
Student Performance Category
Whole Group (WC): Student performance in which all students play or sing. Small Group (SG): Student performance in which a section of the group plays or sings. Individual (I): Student performance in which one student plays or sings.
Student Talk (ST): Any individual student verbalizations, including questions and responses to questions. This includes both on-task and off-task comments initiated by individual students.
• On-Task talking: Any individual student verbalizations, including questions and responses to questions. Includes students asking questions in relation to topic. • Off-Task talking: Any individual or group verbalization unrelated to lesson or question unrelated to topic.

CHAPTER IV

RESULTS

The purpose of this study was to observe the teaching of five novice music teachers in three early career teaching levels to determine if effective teaching behavior increases over time. I also wanted to find if any specific teacher behaviors associated with effectiveness are more prominent when teaching their peers in pre-service, another educator's students in student teaching, and their own students in novice teaching. Pacing, feedback, teacher talk, and modeling were analyzed to determine how different career levels measured in relation to instructional targets, teacher and student behavior in selected rehearsal frames, and ratings on summative evaluation forms.

Invitations to participate in this study were sent to five novice teachers who were selected based on their proximity to the researcher, the number of years they had taught, longevity at the same school, and by recommendations of university faculty who were familiar with the participants' work. Prior peer teaching and student teaching video recordings which were kept on file at the University of Mississippi's Department of Music were utilized for data collection. Additionally, the participants agreed to participate in the novice teacher data collection process and were observed and video recorded four times during the beginning of the school year and at

the midpoint of the same school year. The results of this study are presented for the individuals as well as the three career levels. Further discussion of the behavioral observation data, informal observation, and adjudicator comments on the TEEF will be presented in Chapter Five.

The results are organized based on the research questions posed in Chapter One:

1. What are the frequencies of instructional targets in rehearsal frames of the three early career levels?
2. What are the frequencies and durations of specified teacher and student behaviors observed in selected rehearsal frames?
3. What are the ratings of teaching effectiveness at each level of teaching and individually?

The total amount of video footage observed was 16 hours and 26 minutes. Peer teaching videos ranged from 9 minutes and 30 seconds to 10 minutes and 9 seconds with an average duration of 9 minutes and 59 seconds. Student teaching videos ranged from 10 minutes and 14 seconds to 47 minutes and 42 seconds with an average duration of 17 minutes and 39 seconds. Novice teaching videos ranged from 18 minutes and 9 seconds to 44 minutes and 45 seconds with an average duration of 35 minutes and 30 seconds. Table 4 reports the total duration of recorded footage for each participant.

IDENTIFICATION OF INSTRUCTIONAL TARGET CATEGORIES

I transcribed all video recordings (N = 40) for the purpose of identifying instructional targets and rehearsal frames. Instructional target categories were adapted for elementary settings

Table 4

Total Observation Times for All Participants (N=5, 16:26:34)

	Peer Teaching		Student Teaching		Novice Teaching				
Participant	PT 1	PT 2	ST 1	ST 2	NT 1	NT 2	NT 3	NT 4	Total
A	10:01	10:09	11:05	22:39	30:18	29:43	32:35	33:50	3:00:20
B	10:03	9:44	11:01	10:14	40:42	42:26	40:00	43:27	3:27:37
C	10:06	10:12	11:58	13:07	34:37	38:49	18:09	19:55	2:36:53
D	9:56	9:30	47:42	26:00	41:12	43:29	19:06	28:18	3:45:13
E	10:04	10:08	10:45	11:53	39:27	44:45	44:02	45:27	3:36:31
Total	50:10	49:43	1:32:31	1:23:53	3:6:16	3:19:12	2:33:52	2:50:57	16:26:34

Note: PT = Peer Teaching, ST = Student Teaching, NT = Novice Teaching

from observational studies (Hendel, 1995; Orman, 2002; Price, 1990), as no studies have been performed using rehearsal frames as a system of analysis in elementary general music classrooms. Instructional target categories were adapted from Derby's (2001) and Patterson's (2009) studies for choral settings. Instructional target categories were adapted from Cavitt's (2003) and Worthy and Thompson (2009) studies for instrumental settings. Upon analysis of video transcripts, 302 rehearsal frames were identified. Table 5 presents the distribution of rehearsal frames by instructional categories for all rehearsal frames (N = 302) across all video recordings and levels.

Regarding the individual participants, Participant A had 29 rehearsal frames with a single performance trial and 4 rehearsal frames with two or more performance trials. Participant B had 34 rehearsal frames with a single performance trial and 4 rehearsal frames with two or more

Table 5

Frequency of Instructional Targets for All Video Recordings and Career Levels in Rehearsal Frames (N = 302)

Category	Code	PT #1	PT #2	PT All	ST #1	ST #2	ST All	NT #1	NT #2	NT #3	NT #4	NT All	Total
Articulation	Art		1	1		2	2						3
Breathing/Airflow	BA		2	2									2
Blend/Balance	BB												
Diction/Pronunciation	DP												
Dynamics	Dyn		2	2	2	2	4						6
Embouchure	Emb				1		1						1
Ear Training/Sightreading	ES					4	4						4
Improvising	Imp												
Intonation/Tone	I/T	1		1							1	1	2
Movement	Mov								4			4	4
Multiple	Mul	1	3	4	2	8	10	9	6	4	1	20	34
Other	Oth		1	1				1	2	1	4	8	9
Pitch Accuracy	PA	18	11	29	6	9	15	15	18	5	5	43	87
Posture/Instrument Carriage	PI							1	2			3	3
Phrasing/Word Stress	PW	1	1	2					1			1	3
Presentation	Pre												
Rhythm Accuracy	RA	2	2	4	6	9	15	10	10	8	7	35	54
Reading/Notating	RN				5	3	8				1	1	9
Singing	Sin							2	4		1	7	7
Technical Facility	Tech				3		3	18	4		1	23	26
Tempo	Temp		1	1		4	4	1	1		1	3	8
Tone Quality/Vocal Placement	TV		2	2				6	4		2	12	14
Unidentified Target	UT	8	6	14	3		3	6	1	2		9	26

Note: PT = Peer Teaching, ST = Student Teaching, NT = Novice Teaching

performance trials. Participants A and B's peer teaching video recordings were teaching a choral selection and their student teaching and novice teaching were of elementary general music classes. Participant C had 55 rehearsal frames with a single performance trial and 12 rehearsal frames with two or more performance trials. Participant D had 72 rehearsal frames with a single performance trial and 28 with two or more performance trials. Participant E had 53 rehearsal frames with a single performance trial and 10 rehearsal frames with two or more performance trials.

In all three levels, pitch accuracy was identified the most (40% in peer teaching, 22% in student teaching, 25% in novice teaching) while blend/balance, diction/pronunciation, improvisation, and presentation were not observed among the three levels. Unidentified instructional targets ranked second in peer teaching (19%) but were observed less frequently in student teaching (4%) and novice teaching (5%). Rhythmic accuracy was identified second most frequently in student teaching (22%) and novice teaching (20%) but identified at 6% in peer teaching.

The highest percentages in peer teaching were pitch accuracy (40%), unidentified target (19%), and multiple targets (19%). Other instructional target categories were less prominent with rhythmic accuracy at 6% and breathing/airflow, dynamics, phrasing/word stress, and tone/vocal placement observed at 3%. Articulation, intonation, other, and tempo were observed in only 1% of rehearsal frames. Blend/balance, diction/pronunciation, embouchure, ear training/sightreading, improvisation, movement, posture/instrument carriage, presentation, reading/notating, singing, and technical facility were not observed.

In student teaching, pitch accuracy (22%), rhythmic accuracy (22%), multiple targets (14%), and reading/notating (12%) were identified the most frequently. Dynamics,

eartraining/sightreading, and tempo were observed in 6% of rehearsal frames, while technical facility (4%), unidentified target (3%), articulation (3%), and embouchure (1%) were observed least frequently. Breathing/airflow, blend/balance, diction/pronunciation, improvisation, intonation, movement, other, posture/instrument carriage, presentation, singing, tone quality/vocal placement were not observed.

In novice teaching, pitch accuracy (25%), rhythmic accuracy (20%), technical facility (14%), and multiple targets (12%) were identified the most frequently. Tone quality/vocal placement (7%), unidentified (5%), and other (5%) were observed at a medium to low frequencies. Singing (3%), movement (2%), posture/instrument carriage (2%), tempo (2%), intonation/tone quality (1%), phrasing/word stress (1%), and reading/notating (1%) were observed with the least amount of frequency. Articulation, breathing/airflow, blend/balance, diction/pronunciation, dynamics, embouchure, ear training/sightreading, improvisation, and presentation were not observed. Table 6 ranks the frequency of instructional targets over all three teaching levels.

Of the 302 rehearsal frames that were identified, 244 of them concluded after a single performance trial. Table 7 indicates the distribution of rehearsal frames by instructional categories for all single-performance trial rehearsal frames. In relation to career levels, certain categories were observed more frequently than others. Pitch accuracy was observed the most frequently across all levels (44% in peer teaching, 24% in student teaching, and 27% in novice teaching). Unidentified targets were observed the second most frequently in peer teaching only (27%) and observed less frequently in student teaching (6%) and novice teaching (7%).

Table 6

Ranking of Frequency of Instructional Targets for All Rehearsal Frames (N = 300) Across All Levels

Target Category	Peer Teaching Total	Target Category	Student Teaching Total	Target Category	Novice Teaching Total
Pitch Accuracy	29	Pitch Accuracy	15	Pitch Accuracy	43
Unidentified	14	Rhythmic Accuracy	15	Rhythmic Accuracy	34
Multiple	14	Multiple	10	Technical Facility	23
Rhythmic Accuracy	4	Reading/Notating	8	Multiple	20
Breathing/Airflow	2	Dynamics	4	Tone Quality/Vocal Placement	12
Dynamics	2	Ear Training/Sightreading	4	Unidentified	9
Phrasing/Word Stress	2	Tempo	4	Other	8
Tone/Vocal Placement	2	Technical Facility	3	Singing	6
Articulation	1	Unidentified	3	Movement	4
Intonation	1	Articulation	2	Posture/Instrument Carriage	3
Other	1	Embouchure	1	Tempo	3
Tempo	1	Breathing/Airflow	0	Intonation/Tone Quality	1
Blend/Balance	0	Blend/Balance	0	Phrasing/Word Stress	1
Diction/Pronunciation	0	Diction/Pronunciation	0	Reading/Notating	1
Embouchure	0	Improvisation	0	Articulation	0
Ear Training/Sightreading	0	Intonation	0	Breathing/Airflow	0
Improvisation	0	Movement	0	Blend/Balance	0
Movement	0	Other	0	Diction/Pronunciation	0
Posture/Instrument Carriage	0	Posture/Instrument Carriage	0	Dynamics	0
Presentation	0	Phrasing/Word Stress	0	Embouchure	0
Reading/Notating	0	Presentation	0	Ear Training/Sightreading	0
Singing	0	Singing	0	Improvisation	0
Technical Facility	0	Tone Quality/Vocal Placement	0	Presentation	0

Table 7

Frequency of Instructional Targets for Rehearsal Frames with Single Performance Trials (N = 244)

Category	Code	PT #1	PT #2	PT All	ST #1	ST #2	ST All	NT #1	NT #2	NT #3	NT #4	NT All	Total
Articulation	Art		1	1		2	2						3
Breathing/Airflow	BA		1	1									1
Blend/Balance	BB												
Diction/Pronunciation	DP												
Dynamics	Dyn		2	2	2	1	3						5
Embouchure	Emb				1		1						1
Ear Training/Sightreading	ES					4	4						4
Improvising	Imp												
Intonation/Tone	I/T	1		1							1	1	2
Movement	Mov								4			4	4
Multiple	Mul		1	1		6	6	1	3	1		5	12
Other	Oth		1	1				1	2	1	3	7	8
Pitch Accuracy	PA	16	7	23	5	8	13	14	15	4	4	37	73
Posture/Instrument Carriage	PI							1	2			3	3
Phrasing/Word Stress	PW	1	1	2					1			1	3
Presentation	Pre												
Rhythm Accuracy	RA	2	2	4	3	6	9	7	9	7	7	30	43
Reading/Notating	RN				4	3	7				1	1	8
Singing	Sin							2	3		1	6	6
Technical Facility	Tech				3		3	15	4			19	22
Tempo	Temp		1	1		4	4	1	1		1	3	8
Tone Quality/Vocal Placement	TV		2	2				4	3		2	9	11
Unidentified Target	UT	8	6	14	3		3	6	1	2	1	10	27

Blend/balance, diction/pronunciation, improvisation, and presentation were not observed by any of the three levels.

The most frequently observed targets in peer teaching single performance trials were pitch accuracy (44%) and unidentified target (27%) with a steep decline in percentages and frequencies after this. Rhythmic accuracy (7%), dynamics (4%), phrasing/word stress (4%), and tone/vocal placement (4%) were observed with the second highest level of frequency. Articulation, breathing/airflow, intonation/tone, multiple targets, other, and tempo were observed at 2% respectively. Blend/balance, diction/pronunciation, embouchure, ear training/sightreading, improvisation, movement, posture/instrument carriage, presentation, reading/notating, singing, and technical facility were not observed.

The most frequently observed targets in student teaching single performance trials were pitch accuracy (24%), rhythmic accuracy (17%), reading/notating (13%), and multiple targets (11%). Ear training/sightreading (7%), tempo (7%), dynamics (6%), and unidentified target (6%) were the second most frequently observed, while dynamics (5%) and articulation (4%) were observed the least frequently. Breathing/airflow, blend/balance, diction/pronunciation, movement, other, posture/instrument carriage, phrasing/word stress, presentation, singing, and tone/vocal placement instructional targets were not observed.

The highest observed targets in novice teaching single performance trial rehearsal frames were pitch accuracy (27%), rhythmic accuracy (22%), and technical facility (14%). Observed percentages declined sharply after these three instructional targets, with the next highest categories being unidentified targets and tone quality/vocal placement at 7%. The following instructional target categories were identified in small percentages of single performance trial rehearsal frames: other and singing (5%), multiple targets (4%), movement (3%), tempo and

Table 8

Ranking of Frequency of Instructional Targets for Rehearsal Frames with Single Performance Trials (N = 244) Across All Levels

Target Category	Peer Teaching Total	Target Category	Student Teaching Total	Target Category	Novice Teaching Total
Pitch Accuracy	23	Pitch Accuracy	13	Pitch Accuracy	37
Unidentified	14	Rhythmic Accuracy	9	Rhythmic Accuracy	30
Rhythmic Accuracy	4	RN	7	Technical Facility	19
Dynamics	2	Multiple	6	Unidentified	10
Phrasing/Word Stress	2	Ear Training/Sightreading	4	Tone/Vocal Placement	9
Tone/Vocal Placement	2	Tempo	4	Other	7
Articulation	1	Dynamics	3	Singing	6
Breathing/Airflow	1	Technical Facility	3	Multiple	5
Intonation/Tone	1	Unidentified	3	Movement	4
Multiple	1	Articulation	2	Posture/Instrument Carriage	3
Other	1	Embouchure	1	Tempo	3
Tempo	1	Breathing/Airflow	0	Intonation/Tone	1
Blend/Balance	0	Blend/Balance	0	Phrasing/Word Stress	1
Diction/Pronunciation	0	Diction/Pronunciation	0	Reading/Notating	1
Embouchure	0	Improvisation	0	Articulation	0
Ear Training/Sightreading	0	Intonation/Tone	0	Breathing/Airflow	0
Improvisation	0	Movement	0	Blend/Balance	0
Movement	0	Other	0	Diction/Pronunciation	0
Posture/Instrument Carriage	0	Posture/Instrument Carriage	0	Dynamics	0
Presentation	0	Phrasing/Word Stress	0	Embouchure	0
Reading/Notating	0	Presentation	0	Ear Training/Sightreading	0
Singing	0	Singing	0	Improvisation	0
Technical Facility	0	Tone/Vocal Placement	0	Presentation	0

Table 9

Frequency of Instructional Targets for Rehearsal Frames with Two or More Performance Trials by Career Levels in Rehearsal Frames (N = 58)

Category	Code	PT #1	PT #2	PT All	ST #1	ST #2	ST All	NT #1	NT #2	NT #3	NT #4	NT All	Total
Articulation	Art												
Breathing/Airflow	BA		1	1									1
Blend/Balance	BB												
Diction/Pronunciation	DP												
Dynamics	Dyn					1	1						1
Embouchure	Emb												
Ear Training/Sightreading	ES												
Improvising	Imp												
Intonation/Tone	I/T												
Movement	Mov												
Multiple	Mul	1	2	3	2	2	4	8	3	3	1	15	22
Other	Oth												
Pitch Accuracy	PA	2	4	6	1	1	2	1	3	1	1	6	14
Posture/Instrument Carriage	PI							2				2	2
Phrasing/Word Stress	PW												
Presentation	Pre												
Rhythm Accuracy	RA				1	5	6	3		1	1	5	11
Reading/Notating	RN				1		1						1
Singing	Sin												
Technical Facility	Tech							1			1	2	2
Tempo	Temp												
Tone Quality/Vocal Placement	TV							2	1	1		4	4
Unidentified Target	UT												

posture/instrument carriage (2%), and intonation/tone, phrasing/word stress, and reading/notating (1%). Instructional targets that were not observed in novice teaching included articulation, breathing/airflow, blend/balance, diction/pronunciation, dynamics, embouchure, ear training/sightreading, improvisation, and presentation. Table 8 ranks the frequency of instructional targets over the three levels of experience for single performance trials.

Fifty-eight rehearsal frames were identified that contained two or more performance trials and were selected for subsequent analysis of specified teacher and student behaviors. Table 9 indicates the distribution of selected rehearsal frames by instructional categories. Whereas pitch accuracy was the most observed instructional target in single-performance trials, rehearsal frames with two or more student performance trials occurred less frequently (three in peer teaching, five in student teaching, and six in novice teaching) and a majority of the 15 target categories were not addressed in these rehearsal frames.

The highest observed targets in peer teaching rehearsal frames with two or more performance trials were pitch accuracy (60%) followed by multiple targets (30%) with breathing/airflow (10%). The most frequently observed targets in rehearsal frames with two or more performance trials in student teaching were rhythmic accuracy (43%). Multiple targets (29%) and pitch accuracy (14%) were the second most frequently observed. Dynamics and reading/notating were observed the least (7%). The most frequently observed targets in novice teaching rehearsal frames with two or more performance trials were multiple targets (44%), followed by pitch accuracy (17%), rhythmic accuracy (15%), tone quality/vocal placement (12%), posture/instrument carriage and technical facility were identified less frequently at 6%. Table 10 ranks the frequency of instructional targets of rehearsal frames with two or more performance trials over the three levels of experience.

Table 10

Ranking of Frequency of Instructional Targets for Rehearsal Frames with Two or More Performance Trials (N = 58) Across All Levels

Target Category	Peer Teaching Total	Target Category	Student Teaching Total	Target Category	Novice Teaching Total
Pitch Accuracy	6	Rhythmic Accuracy	6	Multiple	15
Multiple	3	Multiple	4	Pitch Accuracy	6
Breathing/Airflow	1	Pitch Accuracy	2	Rhythmic Accuracy	5
Articulation	0	Dynamics	1	Tone/Vocal Placement	4
Blend/Balance	0	Reading/Notating	1	Posture/Instrument Carriage	2
Diction/Phrasing	0	Articulation	0	Technical Facility	2
Dynamics	0	Blend/Balance	0	Articulation	0
Embouchure	0	Diction/Phrasing	0	Blend/Balance	0
Ear Training/Sightreading	0	Dynamics	0	Diction/Phrasing	0
Improvisation	0	Embouchure	0	Dynamics	0
Intonation/Tone Quality	0	Ear Training/Sightreading	0	Embouchure	0
Movement	0	Improvisation	0	Ear Training/Sightreading	0
Other	0	Intonation/Tone Quality	0	Improvisation	0
Posture/Instrument Carriage	0	Movement	0	Intonation/Tone Quality	0
Phrasing/Word Stress	0	Other	0	Movement	0
Presentation	0	Posture/Instrument Carriage	0	Other	0
Rhythmic Accuracy	0	Phrasing/Word Stress	0	Phrasing/Word Stress	0
Reading/Notation	0	Presentation	0	Presentation	0
Singing	0	Singing	0	Reading/Notating	0
Technical Facility	0	Technical Facility	0	Singing	0
Tempo	0	Tempo	0	Tempo	0
Tone/Vocal Placement	0	Tone/Vocal Placement	0	Unidentified	0
Unidentified	0	Unidentified	0		

SELECTED REHEARSAL FRAMES ANALYSIS

Out of the 302 rehearsal frames, rehearsal frames with two or more performance trials (N = 58) were selected for further analysis. Participant A's 4 selected rehearsal frames' duration ranged from 25 seconds to 1 minute and 44 seconds. Their average duration was 51 seconds. Participant B's 4 selected rehearsal frames' duration ranged from 17 seconds to 1 minute and 52 seconds. The average duration was 1 minute and 16 seconds. Participant C's 12 selected rehearsal frames' duration ranged from 46 seconds to 6 minutes and 16 seconds. The average duration was 2 minutes and 32 seconds. Participant D's 28 selected rehearsal frames' duration was 24 seconds to 9 minutes and 29 seconds. The average duration was 1 minute and 49 seconds. Participant E's 10 selected rehearsal frame duration was 53 seconds to 2 minutes and 18 seconds. The average duration was 1 minute and 30 seconds. The average duration for all participants was 1 minute and 35 seconds.

Calculating and translating rate per minute can provide a more accurate portrayal of rehearsal frames and categories, in that duration and frequency count can be misleading. Rate per minute was calculated using the formula of duration divided by time. Translating rate per minute was calculated with the formula of 1 divided by rate per minute (RPM). When expressed in RPM, Participant A's RPM for all rehearsal frames across all levels was 0.26, roughly translating to one rehearsal frame every 3 minutes and 50 seconds. The rate for rehearsal frames with multiple performance trials was 0.02, approximately one rehearsal frame every 50 minutes. Participant B's RPM for all rehearsal frames was also 0.26 with rehearsal frames with multiple performance trials at 0.02 as well. Participant C's RPM for all rehearsal frames was 0.4 indicating a rehearsal frame every 2 minutes and 30 seconds. The rate for rehearsal frames with

multiple performance trials was 0.08, indicating a rehearsal frame every 12 minutes and 30 seconds. Participant D's RPM for all rehearsal frames was 0.58 indicating a rehearsal frame every 1 minute and 43 seconds. The rate for rehearsal frames with multiple performance trials was 0.12, indicating a rehearsal frame every 8 minutes and 20 seconds. Participant E's RPM for all rehearsal frames was 0.29 indicating a rehearsal frame every 3 minutes and 27 seconds. The rate for rehearsal frames with multiple performance trials was 0.05, indicating a rehearsal frame every 20 minutes.

In relation to selected rehearsal frames and career levels, 10 rehearsal frames were identified in peer teaching from three participants (Participants A, D, and E), 14 from student teaching from four participants (Participants A, B, C, and D), and 34 in novice teaching from all five participants. The average length of a rehearsal frame in novice teaching was 34 seconds. The average length of a rehearsal frame in student teaching was 2 minutes and 37 seconds while the average length of a rehearsal frame in novice teaching was 1 minute and 49 seconds.

In relation to RPM within levels, the RPM for all rehearsal frames in peer teaching was 0.63, indicating a rehearsal frame every 1 minute and 35 seconds. The rate for selected rehearsal frames was 0.17 indicating a rehearsal frame every 5 minutes and 53 seconds. In student teaching, the RPM for all rehearsal frames was 0.39, indicating a rehearsal frame every 2 minutes and 34 seconds. The rate for selected rehearsal frames was 0.09, indicating a rehearsal frame every 11 minutes and seven seconds. In novice teaching, the RPM for all rehearsal frames was 0.23, indicating a rehearsal frame every 4 minutes and 21 seconds. The rate for selected rehearsal frames was 0.05 indicating a rehearsal frame every 20 minutes.

The selected rehearsal frames were analyzed by measuring teacher and student behaviors observed in previous research (Beebe, 2007; Cavitt, 2003; Colprit, 2000; Derby, 2001;

Singletary, 2016; Taylor, 2009). The primary teacher behaviors observed were teacher talk and modeling which were measured in frequency and duration. Verbal behaviors were categorized using directive, information, questions, specific positive feedback, specific negative feedback, general positive feedback, general negative feedback, and off-task talking. Teacher modeling categories included positive modeling, negative modeling, and concurrent performance model. Student behaviors were placed into discrete categories that included whole group performance, small group performance, individual performance, on-task student talk, and off-task student talk.

The combined duration of selected rehearsal frames in peer teaching was 7 minutes and 38 seconds. There were 55 occurrences of teacher talk totaling 3 minutes and 48 seconds and accounted for 49.27% with a rate of 7.21 indicating an occurrence every 9 seconds. There were 17 occurrences of teacher modeling totaling 59 seconds and accounted for 12.84% of the time with a rate of 2.23 indicating an occurrence every 27 seconds. There were 30 occurrences of student behavior totaling 2 minutes and 51 seconds and accounted for 36.17% with a rate of 3.93 indicating an occurrence every 15 seconds. Two student performance behaviors were observed—whole group and small group. Individual performance and on- or off-task student talking were not observed. There were 11 occurrences of whole group which accounted for 21.23% of the combined duration with a rate of 1.44, indicating an occurrence every 42 seconds. There were 19 occurrences of small group performance which accounted for 14.94% with a rate of 2.49 indicating an occurrence every 24 seconds. Table 11 indicates the frequency count, RPM, duration, percentages, mean duration, and standard deviation of teacher behaviors and student behavior in peer teaching.

The most frequently observed categories of teacher verbal behaviors and teacher modeling behaviors in peer teaching were directives, positive modeling, and nonspecific positive

Table 11

Analysis of Selected Rehearsal Frames in Peer Teaching in relation to Teacher Behavior and Student Behavior

	Frequency	Rate per Minute	Duration	Percentage	Mean Duration	Average Standard Deviation
Teacher Behaviors						
Teacher Talk	55	7.21	03:48	49.27	00:04	2.20
Teacher Modeling	17	2.23	00:59	12.84	00:03.5	1.12
Student Behaviors						
Whole Group	11	1.44	01: 37	21.23	00:06	1.57
Small Group	19	2.49	01:14	14.94	00:04	0.21
Individual	0	0	0	0	0	0
On-Task Talking	0	0	0	0	0	0
Off-Task Talking	0	0	0	0	0	0

feedback. There were 44 occurrences of directives that accounted for 47% of all behaviors with a rate of 5.82. There were 14 occurrences of positive modeling that accounted for 15% of all behaviors with a rate of 1.93. There were 9 occurrences of nonspecific positive feedback and accounted for 10% of all behaviors with a rate of 1.25.

The second most frequently observed categories of teacher verbal behaviors and teacher modeling behaviors in peer teaching were information, questions, and concurrent performance model. There were six information statements and accounted for 7% of all behaviors with a rate of 0.65. There were 5 questions and accounted for 6% of all behaviors with a rate of 0.8. There were 5 occurrences of concurrent performance model and accounted for 5% of all behaviors with a rate of 0.53.

The least frequently observed categories of teacher verbal behaviors and teacher modeling behaviors in peer teaching were specific positive feedback, specific negative feedback, off-task statements, and nonspecific negative feedback. There were four specific positive feedback statements and accounted for 4% of all behaviors with a rate of 0.83. There were four specific negative feedback statements and accounted for 4% of all behaviors with a rate of 0.73.

There was one off-task statement and accounted for 1% of all behaviors with a rate of 0.06. There was one non-specific negative feedback statement and accounted for 1% statement and accounted for 1% of statement and accounted for 1% of all behaviors with a rate of 0.2. Table 12 indicates the frequency counts and RPM of verbal categories observed during peer teaching.

Table 12

Frequency Count and Rate per Minute of Observed Teacher Verbal Categories in Peer Teaching

Verbal Categories	Frequency Count	Rate per Minute
Directives	44	5.82
Information	6	0.65
Questions	5	0.80
Off-Task Statements	1	0.06
Specific Positive Feedback	4	0.83
Specific Negative Feedback	4	0.73
Nonspecific Positive Feedback	9	1.25
Nonspecific Negative Feedback	1	0.20
Positive Modeling	14	1.93
Negative Modeling	0	0
Concurrent Performance Model	5	0.53

The combined duration of selected rehearsal frames in student teaching was 36 minutes and 46 seconds. Selected rehearsal frames included 179 teacher talk episodes which accounted for 51.86% of the time, 47 episodes of teacher modeling for 10.96% of the time, and 182

episodes of student behaviors accounted for 32.81% of the time. Teacher talk RPM was 4.87, indicating an episode every 12 seconds. Teacher modeling RPM was 1.00, indicating an episode every minute. Student behavior RPM was 4.94, indicating an episode every 12 seconds. All student behavior categories were observed. There were 53 whole group episodes accounting for 21.57% of the time with a rate of 1.44. There were 37 small group episodes accounting for 11.07% of the time with a rate of 1.00. There were four individual performance episodes accounting for 0.01% of the time with a rate of 0.11. There were 27 occurrences of on-task student talking accounting for 0.03% of the time with a rate of 0.73. There were 61 occurrences of off-task student talking accounting for 0.13% of the time with a rate of 1.66. Table 13 indicates the frequency count, RPM, duration, percentages, mean duration, and standard deviation of teacher behaviors and student behavior in student teaching.

Table 13

Analysis of Selected Rehearsal Frames in Student Teaching in relation to Teacher Behavior and Student Behavior

	Frequency	Rate per Minute	Duration	Percentage	Mean Duration	Average Standard Deviation
Teacher Behaviors						
Teacher Talk	179	4.87	19:04	51.86	00:07	5.00
Teacher Modeling	47	1.00	04:02	10.96	00:06	1.49
Student Behaviors						
Whole Group	53	1.44	07:56	21.57	00:09	2.66
Small Group	37	1.00	04:04	11.07	00:07	0.86
Individual	4	0.11	00:31	0.01	00:08	0.26
On-Task Talking	27	0.73	01:00	0.03	0:07	0.29
Off-Task Talking	61	1.66	04:44	0.13	0:13	0.94

The most frequently observed categories of teacher verbal behaviors and teacher modeling behaviors in student teaching were directives, concurrent performance model, and off-task statements. There were 168 directives that accounted for 46% of all behaviors with a rate of 5.11. There were 33 concurrent performance model observances that accounted for 9% of all behaviors with a rate of 1.17. There were 33 off-task statements that accounted for 9% of all behaviors with a rate of 1.17.

The second most frequently observed categories of teacher verbal behaviors and teacher modeling behaviors were nonspecific positive feedback, questions, specific negative feedback, and positive modeling. There were 30 non-specific positive feedback statements that accounted for 8% of all behaviors with a rate of 1.15. There were 27 questions that accounted for 7% of all behaviors with a rate of 1.04. There were 23 specific negative feedback statements that accounted for 6% of all behaviors with a rate of 0.58. There were 21 positive modeling episodes that accounted for 6% of all behaviors with a rate of 0.47. The least frequently observed categories of teacher verbal behaviors and teacher modeling behaviors were nonspecific negative feedback, information statements, and specific positive feedback. Negative modeling was not observed. There were 12 nonspecific negative feedback statements that accounted for 3% of all behaviors with a rate of 0.17. There were 10 information statements that accounted for 3% of all behaviors with a rate of 0.26. There were nine specific positive feedback statements that accounted for 3% of all behaviors with a rate of 0.17. Table 14 indicates the frequency count and RPM of verbal categories observed during student teaching.

The combined duration of selected rehearsal frames in novice teaching was 60 minutes and 31 seconds. In novice teaching, 416 episodes were identified as teacher talk which accounted for 46.18% of the time, 209 episodes identified as teacher modeling which accounted for 24.44%

Table 14

Frequency Count and Rate per Minute of Observed Teacher Verbal Categories in Student Teaching

Verbal Categories	Frequency Count	Rate per Minute
Directives	168	5.11
Information	10	0.26
Questions	27	1.04
Off-Task Statements	33	0.36
Specific Positive Feedback	9	0.17
Specific Negative Feedback	23	0.58
Nonspecific Positive Feedback	30	1.15
Nonspecific Negative Feedback	12	0.17
Positive Modeling	21	0.47
Negative Modeling	0	0
Concurrent Performance Model	33	1.17

of the time, and 312 episodes identified as student behaviors which accounted for 42.9% of the time. Teacher talk RPM was 6.87, indicating an occurrence every nine seconds. Teacher modeling RPM was 3.45, indicating an occurrence every 18 seconds. Student behavior RPM was 5.15, indicating an occurrence every 12 seconds. All student behavior categories were observed. There were 121 whole group episodes accounting for 23.58% of the time with rate of 2.00. There were 28 small group performances accounting for 6.82% of the time with a rate of 0.46. There were 80 individual performance occurrences accounting for 7.78% of the time with a rate of was 1.32. There were 77 occurrences of on-task student talking accounting for 4.48% of the time with a rate of 1.27%. There were six episodes of off-task student talking accounting for s0.24% of the time with a rate of 0.1. Table 15 indicates the frequency count, RPM, duration, percentages, mean duration, and standard deviation of Teacher Behaviors and Student Behaviors in student teaching.

Table 15

Analysis of Selected Rehearsal Frames in Novice Teaching in relation to Teacher Behavior and Student Behavior

	Frequency	Rate per Minute	Duration	Percentage	Mean Duration	Average Standard Deviation
Teacher Behaviors						
Teacher Talk	416	6.87	27:57	46.18	00:04	3.29
Teacher Modeling	209	3.45	14:48	24.44	00:04	1.56
Student Behaviors						
Whole Group	121	2.00	14:16	23.58	00:06	2.29
Small Group	28	0.46	03:31	6.82	00:02	0.48
Individual	80	1.32	04:43	7.78	00:01	0.71
On-Task Talking	77	1.27	02:43	4.48	00:01	0.48
Off-Task Talking	6	0.10	00:09	.24	00:00:16	0.02

The most frequently observed categories of teacher verbal behaviors and teacher modeling behaviors in student teaching were directives, concurrent performance model, and positive modeling. There were 308 directives that accounted for 38% of all behaviors with a rate of 5.87. There were 114 concurrent performance model occurrences that accounted for 14% of all behaviors with a rate of 1.67. There were 90 positive modeling episodes that accounted for 11% of all behaviors with a rate of 1.46.

The second most frequently observed categories of teacher verbal behaviors and teacher modeling behaviors were questions, nonspecific positive feedback, information, and specific negative feedback. There were 70 questions that accounted for 9% of all behaviors with a rate of 1.20. There were 61 nonspecific positive feedback statements that accounted for 7% of all behaviors with a rate of 1.15. There were 58 information statements that accounted for 7% of all

behaviors with a rate of 0.91. There were 43 specific negative feedback statements that accounted for 5% of all behaviors with a rate of 0.82. The least frequently observed categories of teacher verbal behaviors and teacher modeling behaviors were specific positive feedback, negative modeling, nonspecific negative feedback, and off-task statements. There were 26 specific positive feedback statements that accounted for 4% of all behaviors with a rate of 0.48. There were 21 negative modeling episodes that accounted for 3% of all behaviors with a rate of 0.38. There were 13 nonspecific negative feedback statements that accounted for 2% of all behaviors with a rate of 0.23. There were 12 off-task statements that accounted for 1% of all behaviors with a rate of 0.17. Table 16 indicates the frequency count and RPM of verbal categories observed during student teaching.

Table 16

Frequency Count and Rate per Minute of Observed Teacher Verbal Categories in Novice Teaching

Verbal Categories	Frequency Count	Rate per Minute
Directives	168	5.87
Information	10	0.90
Questions	27	1.19
Off-Task Statements	33	0.17
Specific Positive Feedback	9	0.48
Specific Negative Feedback	23	0.82
Nonspecific Positive Feedback	30	1.15
Nonspecific Negative Feedback	12	0.23
Positive Modeling	21	1.46
Negative Modeling	0	0.38
Concurrent Performance Model	33	1.67

A side-by-side comparison of the data collected from selected rehearsal frames per level is necessary. Teacher talk percentages remained in an area of approximately 50% with a smaller percentage in novice teaching. Teacher talk RPM in peer teaching and novice teaching were very

similar with a slower rate observed in student teaching. However, the average standard deviation for teacher talk was highest in student teaching (SD = 5). Teacher modeling increased almost by two times in novice teaching with a higher RPM. The lowest percentage observed was in student teaching with the lowest RPM. Student performance data indicated similar percentages in whole group with novice and student teaching with a slight increase in novice teaching with a slightly lower rate per minute. However, the average standard deviation was highest in student teaching (SD = 2.66). The percentage of episodes addressing small groups decreased and RPM increased with experience. Individuals were not addressed during peer teaching but were addressed at increasingly higher percentages during student teaching to novice teaching. On-task student talking did not occur in peer teaching but did occur in student teaching and novice teaching. The percentage was lower in student teaching than novice teaching but had a higher rate per minute and higher mean duration. Off-task student talking did not occur in peer teaching but did occur in student teaching and novice teaching. Small differences in percentages indicated an increase in novice teaching, but a decrease in RPM with a lower mean duration. Table 17 indicates a comparison of the three levels in relation to RPM, percentage, mean duration, and standard deviation.

Similarly, it is important to compare the three teaching levels in relation to ranking verbal teacher category frequency counts and comparing rate per minute. Directives were the highest observed category throughout all levels. Information statements were used the most in peer teaching and ranked fourth most utilized but ranked fifth in novice teaching. Information statements were one of the least observed categories in student teaching, ranking at ninth out of 11 categories. Questions ranked similarly among the levels and were fourth in peer teaching and student teaching and third in novice teaching. Off-task statements were observed the most in

Table 17

Comparison of Teacher Behaviors and Student Behaviors Across Experience Levels in Relation to Rate per Minute, Percentage, Mean Duration, and Average Standard Deviation

	Peer Teaching				**Student Teaching**				**Novice Teaching**			
	RPM	%	Mean Dur	Av. SD	RPM	%	Mean Dur	Av. SD	RPM	%	Mean Dur	Av. SD
Teacher Behaviors												
Teacher Talk	7.21	49.27	00:04	2.20	4.87	51.86	00:07	5.00	6.87	46.18	00:04	3.29
Teacher Modeling	2.23	12.84	00:03.5	1.12	1.00	10.96	00:06	1.49	3.45	24.44	00:04	1.56
	RPM	%	Mean Dur	Av. SD	RPM	%	Mean Dur	Av. SD	RPM	%	Mean Dur	Av. SD
Student Behaviors												
Whole Group	1.44	21.23	00:06	1.57	1.44	21.57	00:09	2.66	2.00	23.58	00:06	2.29
Small Group	2.49	14.94	00:04	0.21	1.00	11.07	00:07	0.86	0.46	6.82	00:02	0.48
Individual	0	0	0	0	0.11	0.01	00:08	0.26	1.32	7.78	00:01	0.71
On-Task Talking	0	0	0	0	0.73	0.03	0:07	0.29	1.27	4.48	00:01	0.48
Off-Task Talking	0	0	0	0	1.66	0.13	0:13	0.94	0.10	.24	00:00:16	0.02

student teaching, ranking third among the categories. However, off-task statements were ranked the lowest in peer teaching and novice teaching. Specific positive feedback was a lower observed category among all three levels, ranking at seventh in peer teaching and novice teaching and tenth in student teaching. Specific negative feedback ranked in the lower to middle range among levels, ranking at seventh in peer teaching, fifth student teaching, and six in novice teaching. Nonspecific positive feedback subtly decreased with experience, being ranked second in peer teaching, third in student teaching, and fourth in novice teaching. Nonspecific negative feedback

increased in student teaching was ranked eighth while being one of the lowest ranked categories in peer teaching and novice teaching. Positive modeling was one of the most frequently observed categories in peer teaching and novice teaching ranking at second and third, respectively. However, it was observed less frequently in student teaching and was ranked seventh. Negative modeling was never observed in peer teaching and student teaching and was ranked the third lowest category in novice teaching. Concurrent performance model increased in student teaching and novice teaching, ranking second in both levels. Concurrent performance model was ranked sixth in peer teaching. Table 18 ranks the teacher verbal categories among all levels.

Table 18

Comparison of Rankings of Teacher Verbal Categories Across Experience Levels

	Peer Teaching		Student Teaching		Novice Teaching
Directives	44	Directives	168	Directives	308
Positive Modeling	14	Concurrent Performance Model	33	Concurrent Performance Model	114
Nonspecific Positive	9	Off-Task Statements	33	Positive Modeling	90
Information	6	Nonspecific Positive	30	Questions	70
Questions	5	Questions	27	Nonspecific Positive	61
Concurrent Performance Model	5	Specific Negative	23	Information	58
Specific Positive	4	Positive Modeling	21	Specific Negative	43
Specific Negative	4	Nonspecific Negative	12	Specific Positive	26
Off-Task Statements	1	Information	10	Negative Modeling	21
Nonspecific Negative	1	Specific Positive	9	Nonspecific Negative	13
Negative Modeling	0	Negative Modeling	0	Off-Task Statements	12

TEACHING EFFECTIVENESS RATINGS

Five observers were selected to view and evaluate the five participant's videos. Each observer was responsible for evaluating one participant's videos (N = 8) and one video from each level of another participant's videos for reliability purposes. The five observers were expert veteran teachers with experience in their assigned participant's area of expertise. Analysis of the videos included a summative evaluation form called the Teacher Effectiveness Evaluation Form (TEEF). Selected items from Hamann and Baker's Survey of Teaching Effectiveness (1996) were modified for this study. Observers rated six items on a five-point Likert scale with 1 marked as poor and 5 marked as excellent with anchor definitions. The items were information and demonstrations, musical model, flow (related to pacing), instructional directives, feedback, and teaching style. Additionally, the observers gave an overall rating of the lesson. These items were selected for analysis because previous studies had found these elements problematic in early career teachers (Anderson-Nichols, 1997; Bergee, 2005; Goolsby, 1996, 1997; 1999; Pike, 2014; Wagner and Strul, 1979). Observers completed the TEEF forms immediately after watching each video. Additionally, observers were given the opportunity to comment on each video and provide further information. All experts chose to write comments for every video. The highest rating a participant could receive was a 35 and the lowest possible score was a 7.

TEEF form analysis resulted in a range of scores on Peer Teaching #1 for each element from two to five. The range of overall scores was 14 to 29. No specific trends could be found in item ratings. Overall ratings increased in Peer Teaching #2 ranging from 21-30. Most participants improved in information and demonstrations. Musical model was ranked high (three participants receiving a 4, one participant receiving a 5, one participant receiving a 2). Three

participants improved in feedback from the previous score in Peer Teaching #1. Teaching style also increased with four participants in Peer Teaching #2.

The overall scores ranged in Student Teaching #1 from 13 to 32. Musical model scores were low and decreased from peer teaching. Flow indicated no particular trends in that each participant had a different score. Instructional directives indicated no trends from peer teachings, but three participants were rated a 3 in instructional directives. The scores in Peer Teaching #2 ranged from were 22 to 33. Three participant's overall scores increased while one remained the same and one decreased from Student Teaching #1. Feedback scores decreased from Peer Teaching #2. No trends were found in teaching style and overall rating. Results from Student Teaching #2 ratings resulted similar scores in information and demonstrations as in Student Teaching #1. Scores increased in musical model for three participants in Student Teaching #2. No trends could be found in flow, instructional directives, feedback, teaching style, or overall rating.

The overall scores in Novice Teaching #1 ranged from 23 to 33. Information and directives were scored high, but no trends were found between Student Teaching #2 and Novice Teaching #1. Scores in flow decreased in all participants but increased in instructional directives for three participants. Feedback was given high marks for four participants in the scale of 4 or 5, but no trends were found between student teaching and novice teaching. Three participants' scores increased in teaching style. No trends were found for overall rating; however, three participants were given a 3. The range of overall scores in Novice Teaching #2 was 16 to 33. Information and demonstration scores remained the same for three participants and decreased in two. Musical model ratings increased in three participants. Flow scores stayed the same for three participants and decreased in two participants. Instructional directive ratings decreased in three

participants. Feedback ratings increased in three participants and decreased in two participants. No trends were found in teaching style ratings or overall rating. Overall scores in Novice Teaching #3 ranged from 16 to 33. Information and demonstrations ratings resulted in the same scores for four participants from Novice Teaching #2. Scores decreased in musical model for three participants. No trends were found in flow ratings. Instructional directive ratings were given the same score for three participants. Three participants had an increase in ratings of feedback, teaching style, and overall rating. Overall scores in Novice Teaching #4 ranged from 14 to 32. No trends were found between Novice Teaching #3 and #4 in information and demonstrations, musical model, flow, instructional directives, and teaching style. Feedback and overall rating scores remained the same for three participants, albeit having various scores. TEEF results by career level are available in Appendix D.

In relation to specific items on the TEEF, information and directives were given consistently high marks throughout all levels and were given scores of 4 or 5 in 31 out of the 40 video recorded lessons. Musical model scores were mixed with scores of 4 or 5 for 20 excerpts. The lowest scores for all participants in this area were found in Student Teaching #1. Similarly, scores were mixed in relation to flow with 21 lessons given a score of 4 or 5, most of which were given to one participant. Results were mixed in relation to instructional directives with 19 lessons receiving a score of 4 or 5. Scores for feedback were also mixed with 19 lessons receiving a score of 4 or 5, most of which were connected to one participant. However, the trend of the scores indicated an increase of scores with experience for three participants. Similarly, teaching style ratings resulted in19 lessons receiving a score of 4 or 5. The trend of the scores indicated an increase of higher scores in novice teaching with one participant showing high marks in all levels. Overall rating scores resulted in 17 videos receiving a score of 4 or 5.

Further, Participants' A and B ratings decreased with experience while Participant C and D increased with experience. Participant E received one high rating during novice teaching with no trend otherwise. TEEF results by items with highlighted highest ratings are available in Appendix E.

Inspecting TEEF results for individual participants across levels may also reveal important trends. Participant A's TEEF ratings indicated the highest scores in peer teaching and student teaching. Out of the 28 possible items to score in the four videos from the two levels, 26 of those items were given a 4 or 5. Conversely, the four lessons in novice teaching received a score of 4 on nine items and received no score of 5 out of 28 possible items. She also was marked consistently high in information and demonstrations six out of eight times and seven out of eight times for teaching style.

Participant B's TEEF ratings resulted in the highest scores in peer teaching only with 13 out of 14 items marked as 4 or 5. Out of 42 possible items to score in the six lessons in student teaching and novice teaching, 10 items were marked 4 or 5. She consistently was rated high through all levels on information and demonstrations. Further, she was consistently rated low for flow receiving mostly a 2 or 3 and receiving a 1 in a novice teaching video. Instructional directives and overall rating were also consistently low marked in student teaching and novice teaching.

Participant C's TEEF ratings indicated the highest scores in novice teaching. Out of 28 possible items to score in peer teaching and student teaching, five were rated as 4 and did not receive the highest score of 5. Conversely, all 28 possible items to observe in novice teaching were marked as 4 or 5. He was also given consistently high ratings on the information and demonstrations items.

Participant D's TEEF ratings were the highest in novice teaching as well; however, numerous specific items were scored high throughout all levels. All 28 possible items to observe in novice teaching were rated with a 4 or 5. Out of all 28 possible items to observe in peer teaching and student teaching, 14 items received a 4 or 5. His lowest scores were in Student Teaching #1, given a 13 out of 35 total score. Participant D was consistently rated high in information and demonstrations, musical model, flow, and instructional directives.

Participant E's TEEF ratings were the highest scores in Novice Teaching #2 only. She was consistently rated high in information and demonstrations throughout student teaching and novice teaching. Out of 56 possible items to observe throughout all videos of all levels of teaching, she received a score of 2 or 3 in 44 items. Her lowest scores were found in her final novice teaching video. TEEF results by individual with highlighted highest ratings are available in Appendix F.

CHAPTER V

DISCUSSION

The purpose of this study was to observe the teaching of five novice music teachers in three early career teaching levels to determine if effective teaching behavior increases over time. I also wanted to find if any specific teacher behaviors associated with effectiveness are more prominent when teaching their peers in pre-service, another educator's students in student teaching, and their own students in novice teaching.

Eight video recordings of teaching were collected for each of the five participants: two from peer teaching in an undergraduate teacher preparation course, two from their student teaching experience, and four from their first years of professional service.

Using the data collected, I attempted to answer the following research questions:

1. What are the frequencies of instructional target categories in rehearsal frames of the three early career levels?
2. What are the frequencies and durations of specified teacher and student behaviors observed in selected rehearsal frames?
3. What are the ratings of teaching effectiveness at each level of teaching and individually?

INSTRUCTIONAL TARGET CATEGORIES

I transcribed all video recordings (N = 40) for the purpose of identifying rehearsal frames and their instructional targets. Rehearsal frames were identified in 15 instructional target categories—articulation, breathing/airflow, blend/balance, diction/pronunciation, dynamics, embouchure, ear training/sightreading, improvising, intonation/tone, movement, multiple, other, pitch accuracy, posture/instrument carriage, phrasing/worse stress, presentation, rhythmic accuracy, reading/notating, singing, technical facility, tempo, tone quality/vocal placement, and unidentified target.

Instructional targets for single performance trial rehearsal frames (N = 244) were vastly different from those with two or more performance trials. In single performance trial rehearsal frames, pitch accuracy was the most frequently observed instructional target among all three levels. This finding is similar to Patterson's (2009) findings of rehearsal with beginning choirs as well as Worthy and Thompson's (2009) and Singletary's (2016) study of beginning band. Similarly, pitch accuracy was the second most selected target in Bond's (2015) study of student teacher's off and on-podium time. Waymire's study (2011) comparing high-performing and low-performing band directors indicated that low-performing directors focus on pitch accuracy more (12.9%) than do high-performing directors (1.92%) Similarly, Goolsby (1997) found that pitch accuracy was identified the most during student teaching versus novice and expert teachers. Pitch accuracy was not addressed in Taylor's observation of advanced elementary Orff ensembles and was observed infrequently in Beebe's (2007), Cavitt's (2003), Culp's (2018) and Worthy's (2003, 2006) studies of varying ensemble levels of instrumental directors. Pitch accuracy was

also not addressed in Whitaker's (2017) study of famous conductors of advanced symphonies. The participants in this study worked primarily with younger ensembles and/or those with limited experience. These findings seem to indicate that directors of all experience levels from novice to expert who work with beginning and intermediate groups focus on pitch accuracy. This may be due to the student's lack of experience with their instrument and the necessity to reinforce positive performances. However, it also seems that less effective directors and those who lack experience focus more on pitch accuracy than their expert counterparts. This may be because pitch errors are more easily identified than other types of errors. More studies comparing novice and expert teachers identifying instructional targets are needed.

Unidentified targets were the second most frequently identified instructional target in peer teaching only. Unidentified targets declined steeply in student teaching and novice teaching. Unidentified targets were rarely identified in studies of expert directors (Beebe, 2007; Cavitt, 2003; Culp, 2018; Derby, 2001; Patterson, 2002; Singletary, 2016; Taylor, 2008; Whitaker, 2017; Worthy 2003, 2007; Worthy & Thompson, 2009;). Colprit's (2000) study of expert Suzuki teachers identified 12% of targets as unidentified or unclear. However, Waymire's study (2011) found that less effective teachers pursued unidentified targets more than their effective counterparts. Similarly, novice teachers and student teachers in Goolsby's studies (1997, 1999) asked students to repeat small groups with no identified purpose more than their expert counterparts. These findings would seem to indicate that teachers who lack experience or become stagnant in their growth lack pedagogical clarity in being able to pinpoint problematic issues. Therefore, it is of primary importance that preservice and student teachers are given as

many opportunities as possible to discuss problematic issues and identify strategies to effectively address them.

In student teaching, more targets were identified with pitch accuracy, rhythmic accuracy, reading/notating, and multiple targets observed the most. Rhythmic accuracy, reading/notating, and multiple targets were also observed at high frequencies in numerous studies of experts and novices (Cavitt, 2003; Goolsby, 1997; Patterson, 2009

; Singletary, 2016; Waymire, 2011; Worthy, 2003, 2006; Worthy & Thompson, 2009). The pursuit of multiple targets simultaneously might indicate a higher level of processing; therefore, it is possible that in this study, the participants were evolving in critical listening skills. However, multiple targets were not the most frequent category observed in novice teaching, possibly indicating that the pathway of expertise without guidance may result in irregular trends during early career teachers.

In novice teaching, pitch accuracy, rhythmic accuracy, and technical facility were the most frequently observed targets. Technique was another target which experts focused on in conjunction with other higher order targets in multiple studies (Colprit, 2000; Taylor, 2009; Goolsby, 1997; 1999). Therefore, it would seem from the current study that early career teachers may be able to identify targets which require a more complex understanding of pedagogy, but only in isolated situations and terms.

Rehearsal frames with two or more performance trials (N = 58) revealed a smaller range of target types. All career levels exhibited multiple targets as one of the most frequently utilized targets (first in novice teaching and second in peer teaching and student teaching). Pitch accuracy was also a top selected target (first peer teaching, second in novice teaching, and third in student

teaching). Rhythmic accuracy was the top selected target in student teaching and third novice teaching. Novice teaching rehearsal frames included multiple targets related to technical facility. These findings are similar to other studies related to expert's use of instructional time (Derby, 2001; Patterson, 2009; Singletary, 2016; Worthy, 2003, 2006; Worthy & Thompson, 2009). Multiple instructional targets coupled with multiple performance trials might indicate a developing sense of persistence while addressing complex performance issues.

Blend/balance, diction/pronunciation, improvisation, and presentation were not identified in any of the levels in the current study. Studies which focused on expert directors found that these targets were utilized more than overt instructional targets such as pitch accuracy (Derby, 2001; Goolsby, 1997, 1999; Patterson 2009; Whitaker, 2017). Perhaps this is due to "information overload" in early career teachers, in that it is difficult for novices to process multiple pieces of information at once (Berliner, 1988) or that identifying more sophisticated issues might be challenging. Similarly, in peer teaching, the participant's main instructional targets were pitch accuracy and unidentified targets. After this, target frequencies rapidly declined. Perhaps encouragement and coaching through the mentoring process will assist early career teachers in critical listening.

ANALYSIS OF REHEARSAL FRAMES

Out of the 302 rehearsal frames identified, 58 rehearsal frames with two or more performance trials were selected for further analysis. The average time for all rehearsal frames in this study was 1 minute and 35 seconds. These rehearsal frame lengths were approximately the

same as Waymire's study (2011) of effective and non-effective band directors. However, most studies found average rehearsal frames longer in duration (Culp, 2018; Derby, 2001; Singletary, 2016; Taylor 2006; Waymire, 2011; Worthy, 2006; Worthy & Thompson, 2009), ranging from 2 minutes and 6 seconds (Worthy, 2006) to 5 minutes and 16 seconds (Derby, 2001). In a related vein, Goolsby (1999) found that teaching segments were shorter with experienced teachers, but more rehearsals were necessary for novice teachers to complete work on the same piece of music.

In the same realm, all rehearsal frame rates decreased with experience. However, only three participants led rehearsal frames with multiple performance trials in peer teaching and four participants did so in student teaching. By novice teaching, rehearsal frames with multiple performance trials were identified in the teaching of all participants. This finding might suggest that developing teachers may become comfortable with rigorous instruction with more hands-on experience.

Considering the basic structure of instruction and pacing, most teachers may use rigorous instruction times such as rehearsal frames with multiple performance trials less than those with single performance trials. This was certainly the case in the current study, in that all participants had a lower rate of rehearsal frames with two or more performance trials than ones with single performance trials. However, Participants C and D had a much higher RPM (.08 and .12) than Participants A, B, and E (.05) in rehearsal frames with multiple performance trials. Reasons for this are unknown; however, it is my speculation that Participants C and D may have had preservice training in their instrumental directing methods courses that focused on using sequential instruction which the Participants A, B, and E did not have in choral or elementary

methods courses. Many studies have been performed which indicate that training in enthusiasm and sequential teaching can alter effectiveness levels (Arnold, 1991, 1995; Benson, 1989; Collins 1978; Cassidy, 1990; Yarbrough, et al., 1991). However, more studies need to be performed to determine if training in rehearsal frame analysis can assist in increasing the number of instructional episodes that exist in the classroom.

Among all three experience levels, teacher talk remained around 50% (49.27% in peer teaching, 51.86% in student teaching, and 46.18% in novice teaching). Many studies report similar percentage levels with various teacher expertise levels (Beebe, 2007; Bonds, 2015; Colprit, 2000; Dorfman, 2010; Ihas, 2011; Orman, 2002; Worthy, 2003, 2006; Worthy & Thompson, 2009). However, this finding was not the same in Taylor (2009), Siebenhaler (1997), Derby (2001), or Culp's (2018) studies of varying teaching populations. Studies that compared levels of expertise did not corroborate the finding of the current study. Goolsby (1996) found that episodes of teacher talk decreased with experience when comparing student teachers, novice teachers, and experienced teachers. Waymire (2011) found that successful band directors used less teacher talk (38.99%) than unsuccessful band directors (51.42%). Singletary's (2016) comparison study of advanced and beginning middle school band found higher percentages of teacher talk than the current study at 66% for advanced and 63% for beginning ensembles. More studies need to explore the differences in teacher talk between levels of teaching expertise.

The lowest amount of teacher modeling occurred in student teaching (10.96%) followed by peer teaching (12.84%), and almost doubled in novice teaching (42.9%). The high percentage in novice teaching occurred through consistent use of concurrent performance model by many of the participants. Studies which analyze modeling in rehearsal frames have found a broad range of

usage from less than 1% to 34.14%. However, some studies report findings that are similar to the current study's percentages in peer teaching and novice teaching. Taylor (2009), Worthy (2006), and Worthy and Thompson (2009) found similar findings of modeling; around 10% in expert Orff ensemble and band directors of various levels. However, most studies reported much smaller percentages of modeling than the current study, ranging from less than 1% to 8.3%. Worthy (2003) found when comparing the same director's use of modeling with a high school and college ensemble, more modeling was used with the high school (7.32%) than with the college ensemble (5.23%). Goolsby's (1999) study which compared rehearsal strategies of novice and expert band directors of the same piece of music found that modeling was very similar between the two groups (6.3% and 6.4%). Derby (2001) and Dorfman (2010) both found modeling in a similar range of 6-7%. Culp (2018), Bonds (2015), Waymire (2011), and Singletary (2016) found minimal to low modeling percentages (less than 1% to 3%). The modeling percentage in novice teaching was similar to Colprit (2000), Siebenhaler (16%) (1997), and younger ensembles in Patterson's (2009) study. Goolsby's study (1997) which compared student teachers, novice teachers, and expert teachers found the opposite to be true than in the current study. Goolsby found that the lowest percentage of modeling was in novice teaching (11.2%) followed by student teaching (16%) and expert teaching (17.3%). Siebenhaler and Colprit's studies focused on interactions in private lessons. Perhaps private lessons have more opportunities for modeling and one on one interaction. However, that does not explain why modeling was at such a high percentage in Patterson's study of one of three middle and high school choral ensembles. Additional studies to understand use of modeling are necessary.

Student performance durations had a range of 10 percentage points between levels. The

lowest percentage was in student teaching (32.81%) followed by peer teaching (36.17%) and novice teaching (42.9%). The percentages of performances observed in novice teaching was similar to Colprit (2000), Siebenhaler (1997) and Waymire's (2011) studies. The highest percentage of student performances were in the whole group category across all levels followed by the small group category in peer teaching and student teaching. The second highest percentage in novice teaching was the individual category with the small group category approximately 1% below. Most studies reported similar findings with whole group performing the most followed by small group and individuals (Beebe, 2007; Goolsby, 1996; Singletary, 2016; Taylor, 2009; Worthy, 2003, 2006; Worthy & Thompson, 2009). However, some studies found that performance approximations and student talk ranked highly (Bonds, 2015; Culp, 2018; Derby, 2001). It would seem that most rehearsals focus on whole group instruction during intense rehearsal frame periods in that it may be the most effective strategy for the group. Off-task and on-task student talking were not observed in peer teaching but were observed in student teaching and novice teaching. Furthermore, the rate per minute was faster in on-task talking and off-task talking was slower in novice teaching. This indicates that with experience, novice teachers may have better control of the classroom.

The most frequently observed category of teacher verbal behaviors and teacher modeling behaviors across all levels was the directives category. Directive statements were consistently the most observed verbal category in many studies (Beebe, 2007; Bonds, 2015; Colprit, 2000; Derby, 2001; Waymire, 2011; Whitaker, 2017; Worthy, 2006; Worthy & Thompson 2009). This would seem to indicate that "do it" statements are the prominent teacher verbalization in music classrooms, if not all classrooms.

In peer teaching, positive modeling and nonspecific positive feedback were the second most frequently observed categories of teacher verbal behaviors and teacher modeling behaviors. This finding is similar to Goolsby's (1997, 1999) studies in which nonspecific positive feedback was observed more frequently in student teachers and novice teachers than expert teachers. In a similar realm, vague questions were asked more by inexperienced teachers and their expert counterparts in Goolsby's studies as well. Further, nonspecific negative feedback was one of the least observed behaviors in peer teaching and novice teaching in the current study. In student teaching, concurrent performance model and off-task statements were the second highest observed categories of teacher verbal behaviors and teacher modeling behaviors. In novice teaching, concurrent performance model and positive modeling were the highest observed categories of teacher verbal behaviors and teacher modeling behaviors.

Positive modeling was one of the most frequently observed categories in peer teaching and novice teaching but was one of the least observed in student teaching. This finding corroborates several studies which observed expert teachers (Derby, 2001; Dorfman, 2010; Taylor, 2009; Worthy, 2006; Worthy & Thompson, 2009). Similarly, Bond's study of student teacher's behaviors off and on the podium indicated that a third of participants were observed using positive modeling at a higher frequency than other categories. These findings seem to indicate that positive modeling is a behavior which increases with experience.

Concurrent performance model (CPM) was the second most observed category in frequency in student teaching and novice teaching. This finding is similar to Taylor's study of upper elementary Orff ensembles in which 30% of rehearsal was spent in CPM with a similar RPM as to the current study. Orman's study (2002) of elementary teachers indicated similar

findings of CPM in which half of the modeling observed was CPM. Derby's study (2001) of expert choral directors of various levels indicated a much smaller percentage of time (3.4%) while Patterson's (2009) was higher (10%). While many studies focus on positive and negative modeling, most studies do not address CPM as a main concern. It would be of interest to see if CPM decreases with experience or if specific levels (i.e. elementary general music vs. ensembles) utilize this strategy more than others.

TEEF DATA

Five observers were selected to review and evaluate the participants' videos. Each observer was responsible for evaluating one participant's videos (n = 8) and one video from each level of another participant's videos for reliability purposes. The five observers were expert veteran teachers with experience in the participant's area of expertise. Analysis of the videos included a summative evaluation form called the Teacher Effectiveness Evaluation Form (TEEF). A modified version of Hamann and Baker's Survey of Teaching Effectiveness (1996) was used. Observers responded to six items on a five-point Likert scale with one marked as poor and five marked as excellent with anchor definitions. The six items were information and demonstrations, musical model, flow, instructional directives, feedback, and teaching style. Additionally, the observers gave an overall rating of each lesson. These items were selected for analysis in that previous studies had found these aspects of teaching problematic in early career teachers (Anderson-Nichols, 1997; Bergee, 2005; Goolsby, 1996, 1997; 1999; Pike, 2014; Wagner & Strul, 1979).

The TEEF data was analyzed in three ways—career level, item, and by individual participant. In relation to levels, it was somewhat difficult to identify arching trends from peer teaching to novice teaching. However, there were some elements which provided evidence of micro-growth or lack thereof between levels. Participants' overall scores from Peer Teaching #1 to Peer Teaching #2 increased and modeling was rated higher for all participants in Peer Teaching #2. This somewhat corroborates Killian et al.'s (2008) study which tracked peer teaching growth using observational data. However, ratings did not increase uniformly from peer teaching to student teaching. Ratings of modeling, directives, and feedback were lower than in Peer Teaching #2. No trends were found in relation to Student Teaching #1 and Student Teaching #2. Findings were mixed in ratings from student teaching to peer teaching. The lack of increase between peer teaching to student teaching may support the findings of Powell's study (2018) regarding lack of agency in student teaching and novice teaching. The inability to track trends longitudinally may also indirectly align with Killian and Dye's (2009) study regarding preservice and student teacher's perceptions of learning transfer. However, the overall data seems to indicate that tracking early career levels may be difficult or impractical to quantify.

However, some items were a bit easier to identify and quantify. All participants received strong marks (4 or 5 out of a 5-point Likert scale) in information and demonstrations across all levels. This suggests that participants may have been able to transfer knowledge from each career level and present information effectively. This correlates with Bartolome's qualitative study (2017) which indicated similar findings regarding pedagogy and organization. Musical model ratings were higher in peer teaching and novice teaching but declined in student teaching. All levels indicated middle to low levels of flow which would seem to indicate stagnant growth.

Feedback was scored poorly in peer teaching and student teaching while half of the scores ranked high in novice teaching. However, most of those scores belonged to Participant C and D indicating poor feedback in novice teaching for Participants A, B, and E. Perhaps feedback is a teaching component with which most early career teachers struggle. This may relate back to Berliner's theory of expertise (1988) in which novices and advanced beginners lack personal agency and improvisational thinking skills necessary for delivering appropriate feedback. Teaching style ratings increased in novice teaching, although ratings of Participant B remained stagnant. An increase in teaching style indicates an overall level of comfort in the role as educator and consistent in the findings of longitudinal qualitative and descriptive data (Bartolome, 2017; Dabback, 2018; Killian & Dye, 2009).

From an individual perspective, Participant A was given high ratings in most items in peer teaching and student teaching, but lower ratings in novice teaching. Her lowest ratings were on the musical model and feedback items. Similarly, Participant B's highest ratings were limited to peer teaching. Flow was consistently marked low across all levels. Participant A and B had divergent student teaching settings that did not match their novice teaching assignments. Participant A student taught in an elementary general music setting to fifth and sixth graders at rural intermediate school and was currently teaching Pre-Kindergarten through first grade at a rural elementary school. Informal conversations with Participant A revealed a lack of mentoring and professional development opportunities in this setting. Participant B student taught elementary general music to third and fourth graders at a rural elementary school and was currently teaching Kindergarten through fifth grade at a suburban school. In an informal conversation with me, she stated that she was frustrated by the lack of feedback and overall

guidance from her clinical instructor. Informal conversation also disclosed the lack of a mentor teacher and opportunities for professional development. Participant A and B's experiences perhaps align with Goldhaber et al.'s study (2017) tracking homogenous and heterogenous student teaching and novice teaching placements. Their study indicated that teachers with dissimilar placements decrease in efficacy levels during novice teaching.

Participant C's highest ratings on all items were obtained in novice teaching with the lowest ratings in Peer Teaching #1. He received lower overall ratings in feedback and teaching style (scores ranging from 2 to 4). Similarly, Participant D's highest ratings occurred in novice teaching. However, the participant was given high ratings throughout all levels in information and demonstrations, flow, instructional directives, and feedback. His lower ratings were in teaching style (the highest ratings given were a 4). Participant C student taught middle school band at a rural middle school and was currently an assistant band director at a suburban middle school. The middle school and the high school which it fed into had a fluid relationship, and Participant C assisted with high school ensembles as well. In novice teaching, it was easy to observe the collaboration of directors from the two schools. Further, Participant C's major responsibility was to direct mostly percussion sectionals while another band director taught woodwinds and brass. Participant C was teaching in a diverse and rigorous program with a good number of seasoned teachers. He led sequenced rehearsals which at times lacked warmth; however, the students were responsive and worked at a high level of rigor. Participant D student taught a middle school band at a suburban middle school and was subsequently hired as the head band director at the same school following student teaching. Participant C and D taught at the same middle school. As stated above, the organization of the middle and high school band

programs was cohesive, and all directors of both schools were involved with all grade levels. Rapport between the students and Participant D was good and mistakes were a readily welcome part of learning. Like Participant C, Participant D also enjoyed mentorship opportunities with veteran teachers. Participant D's major role in the program was to work with percussion students. In informal conversation, he stated that Participant C and he assisted with the growth in the percussion program and that all sixth-grade percussion students (N = 16) continued in the seventh grade. They were concerned with ensuring rigor and providing all seventh graders with equal playing opportunities, as the number of percussionists was high. These findings seem to indicate that unlike Participants A and B, Participants C and D were part of a warm and nurturing atmosphere which encouraged meticulousness of the students and teachers. This would also seem to corroborate with Goldhaber et al.'s (2017) findings regarding similar placements in student teaching and novice teaching.

Participant E student taught middle school choir at a suburban school and was currently teaching middle school choir at a suburban school. Her highest ratings were in novice teaching #1 and #2. Most ratings in all levels ranged from 2-3 on many items. Like the other participants, information and demonstrations were rated highly across all levels. However, music model, flow, and instructional directives were rated low. The results from Participant E's TEEF data and informal observation by the researcher do not corroborate with Goldhaber et al.'s (2017) data in that Participant E taught in identical settings in student teaching and novice teaching. However, Participant A, B and E all had low ratings in flow, modeling and feedback, specifically in novice teaching. Perhaps their focus on extra-musical lesson content and lack of direction provided evidence for lackluster scores.

CORROBORATIVE FINDINGS

The purpose of this study was to observe the teaching of five music teachers in three early career teaching levels using video recordings to determine if effective teaching behavior increases over time. Further, the study's purpose was to find if any specific teacher behaviors associated with effectiveness are more present when teaching their peers in pre-service, another educator's students in student teaching, and their own students in novice teaching. Therefore, it is important to attempt to find corroborating results from the research questions posed regarding instructional targets, rehearsal frames, and TEEF data.

There was a general lack of specificity across all levels evidenced in instructional targets, feedback categories, and ratings on the TEEF, that improved somewhat during student teaching and peer teaching. Unidentified targets were the second most frequently observed instructional target during peer teaching and was one of the least frequently observed during student teaching and novice teaching. Nonspecific positive feedback was one of the highest verbal categories observed in selected rehearsal frames during peer teaching as well. Throughout all levels, specific positive and negative feedback were used at medium to low frequency levels. This finding is also similar to studies which compared expertise levels and those which focused only on expert teachers (Goolsby, 1996, 1999; Whitaker, 2017). Goolsby (1997, 1999) found that specific positive feedback was used more frequently by experts than by novices. Whitaker (2017) and Worthy and Thompson (2009) found that expert conductors provided more disapproval statements than positive feedback. Negative feedback was one of the more frequently observed categories in Waymire's (2011) comparison study with both effective and non-effective band

directors; however, his findings were different than the current study in that negative feedback was the second most observed category in less effective directors and was the fourth most observed in expert directors. These studies might suggest that novice teachers may desire to provide positive experiences for their students (hence nonspecific positive feedback), but do not have the experience and practice to utilize constructive and specific positive feedback statements. TEEF results also indicated mixed trends in overall feedback scores in that feedback improved during Peer Teaching #2, decreased in Student Teaching #1 but increased in Student Teaching #2 and Novice Teaching #1, but decreased in Novice Teaching #2 and #3. Further, evaluators commented on the lack of feedback or ignoring blatant problems and giving nonspecific positive feedback in the comments section of the TEEF. Comments from evaluators included, “Her feedback was limited and not very specific,” “There were several times when she responded with ‘great’ when more specific praise could have benefitted the students,” “All on behavior, not on music,” “She didn’t address singing voice when kids were shouting during song,” “Taking more questions might help with feedback,” “Little feedback given to singers, positive or negative,” “The parts weren’t being sung correctly, she moved on as if they were,” “Give more feedback during stick control exercises,” and “Did not address or remind students about chromatic fingerings.” These feedback traits indicate similar findings to studies focused on novice and ineffective teachers (Bergee, 2005; Bonds, 2015; Goolsby, 1996, 1997, 1999; Pike, 2014; Wagner & Strul,1979; Waymire, 2011).

Directives were the most frequently observed teacher verbal category and information and demonstrations were consistently the highest ranked items across all levels on the TEEF. Overall instructional directives scores were inconsistent on the TEEF and indicated high scores

for three participants in peer teaching, high scores for one participant in student teaching, and high scores for two participants in novice teaching. However, the frequency count of directives may not necessarily indicate the effectiveness of the directives and may, perhaps, indicate a lack of specificity. Evaluator's comments indicated similarly. Evaluators wrote comments such as, "Not clear when to sing versus listen," "I'm not sure what the goal of her lesson was," "Played pattern without practicing or talking about rhythm," "Change method of instruction when it doesn't work," and "It seems like the teacher is playing big chunks and hoping to find things that were wrong." In relation to this, Participant A and B consistently ignored or were unaware of nonparticipators and did not encourage them to be actively involved. Further, there were episodes in which classroom instruments were used and poor playing technique and noise level were not addressed. Similarly, comments were made by evaluators indicating similar episodes in all levels with all participants in at least one of the three levels. The high ratings in information and demonstrations coupled with the inconsistency of instructional directive scores across levels indicates that young teachers of all levels may have appropriate pedagogical knowledge, but have difficulty explaining this to students. This may corroborate with Berliner's theory of teaching expertise (1988) in novice and advanced beginner stages. There are several studies which corroborate the lack of proper communications with students in novice teachers (Bergee, 2005; Goolsby, 1996, 1997, 1999; Pike, 2004).

The amount of teacher modeling doubled during novice teaching. Further, positive modeling was the second most observed teacher categories of teacher verbal behaviors and teacher modeling behaviors and third in novice teaching. Concurrent performance model (CPM) was the second most observed teacher category during student teaching and novice teaching.

Many of the evaluators commented on the TEEF about the lack of competency-based education techniques during student teaching and novice teaching. Competency-based instruction indicates clear objectives with assessable behaviors, instruction which focuses on individual students, evaluations which indicate deep pedagogical knowledge, and a focus on student self-discovery (Arnold, 1995; Houston & Howsam, 1972). The evaluator for Participant A indicated that during novice teaching there was a "hodgepodge of random musical activities for the students to do rather than a clear, focused lesson on specific objectives" and "Any person without a music degree could have accomplished what Participant A did in this lesson." Participant B's evaluator during student teaching wrote, "Relied on demo and kids copying her too much—did they know the rhythms or were they just imitating her?" This continued during novice teaching including comments of "Did everything with students—give them more chances to be independent." Participant C's evaluator during student teaching commented, "Hold them accountable for more!" and "I don't have confidence the kids know how to count rhythms. Lots of rote teaching going on." Similar comments were made for Participant D during student teaching—"Did not give directives before rehearsal. Played a huge chunk of music with lots of problems then took several minutes to explain tenuto" and "Bad counters will copy the good counters." Similarly, novice teaching comments included, "Student still counting wrong but he didn't hear it because he was counting with him." Participant E's evaluator stated during student teaching, "Engage students more—empower them to be leaders in the rehearsal setting" and "When fixing a problem, work in smaller chunks." Similar statements were made during novice teaching such as "Give information in smaller chunks" and "She spent most of her time behind the piano, playing with them. How could she hear if they were wrong?"

Evaluators commented on techniques such as performing with students too often and using rote teaching when unnecessary. This is consistent with the observed high use of concurrent modeling in student teaching and novice teaching and perhaps may connect as to why participants were not observed teaching more nuanced instructional targets. This is consistent with the findings of the Goolsby studies (1996, 1997, 1999) as well as Waymire (2011) and Bonds (2015) studies of either ineffective band directors or student teachers.

Participants A, B, and E had very low rates of rehearsal frames with multiple performance trials (one every 50 minutes). This may also have contributed to their low TEEF scores in some if not most of the levels and video recordings observed. Specifically, these participants had low ratings in feedback, flow, and musical model across all levels. This supports studies which indicate that many novice teachers have difficulty with pace, talking instead of modeling, and provide general feedback instead of specific feedback (Bergee, 2005; Goolsby, 1996, 1997, 1999; Pike, 2004; Shleuter, 1991; Wagner & Strul, 1979; Waymire, 2011).

IMPLICATIONS FOR PRESERVICE TRAINING AND EARLY CAREER EDUCATORS

Results should be interpreted with caution; the number of participants was small (N = 5) and different video recordings and other ensembles might have resulted in different findings. Further, participants were responsible for video recording their own student teaching. Therefore, a wide range of video qualities may have distorted the perception of teaching behaviors and student responses. The presence of an observer and/or a video camera may have influenced teacher or student behaviors.

The primary purpose of this study was to observe the teaching of five music teachers in three early career teaching levels using video recordings to determine if effective teaching behavior increases over time. The resulting data from the research questions indicated that time can assist somewhat in early career levels. This can be evidenced in that specificity increased slightly in student teaching and novice teaching in relation to instructional targets and decrease of nonspecific positive feedback. However, there was no growth evidenced in specific positive or negative feedback. Very concerning was that three out of the five participants' TEEF scores decreased or flatlined in novice teaching. Rates of single trial rehearsal frames and multiple performance trial rehearsal frames also declined over time. Lack of growth in novice teaching may be due to the over-simulation in peer teaching and student teaching in that specific requirements are given to preservice teachers in these lessons. Therefore, micro-growth can be evidenced, but only to a certain degree. This data indicates that growth may be easier to determine at larger gaps of time such as the end of the novice period at the five-year point.

The secondary purpose of this study was to find if any specific teacher behaviors associated with effectiveness are more present when teaching their peers in pre-service, another educator's students in student teaching, and their own students in novice teaching. Data in peer teaching indicated higher use of unidentified targets and nonspecific positive feedback. No off- or on-task student talk was observed, indicating the lack of authenticity peer-teaching provides. Offering more opportunities in preservice using authentic context learning is necessary to simulate the real classroom later. Concurrent performance model was the second-most observed teacher behavior in student teaching and novice teaching. Teacher modeling doubled in novice teaching from peer teaching and student teaching. More data would need to be utilized to fully

understand the purpose of the modeling observed. However, when observing the novice teachers informally, it appeared to me that participants either were performing habitually with students or were doing so to over-guide students. Performing with students may indicate the necessity for teachers to have an opportunity to use their instrument, as time for doing so outside of teaching may be limited. Additionally, performing with students could be due to lack of trust in teaching skills or skills of students. Further examination for the purpose of concurrent performance modeling is necessary.

There are several studies and best practice articles which indicate that the process of altering novice teaching behaviors requires experience (Bergee, 2005; Murnane & Phillips, 1981; Rosenshine, 2008; Shulman, 1987). However, it is essential for those in higher education to be persistent in assisting preservice and early career teachers, as experience without guidance will prove less effective. Therefore, it is necessary to give novice teachers tools to become self-aware of problematic behaviors. Many studies performed during preservice and student teaching indicate that accountability in time allocation and verbalizations can increase self-awareness of problematic behaviors (Killian & Dye, 2009; Killian et al., 2008; Nápoles, 2017; Worthy, 2005). These accountability activities may include reviewing video recordings with an expert teacher as well as using observational software like SCRIBE to analyze teaching episodes. Other means might include structured opportunities in real time to adjust teacher behaviors with focusing on eye contact (Browning & Porter, 2007), student proximity (Bonds, 2015), teacher talk (Colwell, 1995; Nápoles, 2017; Worthy, 2005), and enthusiasm (Collins, 1978; Cassidy 1990). Further, studies have indicated that training in competency-based instruction during all levels of teaching can assist teachers in a better understanding of themselves as well as their students (Arnold,

1991, 1995; Benson, 1989; Yarbrough, Price, & Bowers, 1991). Results from these studies indicated increases in positive teacher behaviors.

One problematic issue, however, is the structured accountability that goes away when pre-service teachers become in-service teachers. All participants in this study went through the same accountability measures in peer teaching and student teaching and were all required to view their videos and record behaviors. However, TEEF data indicated decrease or flatline in effectiveness in three participants during novice teaching. Further, overall RPM of all rehearsal frames decreased over time across levels. Lack of competency-based teaching strategies by using concurrent performance model and rote teaching was prevalent in student teaching and novice teaching. Teacher modeling doubled in novice teaching versus student teaching. Therefore, the concern lies in how to create experiences in preservice teaching to motivate novice teachers to continue processing accountability measures and self-awareness.

For example, Participant D may be the novice teacher type who considers self-reflection in development, in that his peer teaching episodes were scored strong as well as his novice teaching episodes. In addition, he scored higher on the TEEF in areas where Participant A, B, and E were lacking in modeling, flow, feedback, and teaching style. Participant D's evaluator commented in his later novice teaching videos that he gave the students "plenty of opportunities for self-discovery" and "good individual instruction." Those in higher education might consider how to place more students in this trajectory. Opportunities would most likely include supporting novice teachers through mentorship and professional guidance.

Novice teachers need expert mentor music teachers to assist them with reflective practicum. While any mentor would certainly be able to assist novice teachers, it would be more

useful if the mentor teacher was in the same field, as there are different issues which might occur in a kindergarten general music class than a fifth-grade science class. School systems need to allow specialty teachers to collaborate and work with each other to become a support system, as this was stated as problematic for two participants. In this regard, school systems need to allow specialty teachers to attend professional development in their own field to assist in creating positive learning experiences for students.

CONSIDERATIONS FOR FURTHER RESEARCH

The results and discussion from this study raise numerous research possibilities. Some of these considerations are listed below:

1. The results from this study were limited to five participants. While data-heavy for each participant, having a larger pool of participants with similar data collection would most likely increase the validity of the current study. Further, comparing types of ACL experiences and accountability measures at different colleges in peer teaching and student teaching longitudinally may also provide interesting results.
2. While the scope of the current study attempted to document micro-growth across three early-career levels, it would be of interest to extend this study and observe the participants again after the first five years of teaching, as many experts agree that the highest gains in effectiveness occur during this period.
3. Concurrent performance modeling was prevalent in student teaching and novice teaching in the current study. While there are some studies which look at modeling, most studies

do not focus on this type of modeling. It would be of interest to observe preservice, novice, and expert teachers in relation to their use of concurrent performance model.

4. Preservice teaching in the current study was limited to peer teaching. However, it would be of interest to replicate the study using data from an actual classroom versus a lab setting of their peers to provide more authenticity. The lack of authenticity in the current study was seen in that there were no episodes of off-task or on-task student talking during peer teaching that was present during student teaching and novice teaching.
5. Only one study was found that utilized rehearsal frames as the unit of analysis for elementary general music classes in which the setting was with upper elementary Orff ensembles (Taylor, 2009). It would be of interest to observe elementary general music teachers using rehearsal frames as the unit of analysis over a broad range of grade levels to see if there are any trends within grade levels and/or expertise of teachers as well.
6. The lack of growth in three participants was concerning in the current study. While Waymire's study (2011) was performed considering highly effective and ineffective band directors, it would be of interest to attempt to identify reasons behind complacency and lack of growth. A study like this would most likely be organized using qualitative and descriptive data, although mixed methods could be used to possibility strengthen the argument seen in qualitative measures.

LIST OF REFERENCES

Anderson-Nichols, J. D. (1997). *Teacher expertise among elementary general music teachers* (Doctoral dissertation). Retrieved from Arizona State University Libraries.

Arnold, J. A. (1991). Using videotape self-analysis to improve teaching during rehearsals. *Update: Applications of Research in Music Education, 10*, 10-14. https://doi.org/10.1177/875512339101000103

Arnold, J. A. (1995). Effects of competency-based methods of instruction and self-observation on ensemble directors' use of sequential patterns. *Journal of Research in Music Education, 43,* 127-138. https://doi.org/10.2307/3345674

Babb. S. L. (2010). *Rehearsal techniques used to build choral tone by four expert collegiate choral conductors across settings* (Doctoral dissertation). Retrieved from http://purl.flvc.org/fsu/fd/FSU_migr_etd-0834

Baughman, M., & Baumgartner, C. M. (2018). Preservice teachers' experiences teaching an adult community music ensemble. *International Journal of Music Education*. https://doi.org/10.1177/0255761418775128

Bartolome, S. J. (2017). Comparing field-teaching experiences: A longitudinal examination of preservice and first-year teacher perspectives. *Journal of Research in Music Education, 65,* 264–286. https://doi.org/10.1177/0022429417730043

Beebe, M. (2007). *Teaching and rehearsal behaviors of instrumental music teachers* (Master's thesis). Retrieved from http://rave.ohiolink.edu/etdc/view?acc_num=bgsu1181923968

Benson, W. L. (1989). The effects of models, self-observation, and evaluation on the modification of specified teaching behaviors of an applied music teacher. *Update: Applications of Research in Music Education, 7,* 28-31. https://doi.org/10.1177/875512338900700210

Bergee, M. J. (2005). An exploratory comparison of novice, intermediate, and expert orchestral conductors. I*nternational Journal of Music Education*, *23,* 23–36. https://doi.org/10.1177/0255761405050928

Berkley, L. O. (2011). *The impact of band directors' teaching experience level on the use of selected rehearsal behaviors* (Master's thesis). Retrieved from UMD theses and dissertations.

Berliner, D. C. (1986). In pursuit of the expert pedagogue. *Educational Researcher, 15,* 5-13. https://doi.org/10.3102/0013189X015007007

Berliner, D. C. (1988). Implications of studies of expertise in pedagogy for teacher education and evaluation. In New directions for teacher assessment: Proceedings of the 1988 ETS Invitational Conference (pp. 39–68). Princeton, NJ: Educational Testing Service.

Biasutti, M., & Concina, E. (2017). The effective music teacher: The influence of personal, social, and cognitive dimensions on music teacher self-efficacy. *Musicae Scientiae, 22,* 1-16. https://doi.org/10.1177/1029864916685929

Blackwell, J. A., & Roseth, N. E. (2018). Problem-based learning in a woodwind methods course: an action research study. *Journal of Music Teacher Education, 28*, 55–69. https://doi.org/10.1177/1057083718769262

Bloom, B. S. (1986). Automaticity: The hands and feet of genius. *Educational Leadership, 43*(5), 70-77.

Bonds, E. C. (2015). *An analysis of student teachers' instruction in a beginning band setting* (Doctoral Dissertation). Retrieved from ProQuest Dissertation and Theses Database. (3726949)

Bowers, J. K. (1990). *The relationship of sequential patterns to the music teaching effectiveness of elementary education majors* (Doctoral Dissertation). Retrieved from https://digitalcommons.lsu.edu/cgi/viewcontent.cgi?referer=https://scholar.google.com/&httpsredir=1&article=6035&context=gradschool_disstheses

Browning, B. P., & Porter, A. M. (2007). The effect of computer-assisted self-observation on the eye contact behaviors of preservice music teachers. *Journal of Music Teacher Education, 17,* 62-76. https://doi.org/10.1177/10570837070170010110

Butler, A. (2001). Preservice music teachers' conceptions of teaching effectiveness, microteaching experiences, and teaching performance. *Journal of Research in Music Education, 49,* 258-272. https://doi.org/10.2307/3345711

Byo, J. L. (1990). Recognition of intensity contrasts in the gestures of beginning conductors. *Journal of Research in Music Education, 38,* 157-163. https://doi.org/10.2307/3345179

Cassidy, J. L. (1990). Effect of intensity training on preservice teachers' instruction accuracy and delivery effectiveness. *Journal of Research in Music Education, 38,* 164-174. https://doi.org/10.2307/3345180

Cavitt, M. E. (2003). A descriptive analysis of error correction in instrumental music rehearsals. *Journal of Research in Music Education, 51,* 218-230. https://doi.org/10.2307/3345375

Chapman, D. (2014). *Effects of observation duration on evaluations of teaching in secondary school band and choir rehearsals* (Doctoral dissertation). Retrieved from UT electronic theses and dissertations.

Chingos, M. M., & Peterson, P. E. (2011). It's easier to pick a good teacher than to train one: Familiar and new results on the correlates of teacher effectiveness. *Economics of Education Review, 30,* 449-465. https://doi.org/10.1016/j.econedurev.2010.12.010

Clark, S. K., Byrnes, D., & Sudweeks, R. R. (2014). A comparative examination of student teacher and intern perceptions of teaching ability at the preservice and inservice stages. *Journal of Teacher Education, 66,* 170–183. https://doi.org/10.1177/0022487114561659

Collins, M. L. (1978). Effect of enthusiasm training on preservice elementary teachers. *Journal of Teacher Education. 29*, 53-57. https://doi.org/10.1177/002248717802900120

Colprit, E. J. (2000). Observation and analysis of Suzuki string teaching. *Journal of Research in Music Education, 48*, 206-221. https://doi.org/10.2307/3345394

Colwell, C. M. (1995). Effect of teaching setting and self-evaluation on teacher intensity behaviors. *Journal of Research in Music Education, 43*, 6–21. https://doi.org/10.2307/3345788

Culp, D. J. (2018). *A descriptive analysis of rehearsal techniques of expert middle school band directors from Kentucky, Mississippi, and Tennessee* (Doctoral Dissertation). Retrieved from ProQuest Dissertation and Theses Database. (13419148)

Dabback, W. (2018). A longitudinal perspective of early career music teachers: Contexts, interactions, and possible selves. *Journal of Music Teacher Education, 27*, 52–66. https://doi.org/10.1177/1057083717727268

Derby, S. E. (2001). *Rehearsal of repertoire in elementary, middle, and high school choirs: How teacher effect change in student performance* (Doctoral dissertation). Retrieved from ProQuest dissertation and theses database. (3008313)

Dorfman, J. (2010). The relationships between time usage during instrumental lessons and preservice teachers' self-evaluations. *Journal of Music Teacher Education, 19*, 89-99. https://doi.org/10.1177/1057083709359885

Dreyfus, S. E., & Dreyfus, H. L. (1980). A five-stage model of the mental activities involved in directed skill acquisition. Unpublished report, University of California, Berkeley.

Dreyfus, H. L., & Dreyfus, S. E. (1986). Mind over machine: The power of human intuition and expertise in the age of the computer. New York, NY: Free Press.

Duke, R. A. (2019). *Scribe 5 (beta version) for Observation and Assessment.* Austin, TX: Learning & Behavior Resources

Duke, R. A. (1994). Bringing the art of rehearsing into focus: the rehearsal frame as a model for prescriptive analysis of rehearsal conducting. *Journal of Band Research, 30*, 78-95. Retrieved from http://umiss.idm.oclc.org/login?url=https://search-proquest-com.umiss.idm.oclc.org/docview/1091951?accountid=14588

Duke, R. A. (1999/2000). Measures of instructional effectiveness in music research. *Bulletin of the Council for Research in Music Education, 143,* 1-48. Retrieved from http://www.jstor.org/stable/40319011

Duke, R. A., & Buckner, J. J. (2009). Watching learners learn. *MTNA e-Journal, 1,* 17-28.

Duke, R. A., & Chapman, D.-L. (2011). Changing learners: The nature of expertise in music teaching. In P. M. Ward-Steinman (Ed.), Advances in Social-psychology and music education research (pp. 29–38). New York, NY: Routledge.

Duke, R. A., & Simmons, A. L. (2006). The nature of expertise: narrative descriptions of 19 common elements observed in the lessons of three renowned artist-teachers. *Bulletin of the Council for Research in Music Education, 170*, 7-19. Retrieved from http://www.jstor.org/stable/40319345

Evans, C. J. (2012). *Listening priorities and rehearsal procedures of selected Texas honor band directors* (Doctoral dissertation). Retrieved from ProQuest dissertation and theses database. (3506221)

Fant, G. R. (1996). *An investigation of the relationships between undergraduate music education students' early field experience and student teaching performance* (Doctoral dissertation). Retrieved from https://repository.arizona.edu/handle/10150/290634

Frederickson, W. E. (1992). Research on eye contact with implications for the conductor: A review of the literature. *Update: Applications of Research in Music Education, 11,* 25-31. https://doi.org/10.1177/875512339201100106

Gao, M., & Liu, Q. (2013). Personality traits of effective teachers represented in the narratives of American and Chinese preservice teachers: A cross-cultural comparison. *International Journal of Humanities and Social Science, 3,* 84-95. Retrieved from http://www.ijhssnet.com/journals/Vol_3_No_2_Special_Issue_January_2013/9.pdf

Gatbonton, E. (2008). Looking beyond teachers' classroom behaviour: Novice and experienced ESL teachers' pedagogical knowledge. *Language Teaching Research, 12,* 161–182. https://doi.org/10.1177/1362168807086286

Goldhaber, D., Krieg, J. M., & Theobald, R. (2017). Does the match matter? Exploring whether student teaching experiences affect teacher effectiveness. *American Educational Research Journal, 54*, 325–359. https://doi.org/10.3102/0002831217690516

Goodrich, A., Bacura, E., & Stauffer, S. (2017). Peer mentoring in a university music methods class. *Journal of Music Teacher Education, 27*, 23–38. https://doi.org/10.1177/1057083717731057

Goolsby, T. W. (1996). Time use in instrumental rehearsals: A comparison of experienced, novice, and student teachers. *Journal of Research in Music Education, 44*, 286-303. https://doi.org/10.2307/3345442

Goolsby, T. W. (1997). Verbal instruction in instrumental rehearsals: a comparison of three career levels and preservice teachers. *Journal of Research in Music Education, 45*, 21-40. https://doi.org/10.2307/3345463

Goolsby, T. W. (1999). A comparison of expert and novice music teachers' preparing identical band compositions: An operational replication. *Journal of Research in Music Education, 47*, 174-187. https://doi.org/10.2307/3345722

Hamann, D. L., Lineburgh, N., & Paul, S. (1998). Teaching effectiveness and social skill development. *Journal of Research in Music Education, 46*, 87-101. https://doi.org/10.2307/3345762

Hamid, S. R. A., Hassan, S. S. S., & Ismail N. A. H. (2012). Teaching quality and performance among experienced teachers in Malaysia. *Australian Journal of Teacher Education, 37*, 85-103. Retrieved from http://dx.doi.org/10.14221/ajte.2012v37n11.2

Haston, W., & Russell, J. A. (2012). Turning into teachers: Influences of authentic context learning experiences on occupational identity development of preservice music teachers. *Journal of Research in Music Education, 59*, 369-392. https://doi.org/10.1177/0022429411414716

Heath-Reynolds, J. (2014). *The effect of teacher nonverbal expressiveness on ratings of teacher effectiveness and student learning* (Doctoral dissertation). Retrieved from ProQuest dissertation and theses database. (3638002)

Hendel, C. (1995). Behavioral characteristics and instructional patterns of selected music teachers. *Journal of Research in Music Education, 43,* 182-203. https://doi.org/10.2307/3345635

Henninger, J. C. (2002). The effects of knowledge of instructional goals on observations of teaching and learning. *Journal of Research in Music Education, 50*, 37–50. https://doi.org/10.2307/3345691

Houston, R., & Howsam R. (1972). *Competency-based teacher education.* Chicago: Science Research Associates.

Hoy, A. W., & Spero, R. B. (2005). Changes in teacher efficacy during the early years of teaching: A comparison of four measures. *Teaching and Teacher Education, 21*, 343-356. https://doi.org/10.1016/j.tate.2005.01.007

Hughes, W. O. (1992). The effect of high versus low teacher affect and active versus passive student activity during music listening on high school general music students' attention. *Research Perspectives in Music Education, 2*, 16-18.

Ihas, D. A. (2011). *Teaching behaviors of middle and high school orchestra directors in the rehearsal setting* (Doctoral Dissertation). Retrieved from the University of Oregon Scholars' Bank.

Juchniewicz, J. A. (2008). *The influence of social intelligence on effective music teaching* (Doctoral dissertation). Retrieved from http://purl.flvc.org/fsu/fd/FSU_migr_etd-3715

Kavanagh, S. S., & Rainey, E. C. (2017). Learning to support adolescent literacy: Teacher educator pedagogy and novice teacher take up in secondary English language arts teacher preparation. *American Educational Research Journal*, *54*, 904–937. https://doi.org/10.3102/0002831217710423

Keller, M. (2011). *Teacher enthusiasm in physics instruction* (Doctoral dissertation). Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.1021.3772&rep=rep1&type=pdf

Killian, J. N., Dye, K. G., & Buckner, J. J. (2008). The effect of self-directed peer teaching on undergraduate acquisition of specified music teaching skills. *Texas Music Education Research.* Available at http://www.tmea.org/ 080_College/Research/Kil2008.pdf

Killian, J. N., & Dye, K. G. (2009). Effects of learner centered activities in preparation of music educators: Finding the teacher within. *Journal of Music Teacher Education*, *19*, 9–24. https://doi.org/10.1177/1057083709343904

Kim, Y. (2016). *Teaching a traditional Korean art form: Descriptive analyses of individual lessons taught by three gayageum sanjo master teachers* (Doctoral dissertation). Retrieved from UT Electronic theses and dissertations.

Kini, T., & Podolsky, A (2016). Does teaching experience increase teacher effectiveness? A review of the literature. In *Learning Policy Institute.* Retrieved from https://learningpolicyinstitute.org/our-work/publications-resources/does-teaching-experience-increase-teacher-effectiveness-review-research.

Levin, D. M., Hammer, D., & Coffey, J. E. (2009). Novice teachers' attention to student thinking. *Journal of Teacher Education, 60,* 142-154. https://doi.org/10.1177/0022487108330245

Liu, S., & Meng, L. (2009). Perceptions of teachers, students and parents of the characteristics of good teachers: A cross-cultural comparison of China and the United States. *Educational Assessment, Evaluation and Accountability, 21*. https://doi.org/10.1007/s11092-009-9077-z

Livingston, C., & Borko, H. (1989). Expert-novice differences in teaching: a cognitive analysis and implications for teacher education. *Journal of Teacher Education, 40,* 36–42. https://doi.org/10.1177/002248718904000407

MacLeod, R. B., & Nápoles J. (2011). Preservice teachers' perceptions of teaching effectiveness during teaching episodes with positive and negative feedback. *Journal of Music Teacher Education, 22,* 91-102. DOI: 10.1177/1057083711429851

MacLeod, R. B., & Nápoles, J. (2014). The influences of teacher delivery and student progress on experienced teachers' perceptions of teaching effectiveness. *Journal of Music Teacher Education, 24,* 24-36. https://doi.org/10.1177/1057083714527111

Madsen, C. K., Standley, J. M., & Cassidy J. W. (1989). Demonstration and recognition of high and low contrasts in teacher intensity. *Journal of Research in Music Education, 37,* 85–92. https://doi.org/10.2307/3344700

Madsen, C. K., Standley, J. M., Byo, J. L., & Cassidy, J. W. (1992). Assessment of effective instrumental music student teachers and experts. *Update: Applications of Research in Music Education, 33,* 20-24. https://doi.org/10.1177/875512339201000206

Marsh, H. W. (2007). Do university teachers become more effective with experience? A multilevel growth model of students' evaluations of teaching over 13 years. *Journal of Education Psychology, 99,* 775-790. DOI: 10.1037/0022-0663.99.4.775

McCall, M. J. (2008). *Qualities of effective secondary science teachers: Perspectives of university biology students* (Doctoral dissertation). Retrieved from https://baylor-ir.tdl.org/baylor-ir/bitstream/handle/2104/5244/madelon_mccall_edd.pdf?sequence=1

Mehrpour, S., & Mirsanjari Z. (2016). Investigating the manifestation of teaching expertise feature among novice and experienced EFL teachers. *Journal of English Language Teaching and Learning, 18,* 167-198. Retrieved from http://elt.tabrizu.ac.ir/article_5503_5ce48ab3947184db1ba3e9098bc53b43.pdf

Miksza P., Roeder, M., & Biggs, D. (2010). Surveying Colorado band directors' opinions of skills and characteristics important to successful music teaching. *Journal of Research in Music Education, 57,* 364-381. https://doi.org/10.1177/0022429409351655

Miksza, P., & Berg, M. H. (2013). Transition from student to teacher: frameworks for understanding preservice music teacher development. *Journal of Music Teacher Education, 23,* 10–26. https://doi.org/10.1177/1057083713480888

Millican, J. S. (2009). Band and orchestra teachers' rankings of general pedagogical knowledge and skill. *Journal of Music Teacher Education, 19,* 68-79. https://doi.org/10.1177/1057083709344045

Millican, J. S. (2017). Examining pedagogical content knowledge of an expert band director teaching lips slurs. *Journal of Music Teacher Education, 26,* 90-103. https://doi.org/10.1177/1057083716629610

Millican, J. S., & Forrester, S. H. (2018). Core practices in music teaching: A Delphi expert panel survey. *Journal of Music Teacher Education, 27,* 51–64. https://doi.org/10.1177/1057083717736243

Montemayor, M. (2014). Evaluative and behavioral correlates to intrarehearsal achievement in high school bands. *Journal of Research in Music Education, 62,* 33-51. https://doi.org/10.1177/0022429413520010

Murnane, R. J., & Phillips, B. R. (1981). Learning by doing, vintage, and selection: Three pieces of the puzzle relating teaching experience and teaching performance. *Economics of Education Review, 1,* 453-465. https://doi.org/10.1016/0272-7757(81)90015-7

Murray, H. G., Rushton, J. P, & Paunonen, S. V. (1990). Teacher personality traits and student instructional ratings in six types of university courses. *Journal of Education Psychology, 82,* 250-261. http://dx.doi.org/10.1037/0022-0663.82.2.250

Nápoles, J. (2006). The relationship between type of teacher talk and student attentiveness. *Journal of Music Teacher Education, 16,* 7-19. https://doi.org/10.1177/10570837060160010103

Nápoles, J., & MacLeod, R. B. (2013). The influences of teacher delivery and student progress on preservice teachers' perceptions of teaching effectiveness. *Journal of Research in Music Education, 61,* 249-261. DOI: 10.1177/0022429413497234

Nápoles, J., & MacLeod, R. B. (2015). Influences of teacher delivery, student engagement, and observation focus on preservice teachers' perceptions of teaching effectiveness. *Journal of Music Teacher Education, 25,* 53-64. DOI: 10.1177/1057083715580436

Nápoles, J. (2006). The relationship between type of teacher talk and student attentiveness. *Journal of Music Teacher Education, 16,* 7–19. https://doi.org/10.1177/10570837060160010103

Nápoles, J. (2016). Teacher talk and perceived teacher effectiveness: An exploratory study. *Update: Applications of Research in Music Education, 35*, 5-10. https://doi.org/10.1177/8755123315626228

O'Meara. K. T. (2007). C*haracteristics of effective teachers in the Air Force's Squadron Officer College* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses Database. (3273379)

Orman, E. K. (2002). Comparison of the national standards for music education and elementary music specialists' use of class time. *Journal of Research in Music Education, 50*, 155-164. https://doi.org/10.2307/3345819

Patterson, M. (2009). *Analysis of instructional time use and preferred teaching strategies of three highly successful choral directors* (Doctoral dissertation). Retrieved from Electronic theses and dissertations.

Paul, S. J. (1998). The effects of peer teaching experiences on the professional teacher role development of undergraduate instrumental music education majors. *Bulletin of the Council for Research in Music Education,137*, 73-92. Retrieved from http://www.jstor.org/stable/40318933

Paul, S. J., Teachout, D. J., Sullivan J. M., Kelly, S. N., Bauer, W. I., & Raiber, M. A. (2001). Authentic-context learning activities in instrumental music teacher education. *Journal of Research in Music Education, 49*, 136-145. http://dx.doi.org/10.2307/3345865

Pike, P. D. (2014). The differences between novice and expert group-piano teaching strategies: A case study and comparison of beginning group piano classes. *International Journal of Music Education, 32*, 213– 227. https://doi.org/10.1177/0255761413508065

Polk, J. A. (2009). Traits of effective teachers. *Arts Education Policy Review, 107*, 23-29. DOI: 10.3200/AEPR.107.4.23-29

Powell, S. R. (2014). The influence of video reflection on preservice music teachers' concerns in peer and field-teaching settings. *Journal of Research in Music Education, 63*, 487-507. https://doi.org/10.1177/0022429415620619

Powell, S. R. (2018). Structure and agency in novice music teaching. *Research Studies in Music Education*. https://doi.org/10.1177/1321103X18794514

Powell, S. R., Weaver, M. A., & Henson, R. K. (2018). Relationship of preservice music teachers' primary instrument background and teaching effectiveness in brass and woodwind techniques classes. *Journal of Music Teacher Education, 27*, 38–50. https://doi.org/10.1177/1057083717733469

Price, H. E., Ogawa, Y., & Arizumi, K. (1997). A cross-cultural examination of music instruction analysis and evaluation techniques. *Bulletin for the Council for Research in Music Education 133*, 121-128. Retrieved from http://www.jstor.org/stable/40318851

Redding, J. B. (2011). T*he effect of choral conductor intensity on the perception of teacher effectiveness* (Doctoral dissertation). Retrieved from ProQuest Dissertation and Theses Database. (3483596)

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica, 73,* 417-458. https://doi.org/10.1111/j.1468-0262.2005.00584.x

Roesler, R. A. (2013). *Development and application of a framework for observing problem solving by teachers and students in music* (Doctoral dissertation). Retrieved from UT Electronic theses and dissertations.

Rohwer, D. & Henry, W. (2004). University teachers' perceptions of requisite skills and characteristics of effective music teachers. *Journal of Music Teacher Education, 13*, 18-27. https://doi.org/10.1177/10570837040130020104

Rosenshine, B., Froehlich, H., & Fakhour, I. (2002). Systematic instruction. In R. Colwell & C. Richardson (Eds.), The new handbook of research on music teaching and learning (pp. 299-314). New York, NY: Oxford University Press.

Rosenshine, B. (2008). Systematic instruction. In T. L. Good (Ed.), *21st Century Education: A Reference Handbook* (pp. 235-243). California: SAGE Publications.

Running, D. J. (2011). Charisma, conductors, and the affective communication test. *Journal of Band Research, 47*, 18-28. Retrieved from https://search.proquest.com/docview/918007901?pq-origsite=gscholar

Schleutter, L. (1991). Student teachers' preactive and postactive curricular thinking. *Journal of Research in Music Education, 39,* 46-63. https://doi.org/10.2307/3344608

Schmidt, M. (2010). Learning from teaching experience: Dewey's theory and preservice teachers' learning. *Journal of Research in Music Education, 58,* 131–146. https://doi.org/10.1177/0022429410368723

Sheldon, D. A., Williams, H., Carrillo, J., Brittin, R. V., Choi, G., Bruno, Z., … Batten, E. (July, 2008). Evidence of the development of higher order thinking skills in instrumental music instruction. In Barret, M., & Johnson, C. (Chairs), International seminar on research in music education. Symposium conducted at meeting of The International Society for Music Education, Portugal.

Shulman, L. S. (1987). Knowledge and teaching: Foundations of a new reform. *Harvard Education Review 57*, 1-22. https://doi.org/10.17763/haer.57.1.j463w79r56455411

Siebenhaler, D. J. (1997). Analysis of teacher-student interactions in the piano lessons of adults and children. *Journal of Research in Music Education, 45*, 6-20. https://doi.org/10.2307/3345462

Siedentop, D. and Eldar, E. (1989). Expertise, experience, and effectiveness. *Journal of Teaching in Physical Education*, *8*, 254-260. https://doi.org/10.1123/jtpe.8.3.254

Silveira, J. M. (2014). The perception of pacing in a music appreciation class and its relationship to teacher effectiveness and teacher intensity. *Journal of Research in Music Education*, *62*, 302–318. https://doi.org/10.1177/0022429414542978

Silvey. B. A., & Baumgartner, C. M. (2016). Effects of nonverbal, verbal, and traditional rehearsal methods on ratings of conductor effectiveness and ensemble expressivity. *Journal of Band Research, 51,* 30-45, 67-68. Retrieved from http://umiss.idm.oclc.org/login?url=https://search-proquest-com.umiss.idm.oclc.org/docview/1801645227?accountid=14588

Singletary, L. (2016). *Time use and instructional focus in beginning and advanced middle school band settings* (Doctoral dissertation). Retrieved from ProQuest dissertations and theses database. (10192942)

Taylor, D. M. (2009). Refining learned repertoire for percussion instruments in an elementary setting. *Journal of Research in Music Education*, *54*, 231-243. https://doi.org/10.1177/002242940605400306

Wagner, M. J., & Strul, E. P. (1979). Comparisons of beginning versus experienced elementary music educators in the use of teaching time. *Journal of Research in Music Education, 27,* 113-125. https://doi.org/10.2307/3344897

Waymire, M. D. (2011). *Behavioral analysis of directors of high-performing versus low-performing high school bands* (Doctoral Dissertation). Retrieved from ProQuest dissertations and theses database. (3479034)

Westbrook, G. W. (2004). An investigation of the effects of teacher personality on teacher behaviors in the instrumental music classroom: A path analysis. *International Journal of Arts and Commerce*, *2*, 129-144. Retrieved from http://www.ijac.org.uk/images/frontImages/gallery/Vol._2_No._2/13.pdf

Westerman, D. A. (1991). Expert and novice teacher decision making. *Journal of Teacher Education*, *42*, 292–305. https://doi.org/10.1177/002248719104200407

Whitaker, J. A. (2008). *Analyses of high school band students' and directors' perceptions of verbal and nonverbal teaching behaviors* (Doctoral dissertation). Retrieved from https://digitalcommons.lsu.edu/cgi/viewcontent.cgi?referer=https://scholar.google.com/&httpsredir=1&article=2971&context=gradschool_dissertations

Whitaker, J. A. (2017). Professional orchestral conductors' use of selected teaching behaviors in Rehearsal. *International Journal of Music Education, 35,* 165-174. https://doi.org/10.1177/0255761415622045

Williams, L. A. C. (2016). A *comparative study of teaching behaviors of band and orchestra directors in ensemble rehearsals* (Doctoral dissertation). Retrieved from Electronic theses and dissertations.

Worthy, M. D. (2003). Rehearsal frame analysis of an expert wind conductor in high school vs. college band rehearsals. *Bulletin of the Council for Research in Music Education, 156,* 11-19. Retrieved from http://www.jstor.org/stable/40319170

Worthy. M. D. (2005). The effects of self-evaluation on the timing of teacher and student behaviors in lab rehearsals. *Journal of Music Teacher Education, 15,* 8-14. https://doi.org/10.1177/10570837050150010103

Worthy, M. D. (2006). Observations of three expert wind conductors in college rehearsals. *Bulletin of the Council for Research in Music Education,168,* 51-61. Retrieved from http://www.jstor.org/stable/40319460

Worthy, M. D., & Thompson, L. (2009). Observation and analysis of expert teaching in beginning band. *Bulletin of the Council for Research in Music Education,180,* 29-41. Retrieved from http://www.jstor.org/stable/40319318

Yarbrough, C., & Price H. E. (1981). Prediction of performer attentiveness based on rehearsal activity and teacher behavior. *Journal of Research in Music Education, 29,* 209-217. https://doi.org/10.2307/3344994

Yarbrough, C., & Price H. E. (1989). Sequential patterns of instruction in music. *Journal of Research in Music Education, 37,* 179–187. https://doi.org/10.2307/3344668

Yarbrough, C., Price, H. E., & Bowers, J. (1991). The effect of knowledge of research on rehearsal skills and teaching values of experienced teachers. *Update: Applications of Research in Music Education, 9,* 17-20. https://doi.org/10.1177/875512339100900204

Yarbrough, C., Price, H., & Hendel, C. (1994). The effect of sequential patterns and modes of presentation on the evaluation of music teaching. *Bulletin of the Council for Research in Music Education, 120,* 33-45. Retrieved from http://www.jstor.org/stable/40318632

LIST OF APPENDICES

APPENDIX A

PARTICIPANT RECRUITMENT LETTER

Date

[Name of Participant]
[Name of Participant's School]
[Address of School]
[City, State, Zip Code]

Re: The Effects of Experience on Teacher Effectiveness Levels of Preservice Teachers, Student Teachers, and Novice Teachers; Alicia Canterbury, Investigator

Dear [Paricipant],

I am writing to let you know about an opportunity to participate in a research study about the effects of experience on teacher effectiveness levels of preservice teachers, student teachers, and novice teachers. This study is being conducted by Alicia Canterbury at the University of Mississippi. The purpose of this study will be to view video segments of you during three levels of teaching experience (peer teaching, student teaching, and novice teaching) to discover if effective teaching behaviors increase over time and to attempt to locate trends within all three levels of teaching experience.

You previously participated in a pilot research study about the same topic. At that time, you indicated an interest in possibly participating in future research. I am writing to tell you that I believe you may be interested in an approved research study about growth in effective teaching behaviors across three levels of teaching experience in which I will video-tape you over a three-day consecutive period of teaching the same ensemble or grade level. The university granted prior access to existing materials available including your peer teaching videos from MUS 422 and your videos from student teaching.

Agreement to be contacted or a request for more information does not obligate you to participate in any study. If you would like additional information about this study, please call Alicia Canterbury at 912-215-2935 or e-mail them at acanterb@olemiss.edu.

Thank you again for considering this research opportunity.

Sincerely,

Alicia Canterbury, PhD Candidate
University of Mississippi

APPENDIX B

STUDY CONSENT FORM

Consent to Participate in Research

Study Title: *Measures of Music Teacher Effectiveness in Three Early Career Levels*

Investigator
Alicia L. Canterbury
Department of Music
164 Music Building
University of Mississippi
University, MS 38677
(662) 915-7268
acanterb@olemiss.edu

Faculty Sponsor
Michael Worthy, PhD
Department of Music
164 Music Building
University of Mississippi
University, MS 38677
(662) 915-7268
mworthy@olemiss.edu

The purpose of this study:

The purpose of this study is to view video segments of three levels of teaching experience (peer teaching, student teaching, and novice teaching) of the same participant to discover if effective teaching behaviors increase over time and to attempt to locate behavioral trends within all three levels of teaching experience.

What will be done in this study:

1. Participants will be video recorded over a consecutive three-day period teaching the same ensemble and/or grade level for the entire duration of a lesson. The data will be used for summative analysis of teaching behaviors, specific teaching behaviors, and how teachers relate content to student performance.
2. The video camera will be focused on the teacher and not on the students.
3. The video will not be shared with any person outside of the study. The primary investigator, the faculty sponsor, and a reliability partner will be the only experts viewing your teaching.

Time required for this study

Three 50-minute teaching episodes.

Possible risks from your participation:

There are no anticipated risks to you from participating in the study.

Benefits from your participation

You should not expect benefits from participating in this study. However, you might experience satisfaction from contributing to scientific knowledge and perhaps self-reflection in considering levels of effectiveness in teaching.

Confidentiality

All information in the study will be collected from you anonymously: it will not be possible for anyone to associate you with the data collected from your teaching episodes. The primary researcher and a reliability partner will have access to your videos only.

Confidentiality and Use of Video/Audio Recordings

- Video recordings are necessary in that this will allow the researcher to collect data using a summative evaluation instrument as well as identifying specific teaching behaviors using a data collection software.
- Only primary investigator, the faculty sponsor, and reliability partner on the research team will have access to these videos.
- Video recordings will be kept indefinitely and will be stored on an external flash drive in a filing cabinet in a locked office.

Right to Withdraw

You do not have to volunteer for this study, and there is no penalty if you refuse. If you start the study and decide that you do not want to finish, just *tell the experimenter*. Whether or not you participate or withdraw will not affect your current or future relationship with the *Department of Music* or with the University.

IRB Approval

This study has been reviewed by The University of Mississippi's Institutional Review Board (IRB). If you have any questions or concerns regarding your rights as a research participant, please contact the IRB at (662) 915-7482 or irb@olemiss.edu.

Please ask the researcher if there is anything that is not clear or if you need more information. When all your questions have been answered, then decide if you want to be in the study or not.

Statement of Consent

I have read the above information. I have been given an unsigned copy of this form. I have had an opportunity to ask questions, and I have received answers. I consent to participate in the study.

Furthermore, I also affirm that the experimenter explained the study to me and told me about the study's risks as well as my right to refuse to participate and to withdraw.

Signature of Participant Date

Printed name of Participant

APPENDIX C

Teacher Effectiveness Evaluation Form

Participant's Letter: _____ Evaluator's Initials: ____ _____

Circle the video being evaluated:

PT 1 PT 2 ST 1 ST 2 NT 1 NT 2 NT 3 NT 4

Directions: Immediately after watching the video, evaluate the following behaviors and assign ratings to the best of your ability based on the observed teaching episode.

Information and Demonstrations: (Poor) 1 2 3 4 5 (Excellent)

Excellent—Presented correct information; accurate demonstrations
Poor—Presented incorrect, contradictory, or misleading information; did not or could not accurately demonstrate (i.e., clapped or sang incorrect rhythms; did not demonstrate or provide information)

Musical Model: (Poor) 1 2 3 4 5 (Excellent)

Excellent—Expressive and accurate
Poor—Non-expressive, incorrect or inappropriate modeling; no modeling evidenced

Flow: (Poor) 1 2 3 4 5 (Excellent)

Excellent—Appropriate balance between teacher directives/explanations and student participation; one activity led to another without interruptions or breaks
Poor—Teacher talked too much; too much time spent going from one activity to another; long, disruptive breaks between and within activities

Instructional directives: (Poor) 1 2 3 4 5 (Excellent)

Excellent—Specific directives identifying tasks to be accomplished
Poor—Non-specific directives with no specific tasks to accomplished

Feedback: (Poor) 1 2 3 4 5 (Excellent)

Excellent—Specific positive or negative feedback provided; utilized student ideas and comments when/where applicable
Poor—No feedback or non-specific feedback provided

Teaching Style: (Poor) 1 2 3 4 5 (Excellent)

Excellent—Secure, animated; captured students attention and interest
Poor—Sluggish, lethargic, insecure; students were bored or disinterested

Overall rating of lesson: (Poor) 1 2 3 4 5 (Excellent)

Comments: _____

_____

_____

APPENDIX D

TEEF Scores by Level

TEEF Scores
Peer Teaching #1

	Participant A	Participant B	Participant C	Participant D	Participant E
Information and Demonstrations	4	4	2	4	3
Musical Model	4	5	2	2	3
Flow	4	2	2	4	3
Instructional Directives	4	5	2	4	3
Feedback	3	5	2	3	2
Teaching Style	4	4	2	2	3
Overall Rating	4	4	2	3	3
Total	27	29	14	22	20

TEEF Scores
Peer Teaching #2

	Participant A	Participant B	Participant C	Participant D	Participant E
Information and Demonstrations	4	4	4	5	3
Musical Model	4	5	4	4	3
Flow	5	4	3	4	3
Instructional Directives	4	4	3	5	3
Feedback	4	4	4	3	3
Teaching Style	4	5	3	3	3
Overall Rating	4	4	3	4	3
Total	29	30	24	28	21

TEEF Scores
Student Teaching #1

	Participant A	Participant B	Participant C	Participant D	Participant E
Information and Demonstrations	5	4	4	3	4
Musical Model	3	3	3	2	3
Flow	5	2	4	1	3
Instructional Directives	5	3	3	1	3
Feedback	4	3	3	2	2
Teaching Style	5	2	3	2	3
Overall Rating	5	3	3	2	3
Total	32	20	23	13	21

TEEF Evaluation
Student Teaching #2

	Participant A	Participant B	Participant C	Participant D	Participant E
Information and Demonstrations	5	4	2	4	4
Musical Model	4	4	3	4	3
Flow	5	2	3	5	4
Instructional Directives	5	2	3	4	3
Feedback	4	3	3	5	2
Teaching Style	5	2	3	3	3
Overall Rating	5	3	3	4	3
Total	33	20	20	29	22

TEEF Scores
Novice Teaching #1

	Participant A	Participant B	Participant C	Participant D	Participant E
Information and Demonstrations	4	4	5	5	4
Musical Model	2	5	4	4	3
Flow	3	3	5	5	4
Instructional Directives	3	3	5	5	3
Feedback	4	2	4	5	4
Teaching Style	4	3	4	4	3
Overall Rating	3	3	4	5	3
Total	23	23	31	33	24

TEEF Evaluation
Novice Teaching #2

	Participant A	Participant B	Participant C	Participant D	Participant E
Information and Demonstrations	3	4	4	5	4
Musical Model	3	3	5	5	3
Flow	3	1	4	5	4
Instructional Directives	3	2	4	4	4
Feedback	2	2	4	5	3
Teaching Style	2	2	4	4	4
Overall Rating	2	2	4	5	4
Total	18	16	29	33	26

TEEF Evaluation
Novice Teaching #3

	Participant A	Participant B	Participant C	Participant D	Participant E
Information and Demonstrations	3	5	4	5	4
Musical Model	4	1	4	4	3
Flow	3	2	5	5	2
Instructional Directives	3	2	4	5	3
Feedback	3	2	4	5	2
Teaching Style	4	2	4	4	3
Overall Rating	3	2	4	5	3
Total	23	16	29	33	20

TEEF Evaluation
Novice Teaching #4

	Participant A	Participant B	Participant C	Participant D	Participant E
Information and Demonstrations	4	3	5	5	2
Musical Model	3	4	4	4	2
Flow	4	3	5	5	2
Instructional Directives	4	3	4	4	2
Feedback	3	4	4	5	2
Teaching Style	4	4	4	4	2
Overall Rating	3	3	4	5	2
Total	25	24	30	32	14

APPENDIX E

TEEF Scores by Item

Information and Demonstrations

	PT #1	PT #2	ST #1	ST #2	NT #1	NT #2	NT #3	NT #4
Participant A	4	4	5	5	4	3	3	4
Participant B	4	4	4	4	4	4	5	3
Participant C	2	4	4	2	5	4	4	5
Participant D	4	5	3	4	5	5	5	5
Participant E	3	3	4	4	4	4	4	2

Musical Model

	PT #1	PT #2	ST #1	ST #2	NT #1	NT #2	NT #3	NT #4
Participant A	4	4	3	4	2	3	4	3
Participant B	5	5	3	4	5	3	1	4
Participant C	2	4	3	3	4	5	4	4
Participant D	2	4	2	4	4	5	4	4
Participant E	3	3	3	3	3	3	3	2

Flow

	PT #1	PT #2	ST #1	ST #2	NT #1	NT #2	NT #3	NT #4
Participant A	4	5	5	5	3	3	3	4
Participant B	2	4	2	2	3	1	2	3
Participant C	2	3	4	3	5	4	5	5
Participant D	4	4	1	5	5	5	5	5
Participant E	3	3	3	4	4	4	2	2

Instructional Directives

	PT #1	PT #2	ST #1	ST #2	NT #1	NT #2	NT #3	NT #4
Participant A	4	4	5	5	3	3	3	4
Participant B	5	4	3	2	3	2	2	3
Participant C	2	3	3	3	5	4	4	4
Participant D	4	5	1	4	5	4	5	4
Participant E	3	3	3	3	3	4	3	2

Feedback

	PT #1	PT #2	ST #1	ST #2	NT #1	NT #2	NT #3	NT #4
Participant A	3	4	4	4	4	2	3	3
Participant B	5	4	3	3	2	2	2	4
Participant C	2	4	3	3	4	4	4	4
Participant D	3	3	2	5	5	5	5	5
Participant E	2	3	2	2	4	3	2	2

Teaching Style

	PT #1	PT #2	ST #1	ST #2	NT #1	NT #2	NT #3	NT #4
Participant A	4	4	5	5	4	2	4	4
Participant B	4	5	2	2	3	2	2	4
Participant C	2	3	3	3	4	4	4	4
Participant D	2	3	2	3	4	4	4	4
Participant E	3	3	3	3	3	4	3	2

Overall Rating

	PT #1	PT #2	ST #1	ST #2	NT #1	NT #2	NT #3	NT #4
Participant A	4	4	5	5	3	2	3	3
Participant B	4	4	3	3	3	2	2	3
Participant C	2	3	3	3	4	4	4	4
Participant D	3	4	2	4	5	5	5	5
Participant E	3	3	3	3	3	4	3	2

APPENDIX F

TEEF Scores by Participant

Participant A

	PT #1	PT #2	ST #1	ST #2	NT #1	NT #2	NT #3	NT #4
Information and Demonstrations	4	4	5	5	4	3	3	4
Musical Model	4	4	3	4	2	3	4	3
Flow	4	5	5	5	3	3	3	4
Instructional Directives	4	4	5	5	3	3	3	4
Feedback	3	4	4	4	4	2	3	3
Teaching Style	4	4	5	5	4	2	4	4
Overall Rating	4	4	5	5	3	2	3	3

Participant B

	PT #1	PT #2	ST #1	ST #2	NT #1	NT #2	NT #3	NT #4
Information and Demonstrations	4	4	4	4	4	4	5	3
Musical Model	5	5	3	4	5	3	1	4
Flow	2	4	2	2	3	1	2	3
Instructional Directives	5	4	3	2	3	2	2	3
Feedback	5	4	3	3	2	2	2	4
Teaching Style	4	5	2	2	3	2	2	4
Overall Rating	4	4	3	3	3	2	2	3

Participant C

	PT #1	PT #2	ST #1	ST #2	NT #1	NT #2	NT #3	NT #4
Information and Demonstrations	2	4	4	2	5	4	4	5
Musical Model	2	4	3	3	4	5	4	4
Flow	2	3	4	3	5	4	5	5
Instructional Directives	2	3	3	3	5	4	4	4
Feedback	2	4	3	3	4	4	4	4
Teaching Style	2	3	3	3	4	4	4	4
Overall Rating	2	3	3	3	4	4	4	4

Participant D

	PT #1	PT #2	ST #1	ST #2	NT #1	NT #2	NT #3	NT #4
Information and Demonstrations	4	5	3	4	5	5	5	5
Musical Model	2	4	2	4	4	5	4	4
Flow	4	4	1	5	5	5	5	5
Instructional Directives	4	5	1	4	5	4	5	4
Feedback	3	3	2	5	5	5	5	5
Teaching Style	2	3	2	3	4	4	4	4
Overall Rating	3	4	2	4	5	5	5	5

Participant E

	PT #1	PT #2	ST #1	ST #2	NT #1	NT #2	NT #3	NT #4
Information and Demonstrations	3	3	4	4	4	4	4	2
Musical Model	3	3	3	3	3	3	3	2
Flow	3	3	3	4	4	4	2	2
Instructional Directives	3	3	3	3	3	4	3	2
Feedback	2	3	2	2	4	3	2	2
Teaching Style	3	3	3	3	3	4	3	2
Overall Rating	3	3	3	3	3	4	3	2

VITA

Alicia Canterbury is a PhD candidate and graduate teaching assistant in music education at the University of Mississippi. Prior to pursuing the PhD degree, she taught elementary general music and directed school choirs for 12 years in Georgia. She was the director of the Bibb County Honor Choir, a district-wide auditioned treble ensemble of students for five years and has been sought as a clinician for regional elementary honor choirs as well. She was the Bibb County Teacher of the Year for the Department of Fine Arts in 2007. She has presented at state-level music educator conferences throughout the South and Midwest and at the American Choral Directors Association Conference on numerous topics related to general music instruction and elementary choral methods. She holds the M.M. degree in Music Education from Texas Tech University and the B.M. degree in Music Education from Mercer University. She is Orff Level 1 and 2 certified from the University of Central Florida. Her research interests include assessment, teacher effectiveness, acquisition of teaching skills, turn of the 20th-century singing schools, and repertoire selection practices. She will be the Assistant Professor in Music Education at Southern Illinois University Edwardsville in August of 2019.