

FINE ARTS TEACHERS' PERCEPTIONS OF STUDENT GROWTH PORTFOLIOS AS AN
ELEMENT OF TEACHER EVALUATION

By

Walter Riley Braem

A Dissertation Submitted in Partial Fulfillment
Of the Requirements for the Degree of
Doctor of Education in Educational Leadership

Austin Peay State University
December 2021

Dissertation Committee:

Dr. Sherri Prosser, Committee Chair

Dr. James Thompson

Dr. Adam Attwood

Marcus Hayes, MFA

FINE ARTS TEACHERS' PERCEPTIONS OF STUDENT GROWTH PORTFOLIOS AS AN
ELEMENT OF TEACHER EVALUATION

Walter Riley Braem

Approved:

DocuSigned by:

Dr. Sherri K Prosser

B0E73C02CC17488...

Dr. Sherri Prosser, Committee Chair

DocuSigned by:

Dr. James M. Thompson I

9171B6AD01BB401

Dr. James Thompson, Committee Member

DocuSigned by:

Dr. Adam Attwood

0EFEEED33D4DD408...

Dr. Adam Attwood, Committee Member

DocuSigned by:

Dr. Marcus Hayes

0119C8B2EB8B4D7...

Marcus Hayes, Committee Member

DocuSigned by:

Dr. Chad Brooks

84E9DFFCA290471...

Dr. Chad Brooks, Associate Provost and Dean, College of Graduate
Studies

Statement of Permission to Use

In presenting this dissertation in partial fulfillment of the requirements for the Doctor of Education degree at Austin Peay State University, I agree that the library shall make it available to borrowers under the rules of the library. Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgement of the source is made. Permissions for extensive quotation or reproduction of this dissertation may be granted by any major professor, or in his/her absence, by the Head of the Interlibrary Services when, in the opinion of either, the proposed use of the material is for scholarly purposes. Any copying or use of the material in this dissertation for financial gain shall not be allowed without any written permission.

Walter Riley Braem

Date: November 12, 2021

DEDICATION

I would like to dedicate this project to my husband and biggest supporter, Rick Talley and our beautiful daughter, Alexis, my grandpa, Richard Dean Ledbetter, who always believed in me (I love and miss you), and two of the most inspirational people that I have known: Tina Mooningham, who taught me how to reach for my dreams and never give up, and Chris Hardesty for showing me how one person could change the world for those he loved.

ACKNOWLEDGEMENTS

I would like to acknowledge the contributions of those who helped me on this journey:

To my husband, Rick Talley and our daughter Alexis for the sacrifices you each made so that I could complete this degree.

To my parents, Walter and Patricia Braem, my sisters Jenny DeArmond and Jamie Vincent, and my mother-in-law, Cindy Talley for their everlasting support along the way.

To my committee, Dr. Sherri Prosser, Marcus Hayes, MFA, Dr. Adam Attwood, and Dr. James Thompson for their encouragement and support throughout this process.

To Noelle Cannon, Shelli Childers, Rachel Funderburk, Ashley Totty, Kim Herndon, and the rest of APSU's first doctoral cohort, it has been an honor and a privilege navigating this road with you.

To my theatre teacher family, Alicia Archuleta, LuEllyn Wilcox, Barbara Wesner, Pollyanna Parker, Meghann Shoulders, Kara Kindall, Rob Silvers, Roy Buchanan, Marc Klein, and all of my TAA folks who inspire me each and every day by their dedication to their students.

To my Northwest family, especially, Jessica Peppard, Kelly Parker, Megan Kuhns, Michael Johnson, and Taylor Kane it is great to be a VIKING!

To my Camp MARC family, Lynne Hardesty, Beth and Philip Tillman, Christy and Reed Rushing, Sabrina Wood, Diana Utley, Meghan Payne, Crystal Curry, Matt Beach, Stephen Bell, Cindy Shemwell, Angie Sturgeon, and Sara Williams.

To my students, past, present, and future, thank you for taking a step out of your comfort zone in my class. Your bravery is inspiring!

To my ASU folks, Stephani Etheridge Woodson, Johnny Saldaña, Roger Bedard, Pam Sterling, Leslie Stellwagen, Kathleen Pennyway, Ashley Hare, and Andy Waldron.

ABSTRACT

The primary purposes of this study are to describe (a) fine arts teachers' perceptions of the portfolio model as a summative evaluation measure and (b) to determine how fine arts teachers use their summative evaluation scores to inform their teaching practice. An additional purpose is to understand how teachers' level of overall effectiveness scores have changed since implementing the Tennessee Fine Arts Student Growth Portfolio Model. Participants of the study were fine arts teachers with at least 4 years of full time teaching experience in their discipline and who had participated in the student portfolio system for a minimum of 2 years. Twenty-eight high school fine arts teachers from five districts across the state completed the survey portion of the study. Of the 28 survey participants, 16 were female and 12 were male; all but three were White; most ($n = 22$) were tenured and 16 of the 28 had at least a master's degree. Six teachers participated in follow-up semistructured interviews. Of the six, four were female and two were male; all were White theatre teachers. This study used a convergent mixed methods design. Participants completed the Perceived Value of Teacher Portfolios Questionnaire developed by Tucker et al. (2003). The survey was completed online and analyzed using descriptive analysis. Semistructured interviews were conducted using videoconferencing software and analyzed manually using in vivo coding. Findings indicate that fine arts teachers had a negative perception of the fairness, feasibility, utility, and accuracy of the student growth portfolio as an element of teacher evaluation. Participants reported issues with artifact manipulation, technology, depth and timeliness of feedback, teacher time requirements and efficiency of the portfolio system, and mistrust and integrity of peer reviewers as factors contributing to their negative perception of the portfolio system. Implications for research include study expansion to include a wider range of grade levels, urban and rural population centers, and relative wealth of the school system.

Implications for practice include providing adequate and ongoing training to teachers and peer reviewers, access to required technologies, and ensuring that peer reviewers are selected based on a record of teaching excellence.

Keywords: student growth portfolio, teacher evaluation, fine arts teachers, attributes of sound evaluation

TABLE OF CONTENTS

CHAPTER I: INTRODUCTION.....	1
Teacher Evaluation in Tennessee.....	2
Race to the Top.....	2
First to the Top.....	3
Tennessee Educator Acceleration Model.....	3
TN Fine Arts Student Growth Portfolio.....	4
Problem of Practice.....	6
Statement of Purpose	7
Overview of Methodology.....	7
Significance of Study.....	8
Key Terms.....	8
CHAPTER II: SYNTHESIS OF THE RESEARCH LITERATURE	11
Theoretical Framework.....	11
Review of Literature	14
Teacher Evaluation in the United States.....	14
Historical Perspectives of Teacher Evaluation	15
Contemporary Perspectives of Teacher Evaluation.....	17
Multi-Measure Summative Evaluation.....	18
Teacher Observations.....	19
Student Growth and Achievement Measures.....	21
Portfolio Teacher Evaluation Models	24
Growth Portfolio Design.....	26

Evidence Credibility	26
Research on Portfolio-Based Teacher Evaluation	28
Attributes of Sound Evaluation.....	30
Sound Portfolio-Based Evaluations	31
Use of Summative Teacher Evaluation Scores	34
State and District Uses	34
Administrator Uses	37
Teacher Uses.....	38
Summary.....	40
CHAPTER III: METHODOLOGY	43
Context.....	43
Research Design.....	46
Participants.....	46
Participant Selection	46
Survey Participants	47
Interview Participants	49
Instrumentation	50
Survey	50
Perceived Value of Teacher Portfolios Questionnaire.....	50
Demographic Information.....	52
Semistructured Interviews	53
Procedure	53
Participant Recruitment	53

Data Collection	54
Survey	54
Semistructured Interviews	54
Data Analysis	55
Survey	55
Semistructured Interviews	55
Trustworthiness and Credibility.....	56
Researcher Reflexivity.....	57
CHAPTER IV: FINDINGS	59
Teacher Perceptions of the Soundness of Portfolio Evaluation.....	59
Fairness	60
Fairness Quantitative Findings	60
Fairness Qualitative Findings	61
Fairness Summary.....	65
Utility.....	65
Utility Quantitative Findings	65
Utility Qualitative Findings	66
Utility Summary.....	68
Feasibility.....	69
Feasibility Quantitative Findings	69
Feasibility Qualitative Findings	70
Feasibility Summary	73
Accuracy	73

Accuracy Quantitative Findings	73
Accuracy Qualitative Findings	74
Accuracy Summary.....	77
Teachers' Perceptions by Academic Discipline and Tenure Status.....	78
Unintended Consequences of Portfolio Evaluation	79
Summary	80
Evaluations Informing Professional Development	80
Utility to Support Professional Development	81
Utility Quantitative Findings	81
Professional Development Quantitative Findings	82
Utility to Support Professional Development Qualitative Findings	83
Professional Development by Academic Discipline and Tenure Status.....	84
Changes in Level of Overall Effectiveness Scores	85
Summary	86
CHAPTER V: DISCUSSION AND RECOMMENDATIONS	88
Discussion	88
Purpose of Teacher Evaluation	88
Attributes of Sound Evaluation.....	89
Utility to Support Professional Growth	90
Variation by Academic Discipline.....	91
Delimitations.....	92
Limitations	92
Implications for Practice and Research.....	94

Implications for Practice	95
Implications for Research	98
REFERENCES	100
APPENDICES	119
Appendix A	119
Appendix B	120
Appendix C	121
Appendix D	122
Appendix E	129
Appendix F	130
Appendix G	132
Appendix H	135
Appendix I	136
Appendix J	137
Appendix K	138
Appendix L	140
Appendix M	141
Appendix N	143
Appendix O	147
Appendix P	148
Appendix Q	149
Appendix R	150

LIST OF TABLES

TABLE 2.1: Examples of Portfolio Types	25
TABLE 3.1: Total Secondary Fine Arts Teachers.....	44
TABLE 3.2: Participating District Information	45
TABLE 3.3: Potential Secondary Fine Arts Teacher Participants.....	45
TABLE 3.4: Secondary Fine Arts Teacher Participants, by Discipline	47
TABLE 3.5: Participant Race, Education Level, and Tenure Status	48
TABLE 3.6: Participant Years of Experience, District Portfolio Use, and Peer Reviewer Status	49
TABLE 3.7: Interview Participant Demographics.....	50
TABLE 4.1: Frequency Ratings of Perceived Value of Fairness Survey Items.....	60
TABLE 4.2: Survey Participant’s Perception of the Portfolio’s Fairness	61
TABLE 4.3: Frequency Ratings of Perceived Value for Utility Survey Items	65
TABLE 4.4: Survey Participant’s Perception of the Portfolio’s Utility	67
TABLE 4.5: Frequency Ratings of Perceived Value for Feasibility Survey Items	69
TABLE 4.6: Survey Participant’s Perception of the Portfolio’s Feasibility.....	70
TABLE 4.7: Frequency Ratings of Perceived Value for Accuracy Survey Items.....	73
TABLE 4.8: Survey Participant’s Perception of the Portfolio’s Accuracy	75
TABLE 4.9: Frequency Ratings of Perceived Value for Utility Survey Items	81
TABLE 4.10: Frequency Ratings of Perceived Value for Professional Development Items.....	82
TABLE 4.11: Changes in Level of Overall Effectiveness Scores Year 1-Year 2	85
TABLE N.1: Frequency Ratings for Fairness Survey Items, by Discipline.....	143
TABLE N.2: Frequency Ratings for Utility Survey Items, by Discipline.....	144
TABLE N.3: Frequency Ratings for Feasibility Survey Items, by Discipline	145

TABLE N.4: Frequency Ratings for Accuracy Survey Items, by Discipline	146
TABLE O: Frequency Ratings for All Survey Items, by Tenure Status	147
TABLE P: Professional Development Items, by Discipline.....	148
TABLE Q: Professional Development Items, by Tenure Status	149
TABLE R: Self-Reported Changes in Level of Overall Effectiveness Year 1-Year 2.....	150

LIST OF FIGURES

FIGURE 1: Reciprocal Determinism.....13

FIGURE 2: State-Level Use of Summative Teacher Evaluation Scores35

FIGURE 3: District-Level Use of Summative Teacher Evaluation Scores36

FIGURE 4: Administrator Use of Summative Teacher Evaluation Scores37

Chapter I

Introduction

The process of evaluating teachers is not unlike that of evaluating students, as both are assessed formatively and summatively. Formative assessment is an assessment for learning or used to support and inform future practice, whereas summative assessment is an assessment of learning meant to measure or verify learning outcomes (Chappuis et al., 2012). In teacher evaluation, classroom observations serve as formative assessments of teacher praxis (Tobin et al., 2015). For teacher assessment to be productive, evaluators must accurately assess what they see in the classroom, provide specific and meaningful feedback, and engage teachers in conversation about their practice (Danielson, 2011). Summative teacher evaluation scores are comprised of a combination of observation scores and measures of student growth and achievement (Derrington & Martinez, 2019).

Some of the earliest methods of evaluating teachers considered the personality characteristics of the teacher to be an important indicator of efficiency. Kratz, an educator in Iowa, conducted a study in 1896 to create a list of characteristics that he felt would serve as benchmarks by which other teachers could be evaluated (McNergney & Imig, 2003; Medley, 1979). In 1972, economist Eric Hanushek introduced the concept of a value-added measure (VAM) as a form of teacher accountability (Green, 2014). The VAM, one of many student growth measures, attempts to quantify the positive or negative effect individual teachers have on their students.

To calculate a tested teacher's student growth measure (i.e., VAM) in Tennessee, a complex statistical analysis is run, which utilizes end of course exam scores to calculate the measure. The student growth measure used for teachers of non-tested subjects is their school's

overall growth score. To provide teachers of non-tested subjects more control over their student growth measure, a group of pioneering fine arts teachers met in the spring of 2011 to develop what would become the Tennessee Fine Arts Student Growth Portfolio Model. Teachers from Memphis City Schools developed and piloted the program during the 2011-2012 school year. The following year, the Tennessee Department of Education (TDOE) approved the portfolio model as an optional, alternative growth measure for fine arts teachers. To understand the significance of this event, one must understand the history of educational reform in Tennessee and what led to the creation of the Tennessee Educator Acceleration Model (TEAM). Eventually, the Tennessee Fine Arts Student Growth Portfolio Model was chosen to evaluate fine arts educators.

Teacher Evaluation in Tennessee

Changes to teacher evaluation in Tennessee occurred after legislative changes that occurred at the state and federal levels of education. Changes to the state's teacher evaluation system resulted from the Race to the Top competition (Hamilton, 2010). The creation and implementation of the TEAM evaluation process and the student growth portfolio were prompted by the passage of the American Recovery and Reinvestment Act of 2009 and, subsequently, the Tennessee First to the Top Act of 2010.

Race to the Top

Part of the American Recovery and Reinvestment Act of 2009, Race to the Top (RTTT) was a competitive grant program that started in 2009 under the Obama administration. The criteria to be awarded the grant included the requirement that states "adopt common standards and assessments, create better student data systems, create teacher evaluation and retention policies, and adopt selected school turnaround strategies" (Brimley et al., 2016, p. 207). In the

initial years of the program, almost every state (i.e., 47) applied for the grant. Due to the influx of states applying for funding, the Department of Education found it necessary to create additional criteria states had to meet before applying for the grant. Of the requirements, linking student test scores to teacher evaluations and allowing for an unlimited number of charter schools in the state were the most impactful additions to the application requirements.

Tennessee was one of the first states to be awarded the RTTT grant. For its part, Tennessee would receive "\$501.8 million over four years...with half of the Race to the Top funds - \$250.9 million – distributed directly to local school districts using the existing Title I formula" (Riley, 2010, p. 1). Key provisions in Tennessee's RTTT application also allowed each school district to participate voluntarily (Riley, 2010). However, participation was on an all or none basis—there would be no exceptions to any of the expectations. Of all of the policy outcomes resulting from RTTT, the inclusion of student test scores into teachers' summative evaluation scores was the most significant change to the status quo.

First to the Top

As a result of RTTT, Tennessee legislators adopted the First to the Top Act of 2010. This piece of legislation had a direct impact on teacher evaluation scores. Before 2010, the Tennessee Value Added Assessment System (TVAAS) scores were only used for diagnostic purposes. However, then-governor Phil Bredesen and the Tennessee General Assembly made changes to the policies that would allow them to use TVAAS scores in teacher evaluations (SCORE, 2010). This change was made to align Tennessee with the RTTT accountability requirements.

Tennessee Educator Acceleration Model

The use of TVAAS scores was not the only change to teacher evaluation due to RTTT. A provision in the First to the Top legislation established the Tennessee Educator Acceleration

Model (TEAM) evaluation process. TEAM is a comprehensive evaluation system developed by the TDOE for teachers and administrators. The crux of the TEAM evaluation process is the combination of data collected to calculate an educator's level of overall effectiveness. TEAM combines teacher observations with data from student growth and achievement measures to evaluate teachers. TDOE describes the summative evaluation process as follows:

The goal of the TEAM evaluation process is to provide educators with a model that helps them continuously improve their practice. A complete picture of what goes on in the classroom is essential to driving educator improvement; therefore, we want to look at how teachers deliver instruction and what students learn from those lessons. By using observations and data together, TEAM allows teachers and school leaders to have an ongoing dialogue about how what happens in the classroom impacts student performance. Ultimately, growth in a teacher's skills leads to growth in student achievement. Like the reflective practices the TEAM observation system promotes for educators, the Tennessee Department of Education is committed to reflecting on and refining the observation system through feedback loops and careful study over time (Tennessee Department of Education, 2021, para. 1).

TN Fine Arts Student Growth Portfolio

As outlined in Tennessee State Board of Education Teacher and Administrator Evaluation Policy 5.201, student growth portfolio models were approved to provide teachers of non-tested subjects with an individual growth measure. A portfolio is a collection of student work, aligned to state standards, selected and submitted by a teacher to demonstrate student growth (Tennessee Department of Education, 2020b). In 2016 the state of Tennessee adopted new fine arts standards, which were organized into four overarching domains:

- Perform (music, dance, theatre), Present (visual arts), Produce (media arts);
- Create;
- Respond; and
- Connect (Tennessee Department of Education, 2020b).

Within each of the domains are two or three foundational skills (e.g., analyze, generate, interpret, synthesis) for a total of 11 foundational skills considered to be requisite behaviors for artistic growth, process, and literacy (Tennessee Department of Education, 2020b). For example, within the respond domain, performing arts students would be expected to interpret intent and meaning in artistic work or apply criteria to evaluate artistic work (Tennessee Department of Education, 2020b). Within the connect domain, performing arts students would be expected to synthesize and relate prior knowledge and personal experiences to artistic works, or relate artistic ideas and works with societal, cultural, and historical context (Tennessee Department of Education, 2020b).

Student growth portfolios contain four evidence collections representing at least two of the four domains. An evidence collection is a sample of student work on one state standard (Tennessee Department of Education, 2016). Each evidence collection must contain artifacts (i.e., student work) from two points in time and can only be aligned to one standard (Tennessee Department of Education, 2020b). Although artifacts used to comprise evidence collections may be used to assess multiple standards or domains in the classroom, peer reviewers can only score evidence collections for one standard within a single domain (Tennessee Department of Education, 2020b). Teachers submit a portfolio that is representative of their teaching assignment and demonstrates student growth across varying student populations. For example, if a teacher

has four Theatre I classes and two Advanced Theatre classes at least two of the evidence collections should come from their Theatre I students.

Fine arts teachers have the option of using two different types of sampling: differentiated and group. A differentiated sample consists of artifacts from multiple learning levels (i.e., emerging, proficient, and advanced) whereas a group sample represents growth from an ensemble or group (Tennessee Department of Education, 2020b). It is important to note that group samples are only available for dance, media arts, music, and theatre teachers.

Completed portfolios are submitted to peer reviewers. A peer reviewer is an educator who has been rated as highly effective and is trained to review portfolios. Peer reviewers utilize state-provided fine arts discipline-specific scoring guides and rubrics, which are aligned to TN fine arts standards (Tennessee Department of Education, 2020b). See Appendix A for the fine arts scoring rubric indicators and Appendix B for an overview of the structure of a Fine Arts Student Growth Portfolio.

Problem of Practice

Educator evaluations can be a source of anxiety for many teachers (Hibler & Snyder, 2015). Variations in evaluation methods and models lead to inconsistencies in the perceived effectiveness of educators (Darling-Hammond et al., 2012; Gitomer & Bell, 2016; Martinez et al., 2016). Fear of low evaluation scores can also negatively impact curriculum and instruction (Fromer & Michelli, 2015; Hart et al., 2014). As a result, fear of low evaluation scores may cause educators to shy away from new and innovative instructional strategies (Hart et al., 2014). Many teachers also feel that educator evaluations are unfair because they have little or no control over student scores used to calculate their teaching effectiveness (Darling-Hammond et al., 2012). Interrater reliability is another ongoing concern of teachers regarding evaluation scores

(Gillespie et al., 1996; Gitomer & Bell, 2016). Fine arts teachers across several Tennessee counties are required to use portfolio-based assessments that may or may not be reflective of their teaching effectiveness (Wilson, 2020).

Statement of Purpose

The primary purposes of this study are to describe (a) fine arts teachers' perceptions of the portfolio model as a summative evaluation measure and (b) to determine how fine arts teachers use their summative evaluation scores to inform their teaching practice. An additional purpose is to understand how teachers' level of overall effectiveness scores have changed since implementing the Tennessee Fine Arts Student Growth Portfolio Model. The research questions guiding this study are:

1. What are the perceptions of fine arts teachers regarding the use of student growth portfolios as an element of teacher evaluation?
2. How do fine arts teachers use their summative evaluation scores to inform their teaching practice?

Overview of Methodology

This convergent mixed methods study will address fine arts teachers' perception of the Tennessee Fine Arts Student Growth Portfolio as a measure of teacher evaluation. Convergent design is a single-phase design that utilizes quantitative and qualitative data. In this design, "a researcher collects both quantitative and qualitative data, analyzes them separately, and then compares the results to see if the findings confirm or disconfirm each other" (Creswell & Creswell, 2018, p.217). The primary assumption of a convergent design is that qualitative and quantitative data provide "different types of information" (Creswell & Creswell, 2018, p. 217).

Significance of Study

The TEAM evaluation process is used to calculate a teacher's level of overall effectiveness. As outlined in the Tennessee State Board of Education (2018) Teacher and Administrator Evaluation Policy 5.201, a teacher's level of overall effectiveness is comprised of a combination of three different measures: teacher observations (50%), student growth data (35%), and a student achievement measure (15%). These scores are calculated and render a level of overall effectiveness score of 1 (*significantly below expectations*) to 5 (*significantly above expectations*). As outlined in the Tennessee Teacher Tenure Law (2012), to be eligible for tenure, a teacher must have "received evaluations demonstrating an overall performance effectiveness level of 'above expectations' or 'significantly above expectations' provided by the evaluation guidelines adopted by the state board of education pursuant to § 49-1-302, during the last two (2) years of their probationary period." Additionally, teachers who have tenure can be terminated if their level of overall effectiveness score falls below a 3 (*at expectations*) for more than 2 consecutive years (Tennessee Teacher Tenure Law, 2012).

Key Terms

1. **Accuracy:** the evaluation system produces sound information about performance (Joint Committee on Standards for Educational Evaluation, 1988)
2. **Artifact:** the products and by-products of teaching that demonstrate a teacher's performance (Tucker et al., 2002)
3. **Evidence Credibility:** the believability of any teacher collected evidence that demonstrates student growth (Popham, 2013a, p. 36)
4. **Feasibility:** the evaluation system offers an efficient, easy to use, and viable means of evaluation (Joint Committee on Standards for Educational Evaluation, 1988)

5. **Level of Overall Effectiveness:** a scale score created by combining teacher observation scores, student growth data, and student achievement data (Tennessee Department of Education, 2021c)
6. **Peer Reviewer:** a highly effective educator who is trained to review portfolios in an objective manner based on the content-specific scoring guide. Peer reviewers must also submit portfolios (Tennessee Department of Education, 2016)
7. **Professional Learning:** a program of planned activities designed to increase the competencies needed by all licensed personnel in the performance of their professional responsibilities (Tennessee Department of Education, 2021a)
8. **Propriety:** the evaluation system considers and protects the legal rights of those affected by the evaluation. (Joint Committee on Standards for Educational Evaluation, 1988) For the purposes of clarity for participants, I will refer to this construct as fairness in this study
9. **Student Achievement Measure:** a measure examining qualitative indicators (i.e., grades, industry certifications, ACT/SAT composite scores, graduation rate) that represent students' skills and knowledge and the efficiency of the educational institution (e.g., teacher; Ballafkih & Van Middelkoop, 2019)
10. **Student Growth Measure:** the change in student achievement for an individual student between two or more points in time (USDE, 2020)
11. **Student Growth Portfolio:** a collection of student work, aligned to state standards, selected and submitted by a teacher to demonstrate student growth (Tennessee Department of Education, 2021b)

12. **Utility:** the evaluation system provides timely and meaningful information in the evaluation process (Joint Committee on Standards for Educational Evaluation, 1988)

Chapter II

Synthesis of the Research Literature

A review of research literature was conducted to guide the development of this study. This chapter begins with a historical overview of teacher evaluation in the United States starting with the work of Kratz (1896). Next, contemporary perspectives of teacher evaluation are explored. The use of multi-measure teacher evaluation and portfolio-based evaluation models are discussed at length. The next section discusses the four attributes of sound evaluation: accuracy, utility, feasibility, and fairness, developed by the Joint Committee on Standards for Educational Evaluation (1988). The review of literature then presents research on portfolio-based teacher evaluation models that have been evaluated using the four attributes of sound evaluation. The final section of this chapter discusses the various uses of summative teacher evaluation scores by key stakeholders (i.e., state, district, administrator, and teacher).

Theoretical Framework

Albert Bandura's social cognitive theory will serve as the theoretical lens in which the literature is viewed. First referred to as social learning theory (Bandura, 1977), social cognitive theory evolved from Bandura's earlier work on learning and behavioral change. Social cognitive theory consists of six constructs:

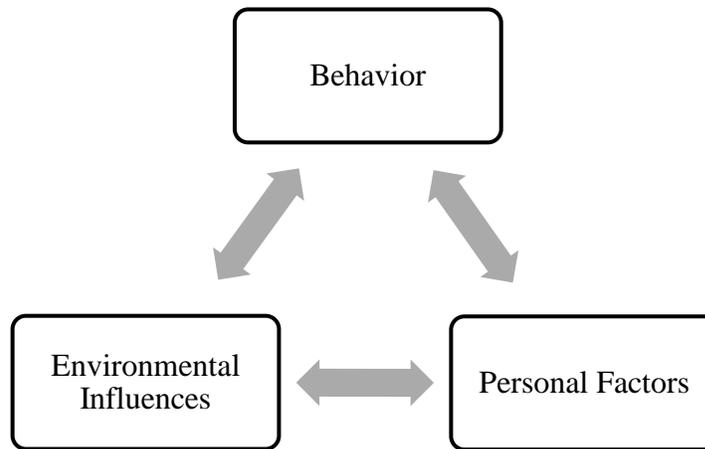
1. Reciprocal determinism: the reciprocal interaction between individual, environment, and behavior.
2. Behavioral capability: the individual's ability to perform a behavior through essential knowledge and skills.
3. Observational learning: an individual can observe a behavior performed by another and reproduce the behavior.

4. Reinforcements: the internal and external responses to a behavior that impact the likelihood the behavior will or will not be repeated.
5. Expectations: the anticipated consequences of a behavior .
6. Self-efficacy: an individual's belief that they can successfully perform a behavior.

(Bandura, 1977; 1986)

Social cognitive theory posits that learning occurs in a social setting and is a complex reciprocal interaction of the individual, environment, and behavior (Bandura, 1977; 1986). The continuous interaction between these three factors is known as reciprocal determinism, one of the primary constructs of social cognitive theory (Bandura, 1977; see Figure 1). Observational learning and self-efficacy are also key components of social cognitive theory (Bandura, 1986). Although all six constructs of social cognitive theory can apply to teacher evaluation, this study will utilize reciprocal determinism, observational learning, and self-efficacy as the primary theoretical elements that guide the literature review.

Teachers' perception of the different elements of summative evaluation can negatively impact their work in the classroom (Hart et al., 2014). Bandura (1986) posits that inaccurate or false beliefs trigger avoidant behavior keeping individuals out of touch with reality. Given that evaluations are a source of anxiety for teachers (Hibler & Snyder, 2015), and that fear of low evaluation scores may cause teachers to shy away from new and innovative instructional strategies (Hart et al., 2015), the perceived consequences associated with teacher evaluation serve as an example of reciprocal determinism. In the case of teacher evaluations, "consequences serve as an unarticulated way of informing performers what they must do to gain beneficial outcomes and to avoid punishing ones" this phenomenon is an example of observational learning (Bandura, 1986, p. 192).

Figure 1*Reciprocal Determinism (Bandura, 1977)*

Note: Reciprocal determinism, a key concept of social cognitive theory (Bandura, 1977, 1986), suggests that one's behavior is influenced by and influences one's personal factors (e.g., cognitive skills, attitudes) and social environment (e.g., feedback, reinforcement). Reciprocal determinism adapted from "Social Learning Theory," 1977 and "Social Foundations of Thought and Action: A Social Cognitive Theory" 1986, Albert Bandura. Copyright 1977; 1986 by Albert Bandura.

Self-efficacy is an individual's beliefs that they can perform tasks that can affect their lives (Bandura, 1994). Self-efficacy beliefs determine how individuals feel, think, motivate themselves, and behave (Bandura, 1994). Individuals with a strong sense of self-efficacy set challenging goals and maintain commitment to a task and can also recover quickly after a setback or failure (Bandura, 1994). Conversely, individuals with a low sense of self-efficacy do not set challenging goals and maintain low levels of commitment to a task and recover slowly after a setback or failure (Bandura, 1994).

An individual's perceived self-efficacy can affect the choices they make, their motivation to accomplish goals, the quality in which they perform tasks, their reaction to setbacks, and their susceptibility to depression and stress (Bandura, 1994). Self-efficacy plays a vital role in processes of teaching and learning (Bandura, 1997; Skaalvick & Skaalvick, 2007; Tschannen-Moran & Hoy, 2001). Teacher self-efficacy is important because of the implications it can have on the educational environment (de Oliveira Fernandez et al., 2016). Bandura (1993) argues that "teachers' beliefs in their personal efficacy to motivate and promote learning affect the types of learning environments they create and the level of academic progress their students achieve" (p. 117). Given the impact summative evaluation scores can have on teachers, the constructs of reciprocal determinism, observational learning, and self-efficacy are a suitable framework in which to examine how teacher evaluation is presented within the literature.

Review of Literature

Teacher Evaluation in the United States

Teacher evaluation in the United States consists of formative and summative elements of evaluation (Popham, 2013b). Scriven (1967) distinguished the differences between formative and summative evaluation in his essay, *The Methodology of Evaluation*. Formative evaluation is used to gather information that could improve an educational program whereas summative evaluation was used to determine if an educational program would continue (Scriven, 1967). The distinction between the two types of evaluation was soon adopted by those whose focus was to evaluate teachers (Popham, 2013b). Today, "formative teacher evaluation is focused on improvement" whereas "summative teacher evaluation is focused on removal or rewards" (Popham, 2013b, p. 20). To understand the current state of teacher evaluation, context must be given to the origins of teacher evaluation in the United States.

Historical Perspectives of Teacher Evaluation

Even though teaching and learning has existed since the dawn of human history, formal education emerged in tandem with early human civilizations more than 5,000 years ago (Eskelson, 2020). The first public school in the United States, the Boston Latin School, was founded in 1635 (BLS History, 2020). At this point in time, education was not seen as a professional discipline or field of study (Marzano et al., 2011; Tracy, 1995) and decisions about the hiring, firing, and evaluation of teachers were largely left to local clergy (Marzano et al., 2011). The rise of the industrial revolution and the urbanization of cities during the 1800s had a direct impact on the structure of public education (Marzano et al., 2011; Carl, 2010; Katz, 1987) and school leaders replaced clergy as the evaluators of teachers (Marzano et al., 2011).

Just as education is ever-evolving, so too are the methods used to evaluate teachers on their effectiveness as educators. Some of the earliest methods of evaluating teachers considered the personality characteristics of the teacher to be an important indicator of efficacy. In 1896, Kratz created a list of characteristics that could serve as benchmarks by which his fellow teachers could be evaluated including: personal appearance, patience, politeness, neatness, and being good or kind (Kratz, 1896). Although ratings teachers based personal characteristics were the norm in the mid-1900s (Barr, 1948), some researchers argued that the opinions of a supervisor had no bearing on student outcomes (McNergney & Imig, 2003).

From the late 1800s to the early 1900s, education was governed by two predominant views: those of John Dewey (1938) and Frederick Taylor (1911). Dewey (1938) asserted that, “growth, not only physically but intellectually and morally, is one exemplification of the principle of continuity for educational learning” (p. 5). Conversely, Taylor, who was an industrial engineer, developed the principles of scientific management that reorganized industrial

systems around quantifiable and measurable goals (Stoller, 2015). Taylorism, as it came to be known, required school administrators to assume more responsibilities which, in turn, reduced the responsibilities and decision making required by teachers (Marzano, 2012; Spring, 2010).

Psychologist Edward Thorndike and educator Ellwood Cubberley led the paradigm shift that measurement was a better scientific evaluation tool for education (Marzano et al., 2011). Thorndike (1910) asserted that there are “scales for everything in human nature” (p. 4). The work and theories of Thorndike were instrumental in integrating Taylor’s principals of scientific management into educational theory and practice (Stroller, 2015; Marzano et al., 2011). Cubberley applied the principles outlined by Taylor to K-12 education (Marzano et al., 2011). Cubberley (1929) felt that Taylor’s principles of factory management could be applied to managing schools, stating that:

Our schools are, in a sense, factories in which the raw products (children) are to be shaped and fashioned into products to meet the various demands of life. The specifications for manufacturing come from the demands of twentieth century civilization and is the business of the school to build its pupils according to the specifications laid down. (p. 338)

Much like the lists of characteristics that Kratz developed, Cubberley utilized a set of principles that guided school administrators during evaluations (Marzano et al., 2011). The scientific approach that Cubberley took towards supervision stressed that the evaluation of teachers should be based on the measurement and evaluation of data to ensure teacher productivity (Marzano et al., 2011). Cubberley’s approach, applied a one-size-fits-all approach to teacher evaluation. However, high school administrator, William Wetzel (1929) advocated for the use of student learning to determine teacher effectiveness. Unlike his contemporaries, Wetzel

(1929/1932) distanced himself from referring to schools as factories. Wetzel (1929) proposed three elements that defined scientific supervision: the use of aptitude tests for every student, clear and measurable course objectives, and the use of reliable measures of student learning.

The post-World War II period of educational evaluation saw a shift from the stringent role of Taylorism in evaluation to an emphasis on the teacher as an individual (Danielson & McGreal, 2000; Marzano et al., 2011). During this period, school administrators focused more on evaluating teachers as a means of professional growth (Marzano et al., 2011). As a result, the focus of teacher evaluation shifted to teacher instructional practices (Coleman, 1945). From the late 1950s on, clinical supervision was commonplace on the educational landscape (Marzano et al., 2011). By the early 1980s, upwards of 90% of administrators utilized some type of clinical supervision model to evaluate their teaching staff (Bruce & Hoehn, 1980).

At the rise of the 21st century, a shift in teacher evaluation began to move away from teacher practice and towards student achievement (Marzano et al., 2011). Researchers advocated for the use of student achievement as an element of teacher evaluation (Tucker & Strong, 2005). Critics of earlier evaluation models, Toch and Rothman (2008) characterized them as superficial and argued that they did not directly address student learning or quality of instruction.

Contemporary Perspectives of Teacher Evaluation

Beginning with the passage of the Race to the Top Act of 2009 and the Elementary and Secondary Education Act Flexibility Program in 2011, teacher evaluation programs in the United States have changed significantly (Popham, 2013b). Race to the Top Grants and waivers under the Elementary and Secondary Education Act Flexibility Program were only awarded to states whose teacher evaluation systems:

- Were used for continuous improvement of instruction

- Consisted of a differentiated performance scale (e.g., emerging, proficient, advanced)
- Included multiple measures to determine performance levels with student growth being a significant factor
- Evaluated both teachers and administrators on a regular basis
- Provided feedback to inform professional development
- Were used to inform personnel decisions (USDE, 2012)

Part of the American Recovery and Reinvestment Act of 2009, Race to the Top (RTTT) was a competitive grant program that started in 2009 under the Obama administration. Funding criteria required that states “adopt common standards and assessments, create better student data systems, create teacher evaluation and retention policies, and adopt selected school turnaround strategies” (Brimley et al., 2016, p. 207). To encourage states to reform their teacher evaluation systems, the federal government tied RTTT funding with reformed teacher evaluation systems (Close et al., 2019). States that included higher stakes (i.e., student growth measures) in their reformed evaluation systems received more RTTT funding (Close et al., 2019).

Multi-Measure Summative Evaluation. Evaluating teachers is difficult because teaching is a multifaceted activity (Airasian & Gullickson, 1997). The environments that teachers work in are extremely complex and diverse (Livingston, 2017). As such, it is illogical to evaluate teachers using a single source of data (Mehrens, 1989). In contrast to single-measure summative evaluation (i.e., value-added models), multi-measure summative evaluation models use more than one data source to evaluate teacher effectiveness (Hanover Research, 2012). Quality assurance and professional development are the two primary purposes of teacher evaluation (Danielson & McGreal, 2000). The measurement (i.e., quality assurance) and development of teachers are important aspects of multi-measure teacher evaluation systems currently in use

(Connally & Tooley, 2016; Danielson, 2010; Marzano, 2012; Robinson, n.d.). Summative teacher evaluation scores are comprised of a combination of observation scores and measures of student growth and achievement (Derrington & Martinez, 2019).

Teacher Observations. The most commonly used component of teacher evaluation systems, classroom observations, are required in all 50 states (Steinberg & Donaldson, 2015). Classroom observations involve an evaluator assessing a teacher's instructional practices against a set of performance standards (Donaldson & Papay, 2014). Evaluators typically utilize a rubric as the observation instrument (Donaldson & Papay, 2014). Post-observation feedback is a critical element of the teacher evaluation process (Danielson & McGreal, 2000). In terms of social cognitive theory (Bandura, 1986), post-observational feedback serves as an environmental factor that has a reciprocal relationship between a teacher's personal factors (e.g., self-efficacy) and their behavior (e.g., teacher performance). Effective feedback is actionable, goal-referenced, tangible and transparent, personalized, timely, ongoing, and consistent (Wiggins, 2012).

The primary purpose of any rubric is to assess performance (Brookhart, 2013). Commonplace in early childhood education (Rezaei & Lovorn, 2010) through higher education (Sadler, 2009) instruction, rubrics are also used in 36 states (Close et al., 2019) as an element of teacher evaluation (Papay, 2012) during classroom observations. A performance assessment (i.e., teacher evaluation) consists of two elements: a task and scoring criteria or rubric (Humphry & Heldsinger, 2014; Perlman, 2003). As an evaluation tool, rubrics have three distinct features: evaluative criteria, quality definitions (i.e., gradations), and a scoring strategy (Popham, 2000). Most rubrics utilize a matrix design where each criterion has the same number of gradations of quality (Humphry & Heldsinger, 2014; Walvoord, 2010). Rubrics are popular in education because they are easy to create and use (Andrade, 2005).

The two most discussed performance assessment rubrics in literature are analytical and holistic rubrics (Moskal, 2002). Analytical rubrics break down a performance into different elements that are evaluated separately on different scales (Brookhart, 2013; Moskal, 2002). With holistic rubrics, as the name suggests, all of the elements that make up a task are evaluated simultaneously using a single scale (Brookhart, 2013; Moskal, 2002). The criteria and descriptions of a general rubric can be generalized to different performance tasks within the same learning outcome, for example, writing or problem solving (Brookhart, 2013). Conversely, task-specific rubrics cannot be generalized to other tasks because they are specific to the performance task they are being used to evaluate (Brookhart, 2013).

The two most common rubric-based frameworks for teacher evaluations are the Marzano and the Danielson Frameworks (Close et al., 2019). Both the Marzano Causal Teacher Evaluation Model (Marzano & Toth, 2013) and Danielson Framework for Teaching (Danielson, 2012; Danielson & McGreal, 2000) are organized with indicators in four domains. These domains encompass the areas of planning, instruction, environment, and professional responsibility (Danielson, 2012; Danielson & McGreal, 2000; Marzano & Toth, 2013). Rubric-based models help administrators provide detailed feedback and more specific ratings on teacher performance (Hart et al., 2014).

Some researchers believe that an effective rubric must be clear and specific to the skill it is assessing (Chang & Wu, 2012; Humphry & Heldsinger, 2014; Moskal, 2002; Parkes et al., 2015; Silveira, 2013), whereas others believe that overly specific rubrics are essentially worthless (Popham, 1997). Additionally, numerous researchers have identified issues of validity and reliability concerning rubric-based evaluation (see Andrade, 2005; Bresciani et al., 2009; Humphry & Heldsinger, 2014; Jonsson & Svingby, 2007; Moskal, 2002; Moskal & Leydens,

2000; Szafran, 2017). Issues of validity include: structural design of the rubric (Humphry & Heldsinger, 2014); consequential evidence (Moskal & Leydens, 2000); and language of the rubric (Andrade, 2005). Whereas issues of reliability revolve around interrater and intrarater reliability (Andrade, 2005; Bresciani et al., 2009; Jonsson & Svingby, 2007; Moskal & Leydens, 2000; Szafran, 2017).

Student Growth and Achievement Measures. In many states, student growth and achievement measures are used to evaluate both teachers and principals on their overall effectiveness (Ross & Walsh, 2019). As a result of RTTT, states worked quickly to incorporate student growth measures into their evaluation systems (Ross & Walsh, 2019). Student growth is the change in student achievement for an individual student between two or more points in time (USDE, 2020). For example, only 15 states required an objective measure of student growth as part of their teacher evaluation system in 2009, compared to 43 states in 2015 (Ross & Walsh, 2019).

Including student growth measures as part of summative evaluation is considered highly punitive by teachers across the country (Close et al., 2019). If test scores failed to show growth over time, teachers' professional files could be flagged as "underperforming," or they could be denied merit-based pay (Close et al., 2019), denied tenure, or face termination (Close et al., 2019; NCTQ, 2016). Garet et al. (2017) and McCullough et al. (2015) explain that growth scores are often difficult to understand, are not reported until after students have moved on, and can fluctuate annually as examples of why growth scores are not particularly useful. In alignment with reciprocal determinism (Bandura's (1977, 1986), the environmental factors listed above can have a direct impact on teacher behavior in the classroom. States and districts across the country

use a variety of student growth measures for teacher evaluation (e.g., district level assessments, portfolios, student learning objectives, and value-added measures; NCTQ, 2019).

First introduced by economist Eric Hanushek in 1972 as a form of teacher accountability, the value-added measure (VAM) serves as a measure of student growth (Green, 2014). The VAM, one of many possible measures of student growth, attempts to quantify the positive or negative effect individual teachers have on their students (Darling-Hammond et al., 2012; Green, 2014; Rothstein, 2016). To calculate a teacher's VAM, a complex statistical analysis is run, which almost always includes a standardized test as part of the calculation (Rothstein, 2016). The use of the VAM to evaluate teacher effectiveness is based on the assumption that student achievement is influenced solely by an individual teacher and does not take other factors, such as class size, individual student needs and abilities, and home environment into account (Darling-Hammond et al., 2012).

Although VAMs still serves as the so-called unbiased measure of student growth, what constitutes student growth has changed in both policy and practice (Close et al., 2019). The shift in policy and practice occurred with the passing of the Every Student Succeeds Act (ESSA) in 2015. For example, eight states and the District of Columbia have dropped the objective student growth measure requirement from their teacher evaluation systems (Ross & Walsh, 2019). Currently, only 15 states continued to use VAMs as part of their teacher evaluation systems (Close et al., 2019). In contrast, 28 states utilize an alternative objective student growth measure even though the remaining states do not incorporate an objective measure of student growth into their evaluation systems (Close et al., 2019).

The third element of teachers' summative evaluation scores is a student achievement measure. Student achievement is measured by examining qualitative indicators (i.e., grades,

industry certifications, ACT/SAT composite scores, graduation rate) that represent students' skills and knowledge and the "efficiency of the educational institution" (e.g., teacher; Ballafkih & Van Middelkoop, 2019, p. 46). These indicators are then used to communicate the level of mastery a student has demonstrated (Ballafkih & Van Middelkoop, 2019; Nitko & Brookhart, 2010).

Student achievement is a "multifaceted construct dictated by the tasks of education" (Ballafkih & Van Middelkoop, 2019, p. 47). In the broadest sense of the term, student achievement includes soft skills and personal growth, in addition to qualitative aspects of student engagement in educational activities (e.g., student satisfaction, student knowledge acquisition, post-educational performance; Ballafkih & Van Middelkoop, 2019). In contrast, the narrow definition of student achievement relates to hard skills and knowledge measured through standardized tests; in short, "student achievement is related to the student's ability to reproduce knowledge and tasks" (Ballafkih & Van Middelkoop, 2019, p. 46).

Similar to the incorporation of student growth measures as an element of teacher evaluation, the inclusion of a measure of student achievement has also changed drastically over time. In 2009, 15 states required "student achievement to be included in teacher evaluation" (Doherty & Jacobs, 2015, p. 2). After RTTT was adopted all but 7 states required student achievement to be included as an element of teacher evaluation (Doherty & Jacobs, 2015).

The quality assurance characteristics of summative evaluation are favored by legislators and policymakers because summative evaluation is perceived as a measure of accountability (Danielson & McGreal, 2000). Summative evaluation provides administrators with a wide variety of information including, but not limited to, determining the effectiveness of instructional

activities, data to determine employment decisions (e.g., continuation, tenure, promotion), and to identify strengths and weaknesses in an instructor's performance (Tobin et al., 2015).

Portfolio Teacher Evaluation Models. Portfolios have been defined by researchers based on their purpose and content (Birgin & Baki, 2007). A portfolio is “a fusion of process and product. It is the process of reflection, selection, rationalization, and evaluation, together with the product of those processes” (Winsor & Ellefson, 1995, p. 68). The use of a portfolio as an evaluative method in education first appeared in 1979 when educator Eugene Williams Sr. used the term “audiovisual interview portfolio” in an article (deMontmollin, 2018). The purpose of the audiovisual interview portfolio was to provide potential employers with a sample of a preservice teacher's competencies and performance skills. These portfolios included audiovisual components, writing samples, and demonstration slides of the candidates' work (Williams, 1981).

In the 1980s, portfolios were used to support preservice teacher education (English & Lachlan-Haché, 2015; Goldberg, 2011). By the 1990s school districts around the country began to utilize portfolios as an element of teacher evaluation (English & Lachlan-Haché, 2015; Wolf & Dietz, 1998). Within 20 years of the portfolio's appearance, districts across the country were using portfolios as an element of teacher evaluation and to inform professional learning (English & Lachlan-Haché, 2015; McNelly, 2002).

Various types or categories of portfolio assessments have been identified in the literature (see Table 2.1). Portfolios can be reduced into two major classifications: best-work portfolios and growth portfolios (Rolheiser et al., 2000). The best-work portfolio features evidence of students' best work and can include both product and processed based artifacts whereas a growth portfolio demonstrates students' growth and development over time and often shares the student's struggles, failures, and successes (Rolheiser et al., 2000).

The use of student growth portfolios to measure teacher performance is required or encouraged in at least 13 states (English & Lachlan-Haché, 2015; Center on Great Teachers and Leaders, 2015). States and districts use portfolios because they complement classroom observations and other student performance measures (English & Lachlan-Haché, 2015). Specifically, portfolios “acknowledge that student growth may be observed in other ways than improved test scores” (English & Lachlan-Haché, 2015, p. 1). Additionally, portfolios have the ability to show the process of teaching that is otherwise not observable during classroom observations or in student achievement measures (English & Lachlan-Haché, 2015, p. 2).

Table 2.1

Examples of Portfolio Types

Author(s)	Types of Portfolios
Burke et al. (1994)	Personal, academic, professional
Campbell et al. (1997)	Working, presentation
Danielson & Abrutyn (1997)	Working, display, assessment
Haladyna (1997)	Ideal, showcase, documentation, evaluation, class
Melograno (2000)	Personal, working, record-keeping, group, thematic, integrated, showcase, electronic, multiyear
Seely (1996)	Showcase, documentation, evaluation, process
Slater (1996)	Showcase, open-format, checklist

Portfolios are appealing because they can be used to measure growth in non-tested subjects by utilizing what is already occurring in the classroom (English & Lachlan-Haché, 2015). Likewise, portfolios empower teachers because they often incorporate a self-scoring and self-reflection of student growth and give the teacher the agency to select which artifacts are included in the portfolio (English & Lachlan-Haché, 2015). Though there are variations, teachers typically include samples from low-, average-, and high-achieving students (English & Lachlan-Haché, 2015, p. 4). However, the process of selecting, reflecting, and evaluating student artifacts for the portfolio is extremely time consuming (Painter, 2001).

Growth Portfolio Design. Regardless of the portfolio classification (i.e., best-work, growth) when designing a portfolio one must keep in mind the purpose of the portfolio. As a form of assessment, a portfolio can be a form of formative or summative assessment while also serving to promote the professional growth of the teacher (Beck & Weiland, 2001; Clancy & Gardner, 2017; Gillespie et al., 1996). Key elements of the design of a portfolio to gauge student growth are a purposeful sampling of student work from multiple points in time, evidence of a connection to the standard being assessed, and the criteria for assessing the artifacts in the collection (Beck & Weiland, 2001; Gillespie et al., 1996; Henderson et al., 1995; Parkes et al., 2015; Silveira, 2013).

An important aspect of the design is the evaluation of the portfolio to gauge student growth. As with other forms of teacher evaluation, both the creator of the portfolio and the reviewer of the portfolio must be intimately familiar with the scoring rubric (Chang & Wu, 2012; Hart et al., 2014). In order for the assessment tool to be reliable, it must be task-specific with task-specific wording (Chang & Wu, 2012, Gillespie et al., 1996; Hart et al., 2014; Humphry & Heldsinger, 2014). As reviewers are gauging student growth they must also be cognizant of the fact that they are also using the portfolio to evaluate teacher performance and assess the teacher's practices (Beck & Weiland, 2001; Wolf & Dietz, 1998).

Evidence Credibility. Credibility in assessment is a major concern when considering a portfolio as a method to evaluate teacher effectiveness. As with other forms of assessment, the concern of credibility and integrity is paramount. Simply put, evidence credibility is the "believability of any teacher collected evidence that demonstrates student growth" (Popham, 2013a, p. 36) Researchers have referenced teachers exercising opportunistic and self-serving behavior (Ballou & Springer, 2015; Popham, 2013a). For example, teachers might coach their

students during a performance assessment thus influencing the product that will be used to evaluate them. All efforts to secure evidence credibility by the teacher submitting the portfolio should be taken. Because the assessment of the student growth portfolio is, in fact, an assessment of the abilities and practice of the teacher (Beck & Weiland, 2001).

To establish evidence credibility Popham (2013a) suggests four simple guidelines. First, teachers must establish a clear tie to the curriculum that they are teaching. To that point, the teacher must establish a link between the artifact and the purpose of the assessment (Clancy & Gardner, 2017).

Next, the reliability and validity of the assessment tool (i.e., scoring rubric) must be verified. Many factors can impact the validity and reliability of a rubric (Chang & Wu, 2012; Gillespie et al., 1996; Humphry & Heldsinger, 2014; Gates et al., 2015; Parkes et al., 2015; Reddy & Andrade, 2010; Rezaei & Lovorn, 2010; Silveira, 2013). From issues surrounding the interpretation of the language of the rubric, the inclusion of all stakeholders in the creation of the rubric, and the design and format of the rubric it is clear that much care must be taken in creating the rubric that will be used to assess an artifact included in a portfolio. High-quality rubrics aligned with the standards being assessed must be created (English & Lachlan-Haché, 2015). Portfolio evaluators should be involved in calibration sessions to ensure all evaluators have a common understanding of the scoring rubric (English & Lachlan-Haché, 2015). Additionally, states and districts that employ a portfolio-based evaluation system should consider having multiple evaluators review each portfolio to ensure interrater reliability (English & Lachlan-Haché, 2015).

Third, the parameters that surround the assessment must be consistent. In short, purposeful sampling from two points in time must be adhered to for the artifact to be considered

credible. (Beck & Weiland, 2001; Gillespie et al., 1996; Henderson et al., 1995; Parkes et al., 2015; Silveira, 2013). The artifacts that are collected should always represent the abilities and accomplishments of the representative student(s) to present a more reliable and complete picture of growth (Henderson et al., 1995; Panitz, 1996).

Finally, the artifacts must show an absence of item-focused instruction (Popham, 2013a). In other words, the evidence must be apparent that teachers have not specifically provided instruction to achieve a specific result. This echoes back to the suggestion that some teachers partake in opportunistic or self-serving behavior to give the appearance that growth has occurred when, in fact, it has not (Ballou & Springer, 2015; Popham, 2013a).

Research on Portfolio-Based Teacher Evaluation. Research on portfolio-based teacher evaluation systems is limited and overwhelmingly qualitative in nature. A review of literature on the effectiveness of portfolios by Bryant and Chittum (2013) indicate that, between 1996 and 2012, a majority of the empirical research was affective in nature, in other words, although original data were collected, the data addressed “participants’ feelings and opinions” about portfolios (p. 190). Case studies, focus groups, surveys, and rubrics are the most common methodologies used in portfolio based research (Rhodes et al., 2014). Rhodes et al. (2014) found that much of the research about portfolios suffered from a lack of rigor in regard to reliability and validity and small sample sizes that rendered findings from the studies limited to the context of the study.

Several dissertation studies have utilized a case study design to explore the use of portfolios as a form of teacher evaluation. Bratcher (1998), Freeman (1998), Kearley (1997), and Sutherland (1998) all utilized a case study design to explore the use of portfolios as a form of teacher evaluation. Additionally, all of the aforementioned studies had a small sample size, with

the number of participants ranging from three to seven teachers and/or administrators. As a result, the findings of the studies cannot be generalized to the greater population.

Teacher and administrator perceptions of portfolio-based evaluation have been explored in the literature. Kearley (1997) and Sutherland (1998) both explore the perceptions of teachers and administrators in their studies. Kearley's (1997) 12-week case study explored the perception of three teachers and one administrator on the use of portfolios as an element of formative evaluation to inform professional growth. Kearley collected data using interviews, a focus group, and document review and found that all participants stated that portfolios presented an opportunity to improve formative teacher evaluation that promoted professional growth. Additionally, participants stated an increased sense of collegial interaction between teacher and administrator (Kearley, 1997).

Teacher and administrator perceptions were also explored by Sutherland (1998). Sutherland's case study of five teachers and two administrators investigated the use of portfolios to promote professional growth. Sutherland collected data using interviews, observations, and document analysis and found that teachers experienced a greater sense of control over their professional development and an increased sense of professionalism compared to the previous evaluation system. Interestingly, administrators who participated in the study did not provide insight on the portfolio as a means to promote professional development though they did believe that the portfolio process increased collaboration between teacher and administrator (Sutherland, 1998).

Other studies have focuses solely on the perception of teachers in the use of portfolios as an element of teacher evaluation. For example, Bratcher (1998) utilized a case study methodology to investigate teachers' perception of portfolios on their performance and student

achievement. Bratcher's (1998) study included five middle and high school teachers and utilized document reviews, observations, and interviews as her data collection methods. In this study, teachers perceived portfolios as having a positive effect on teacher performance and student achievement (Bratcher, 1998). Freeman (1998) also utilized a case study methodology to explore the use of portfolios as a form of professional growth. Freeman's (1998) study included three elementary school teachers and utilized journals, observations, and interviews as his data collection methods. In his study, Freeman (1998) found that the use of portfolios prompted changes in teachers' instructional practices.

Attributes of Sound Evaluation

The Joint Committee on Standards for Educational Evaluation (1988) are organized under four essential constructs: fairness, accuracy, feasibility, and utility. These constructs represent national and international consensus that evaluations should be "ethical, fair, useful, feasible, and accurate" (Joint Committee on Standards for Educational Evaluation, 1988, p. 1). The committee defines evaluation as, "the systematic assessment of a person's performance and/or qualifications in relation to a professional role and some specified and defensible institutional purpose" (Joint Committee on Standards for Educational Evaluation, 1988, p. 3).

Fairness means that the evaluation system considers and protects the legal rights of those affected by the evaluation (Joint Committee on Standards for Educational Evaluation, 1988). In other words, that the evaluation is conducted in a legal and ethical way that takes into account the welfare of those involved in the evaluation. The core principle of this construct is that educational institutions exist to serve students and, as such, teacher evaluations should concentrate on determining whether educators are meeting the needs of their students (Joint Committee on Standards for Educational Evaluation, 1988).

Accuracy means the evaluation system produces sound information about performance (Joint Committee on Standards for Educational Evaluation, 1988). Specifically, the evaluation methodology (e.g., observation, student growth measure, student achievement measure) should be appropriate for the purpose of the evaluation, as well as, the individual being evaluated and the context in which they are evaluated (Joint Committee on Standards for Educational Evaluation, 1988).

Feasibility means that the evaluation system offers an efficient, easy to use, and viable means of evaluation (Joint Committee on Standards for Educational Evaluation, 1988). Additionally, the evaluation system should be adequately funded, be an efficient use of teacher and evaluator time and resources, and be a politically viable form of evaluation (Joint Committee on Standards for Educational Evaluation, 1988). This construct also acknowledges that evaluations occur in dynamic settings and can be affected by factors that can have a positive and/or negative effect on the quality of the evaluation (Joint Committee on Standards for Educational Evaluation, 1988). For example, evaluation methods, political or public pressures, time and other resource constraints can all affect the quality of an evaluation.

Utility means that the evaluation system provides timely and meaningful information in the evaluation process (Joint Committee on Standards for Educational Evaluation, 1988). This construct requires that evaluation systems provide useful information to those being evaluated as a means for improving their performance (Joint Committee on Standards for Educational Research, 1988).

Sound Portfolio-Based Evaluations

Multiple studies, Tucker et al. (2003), Attinello et al. (2006), and Denison (2008) have all used the constructs of fairness, accuracy, feasibility, and utility as the guiding constructs of their

studies on teachers' and administrators' perception of portfolio based evaluation models. Tucker et al.'s (2003) survey, *Teacher Survey: Perceived Value of Teacher Portfolios*, was designed based off of the four constructs outlined by the Joint Committee on Standards for Educational Evaluation (1988).

Several mixed methods studies have been conducted to examine teachers' perceptions of portfolios as an element of teacher evaluation. Tucker et al. (2003) explored teacher and administrator perceptions of the use of portfolios in teacher evaluation and professional growth. Their multiyear study was conducted in the Williamsburg James City County School District in Virginia and included teachers ($n = 309$) and administrators ($n = 15$) from all 11 schools in the school district. Tucker et al. (2003) utilized archival record analysis, a Likert-type scale survey questionnaire, and focus groups as their data collection methods. Researchers conducted two teacher focus groups, one for elementary and one for secondary school teachers with a total of 30 teachers, and one focus group of eight administrators (Tucker et al., 2003).

Tucker et al. (2003) created the *Teacher Survey: Perceived Value of Teacher Portfolios* and focus group questions based off of four constructs (i.e., fairness, utility, feasibility, accuracy) developed by the Joint Committee on Standards for Educational Evaluation (1988). Tucker et al. (2003) found that teachers and administrators felt that portfolios were both fair and accurate whereas their utility and feasibility were rated less positively overall. In regards to utility, focus group participants did not feel that the portfolio process prompted good teaching habits or encouraged them to alter their current instructional practices (Tucker et al., 2003). Both teachers and administrators felt that the time required for the portfolio was an obstacle (Tucker et al., 2003). Across all four constructs, administrators rated the portfolio process more positively than teachers (Tucker et al., 2003).

Attinello et al. (2006), replicated the study conducted by Tucker et al. (2003) for Attinello's (2004) dissertation study with a larger population in the Forsyth County School System in Georgia. Surveys were sent to teachers ($n = 752$) and administrators ($n = 46$) from 19 of the 23 schools in the district. Researchers used modified versions of the teacher survey and focus group questions developed by Tucker et al. (2003), portfolio data, and interviews as their data collection methods (Attinello et al., 2006). Interviews were conducted with 10 teachers and eight administrators in addition to conducting three focus groups of eight teachers and eight administrators (Attinello et al., 2006). Results from the study found that teachers and administrators believed that portfolios were an accurate and more comprehensive reflection of teacher performance with positive perceptions of portfolios across all four constructs (i.e., fairness, accuracy, utility, feasibility) although time and effort of the portfolio process emerged as disadvantages based on analysis of the qualitative data (Attinello et al., 2006). As with Tucker et al.'s (2003) study, administrators rated the portfolio process more positively than teachers across all four constructs of the survey (Attinello et al., 2006).

Denison (2008) replicated, in part, the study of Tucker et al. (2003). Whereas Tucker et al. (2003) and Attinello et al. (2006) investigated the perception of teachers and administrators, Denison only explored the perception of public school K-12 teachers ($N = 679$) in her study. Surveys were sent to all active K-12 teachers from 36 of the 40 school districts in the state of Utah. Denison utilized a modified version of the Tucker et al. (2003) survey that included open-ended questions and conducted 40 phone interviews with representatives from each of the 40 school districts in the state. Results from the study found that, across all four constructs of the survey, teachers rated portfolios positively. However, ratings for accuracy, utility, and feasibility were all rated toward the midpoint of the scale. Qualitative data from open-ended questions and

interviews expressed a much more negative view of the portfolio process as a whole, especially for the constructs of accuracy and feasibility (Denison, 2008). Qualitative data revealed that some teachers doubted the accuracy of the portfolio process because they believed portfolios could be faked and that there were inconsistencies in the implementation of the portfolio among school districts (Denison, 2008). Denison (2008) found that some teachers did not believe the portfolio was worth the time and effort if the administrators were not spending equitable time reviewing and providing feedback to teachers.

Use of Summative Teacher Evaluation Scores

State and District Uses

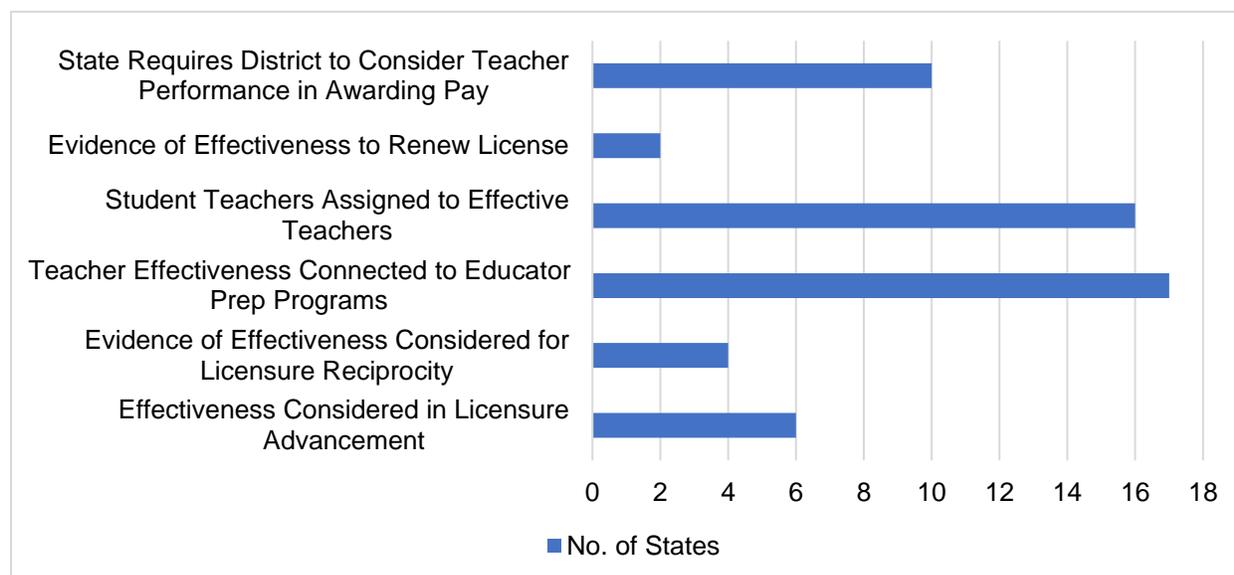
States and districts are required to release report cards under Title I of the Elementary and Secondary Education Act of 1965 as amended by the Every Student Succeeds Act (USDE, 2019). Report cards provide stakeholders (e.g., parents, taxpayers) with information concerning school performance and progress at the state, district, and school levels (USDE, 2019). Report cards include information on educational spending, parent engagement, student achievement, graduation rates, accountability measures, school environment, and teacher qualifications (USDE, 2019). Additionally, state, district, and school report cards must be readily available and accessible to all stakeholders (USDE, 2019). Although state and district report cards contain data that are used to calculate teachers' level of overall effectiveness (i.e., student growth and student achievement) the data are aggregate and do not have a direct result on individual teachers' summative evaluation scores.

States and districts use teacher summative evaluation scores in a variety of ways. At the state level, teacher summative evaluation scores may impact decisions about licensure advancement (i.e., moving from a temporary, or provisional, to a professional license), licensure

reciprocity (i.e., the ability to transfer a teaching license from one state to another), eligibility to serve as a mentor teacher for a student teacher, and an accountability measure for educator preparation programs (see Figure 2). Although the use of summative evaluation scores are used to identify mentor teachers for preservice teachers and as an accountability measure for educator preparation programs in a number of states, a number of states tie summative evaluation scores to teacher licensure and pay. Ten states require districts to consider teachers' performance in awarding pay, six states consider teacher effectiveness in the advancement of teacher licensure, and two states require evidence of teacher effectiveness to renew licensure (NCTQ, 2021).

Figure 2

State-Level Use of Summative Teacher Evaluation Scores



Note: State level use of summative teacher evaluation scores are used in a variety of ways.

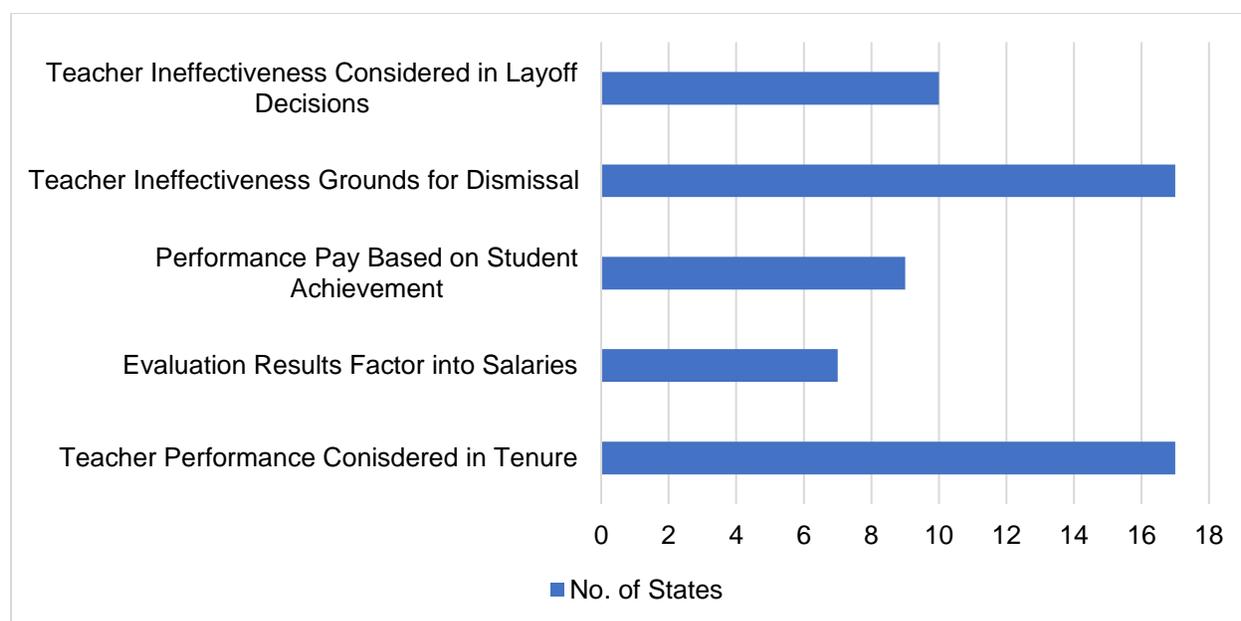
Summative evaluation scores can impact pay, licensure renewal and advancement, and even the educator preparation programs teachers attended to earn their license. State Level Use of Summative Teacher Evaluation Scores adapted from “State teacher policy database,” 2021,

National Council on Teacher Quality, (<https://www.nctq.org/yearbook/home>). Copyright 2021 by the National Council on Teacher Quality.

District use of summative teacher evaluation scores in a variety of ways. Summative teacher evaluation scores are used at the local level to determine tenure, compensation, and retention (see Figure 3). Teacher ineffectiveness is grounds for dismissal in 17 states and considered in decisions to lay off teachers in 10 states when a reduction in force is necessary (NCTQ, 2021). Summative teacher evaluation scores are considered when awarding tenure to teachers in 17 states (NCTQ, 2021). Teacher salaries are impacted by summative evaluation scores in seven states whereas nine states award performance pay based on student achievement (NCTQ, 2021).

Figure 3

District-Level Use of Summative Teacher Evaluation Scores



Note: School districts use summative teacher evaluation scores in a variety of ways. Summative teacher evaluation scores can effect teacher compensation, tenure, and continued employment.

District Level Use of Summative Teacher Evaluation Scores adapted from “State Teacher Policy

Database,” 2021, National Council on Teacher Quality, (<https://www.nctq.org/yearbook/home>).

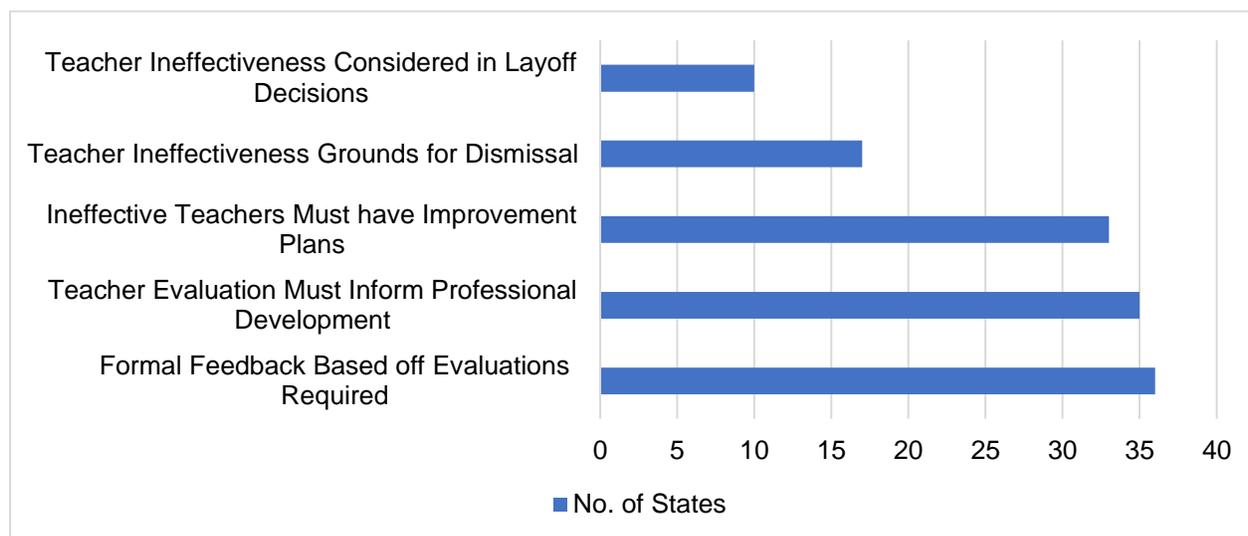
Copyright 2021 by the National Council on Teacher Quality.

Administrator Uses

There is a great deal of resistance from teachers on the use of evaluation data to make high-stakes personnel decisions (e.g., pay, promotion, and dismissal; Connally & Tooley, 2016). Such high-stakes accountability evaluation systems have been criticized because of their focus on the summative aspects of teacher evaluation instead of the professional development uses of the evaluation data (Derrington & Martinez, 2019; Sullivan, 2016). However, legislators, state officials, and other policymakers contend that public schools are supported by tax-payer money and therefore subject to public oversight to ensure teacher quality meets public expectations (Danielson & McGreal, 2000).

Figure 4

Administrator Use of Summative Teacher Evaluation Scores



Note: School administrators use summative teacher evaluation scores in many ways. In addition to continuing employment, administrators use summative teacher evaluation scores to develop individual teacher improvement plans for ineffective teachers. Evaluation scores are also used by

administrators to develop school-wide professional development plans. Administrator Use of Summative Teacher Evaluation Scores adapted from “State Teacher Policy Database,” 2021, National Council on Teacher Quality, (<https://www.nctq.org/yearbook/home>). Copyright 2021 by the National Council on Teacher Quality.

Administrators also use summative teacher evaluation scores in a variety of ways. Summative teacher evaluation scores are used by administrators to provide formal feedback, plan for professional development, and make teacher retention decisions (see Figure 4). As stated by Bandura (2001), “feedback that is most informative and achieves the greatest improvements relies on corrective modeling” (i.e., observational learning; p. 277). Administrators in 17 states can dismiss teachers who have ratings of “ineffective” (NCTQ, 2021). Interestingly, formal feedback based on summative evaluation scores is only required in 36 states although seven states do not provide formal feedback other than copies of evaluations to teachers (NCTQ, 2021). According to a majority of state policies, professional development is a key aspect of evaluation with 33 states requiring ineffective teachers to be placed on teacher improvement plans with 35 states requiring teacher professional development be informed by their evaluations (NCTQ, 2021).

Teacher Uses

The purpose of teacher evaluation is to promote professional learning (Danielson, 2011). The National Council on Teacher Quality reports that 35 states and the District of Columbia require that teacher evaluations inform their professional development (Doherty & Jacobs, 2015). Formal and informal observations and the subsequent professional conversations support the developmental purposes of teacher evaluation (Danielson, 2011). However, teachers’

perceptions of the evaluations informing their instructional practices and professional learning are dependent on the confidence they have in their evaluator (Tuytens & Devos, 2010).

Studies have found that related arts teachers (e.g., fine arts, physical education, world languages) do not believe that their evaluators have the requisite training or content specific knowledge to accurately evaluate them (Norris et al., 2017; Robinson, 2017). Norris et al. (2017) found that physical education teachers lacked confidence in their evaluators and that administrators' inexperience in the physical education classroom resulted in an unfair assessment. Similarly, Robinson's (2017) study of eight music teachers across five states found that music teachers did not have confidence in their evaluator's ability to provide effective feedback. In fact, inconsistencies among evaluators and their level of music background "posed problems in terms of the perceived usefulness, and even validity, of their evaluations" (Robinson, 2017, p. 6). As a remedy, Lochmiller (2016) suggests that feedback to encourage teacher growth might best be provided by content-specific peers (e.g., department chairs, curriculum consulting teachers).

The crux of effective and useful evaluation to inform professional growth is meaningful feedback (Ritter & Barnett, 2016). Derrington and Martinez (2019) explored teachers' perceptions of evaluation in Tennessee after five years of the TEAM evaluation system implementation. Participants were 143 teachers from 14 middle and high schools in eastern Tennessee. Derrington and Martinez (2019) found that the "perception of evaluation effectiveness contributes to teacher motivation to change their instructional practices" (p. 43). This finding is in line with Bandura's (1977, 1986) social cognitive theory and its construct of reciprocal determinism. In fact, teachers find classroom observations and their subsequent

observation-based feedback useful in contrast to growth scores (Firestone & Donaldson, 2018). Meaningful feedback encourages teachers to be more self-reflective (Ritter & Barnett, 2016).

Specific feedback helps teachers focus their professional learning on areas that have been identified as a strength or weakness (Smith et al., 2020). Smith et al. (2020) used Bandura's (1977) social cognitive theory as the theoretical framework for their cross-sectional quantitative study to investigate the relationship between teacher evaluation feedback and instructional practice self-efficacy in secondary school teachers. Participants were teachers ($N = 69$) from two school districts in west central Florida. Smith et al. (2020) found that "higher levels of teacher self-efficacy of instructional practice are positively associated with both the perceived value of feedback and the degree to which feedback is specific to instructional practice" (p. 689).

Summary

This review of literature provided an overview of historical and contemporary perspectives of teacher evaluation in the United States, elements of multi-measure summative evaluation, portfolio-based evaluation models, attributes of sound evaluation, and the uses of summative evaluation scores by states, school districts, administrators, and teachers. The constructs of reciprocal determinism, observational learning, and self-efficacy from Bandura's (1977; 1986) social cognitive theory served as a theoretical framework in which the literature was viewed.

Teacher evaluation in the United States has changed significantly since Kratz (1896) created his list of characteristics to evaluate teachers. The passage of RTTT legislation tied federal funding for education to high stakes teacher evaluation systems (Close et al., 2019). In an effort to secure federal funding, many states moved to multi-measure evaluation models (i.e., models using more than one data source) to evaluate teacher effectiveness (Hanover Research,

2012). In new multi-measure evaluation systems, teacher quality and development are key aspects of summative evaluation systems (Connally & Tooley, 2016; Danielson, 2010; Marzano, 2012); Robinson, n.d.). Summative evaluation scores are comprised of teacher observations and measures of student growth and achievement (Derrington & Martinez, 2019).

Student growth portfolios have been used to evaluate teachers since the 1990s (English & Lachlan-Haché, 2015; Wolf & Dietz, 1998). Portfolios are attractive to states and districts because they can serve as formative or summative assessment in addition to promoting professional growth of teachers (Beck & Weiland, 2001; Clancy & Gardner, 2017; Gillespie et al., 1996). Despite the positive perception of portfolios from educational leaders at the state, district, and school levels, student growth scores can be difficult to understand, can fluctuate from year to year, and are not reported until students move on, rendering their utility as limited (Garet et al., 2017; McCullough et al., 2015).

Although a number of states require or encourage the use student growth portfolios to measure teacher performance (English & Lachlan-Haché, 2015; Center on Great Teachers and Leaders, 2015), research on portfolio-based teacher evaluation systems has been overwhelmingly qualitative, small in scale, and affective in nature (Bryant & Chittum, 2013). Studies by Tucker et al. (2003), Attinello et al. (2006), and Denison (2008) have employed mixed methods research designs based on the constructs of fairness, accuracy, feasibility, and utility developed by the Joint Committee on Standards for Educational Evaluation (1988). These studies examine the perceptions of the fairness, accuracy, feasibility, and utility of portfolio-based teacher evaluation from the perspectives of teachers and administrators.

States, school districts, administrators, and teachers all use summative teacher evaluation scores in different ways. Summative teacher evaluation scores can impact pay, licensure, renewal

and advancement at the state level whereas teacher compensation, tenure, and continued employment are impacted at the district level (NCTQ, 2021). Administrators use summative evaluation scores in two ways, for individual and school wide professional development and human capital decisions (NCTQ, 2021). Teachers utilize summative evaluation scores to inform their professional development (Doherty & Jacobs, 2015). Indeed, Ritter and Barnett (2016) explain that meaningful feedback teachers receive from their evaluators is the key to informing professional growth.

Chapter III

Methodology

The primary purposes of this study were to describe (a) fine arts teachers' perceptions of the portfolio model as a summative evaluation measure and (b) to determine how fine arts teachers use their summative evaluation scores to inform their teaching practice. An additional purpose was to understand how teachers' level of overall effectiveness scores have changed since implementing the Tennessee Fine Arts Student Growth Portfolio Model. This chapter addresses study context, research design, participants, instrumentation, participant recruitment, data collection, analysis, the trustworthiness and credibility of the study, and the researcher's reflexivity. The constructs, instrumentation, data collection, and analysis measures are aligned with the research questions, as indicated in the research matrix (see Appendix C). The research questions that guided this study were:

1. What are the perceptions of fine arts teachers regarding the use of student growth portfolios as an element of teacher evaluation?
2. How do fine arts teachers use their summative evaluation scores to inform their teaching practice?

Context

As part of the Tennessee TEAM evaluation process, a portion of teachers' level of overall effectiveness scores come from a student growth measure (Tennessee Department of Education, 2020a). Tested subjects (e.g., English, biology, algebra) utilize end-of-course exams to calculate teacher growth scores (e.g., VAM), while some fine arts teachers participate in a portfolio-based model. The Tennessee Fine Arts Student Growth Portfolio Model was developed by theatre teachers from Memphis City Schools during the 2011-2012 school year.

Three school districts piloted the model during the 2012-2013 academic year. The Tennessee Board of Education approved the model as an alternative measure of student growth starting with the 2013-2014 school year. In its first year of approval, theatre teachers across the state submitted portfolios. Visual arts teachers were added during the 2015-2016 academic year, followed by music teachers during the 2016-2017 academic year. At the height of its use during the 2017-2018 school year, 26 districts across the state implemented the Tennessee Fine Arts Student Growth Portfolio Model. Currently, six school districts participate in the fine arts portfolio: District A, District B, District C, District D, District E, and District F (see Table 3.1).

Table 3.1

Total Secondary Fine Arts Teachers, by District and Discipline (N = 251)

	District A (n = 7)	District B (n = 25)	District C (n = 3)	District D (n = 41)	District E (n = 5)	District F (n = 170)
Dance	0	1	0	0	0	10
M - I	2	8	3	10	1	35
M - V	2	4	0	7	1	35
Theatre	1	5	0	9	0	20
Visual Arts	2	7	0	15	3	70

The COVID-19 pandemic caused the widespread interruption of education due to statewide shutdowns that started in March 2020. The 2020-2021 school year was atypical because TN school districts delivered instruction in the best way for their community. While some school districts held face-to-face classes, others conducted school using a virtual model or employed a combination of live and virtual instruction—at times transitioning from face-to-face to virtual models, as COVID-19 cases increased locally.

Due to the COVID-19 pandemic, the state of Tennessee waived 2019-2020 and 2020-2021 state-level assessments and teacher evaluations. As a result, teachers did not receive a portfolio score or a level of overall effectiveness score for the 2019-2020 or 2020-2021 academic years. Additionally, as an attempt to mitigate the spread of COVID-19, District F placed a

moratorium on all external research requests until further notice. As a result, teachers from District F will not be able to participate in the study.

Participating school districts consisted of Districts A through E. Student and fine arts teacher populations in each district varied greatly (see Table 3.2). Student populations ranged from a low of 4,048 (District C) to a high of 36,606 (District D). The total number of fine arts teachers also varied by school district. There was a total of 81 fine arts teachers across the five participating school districts (see Table 3.3). The largest participating school district, District D, had a total of 41 fine arts teachers whereas the smallest participating district, District C, had only three fine arts teachers. The number of fine arts teachers by discipline also varied greatly. Of the 81 fine arts teachers, one taught dance, 24 taught instrumental music, 4 taught vocal music, 15 taught theatre, and 27 taught visual art.

Table 3.2

Participating District Information

	District A	District B	District C	District D	District E
Student Population	7,303	12,729	4,048	36,606	6,500
Number of High School Fine Arts Teachers	7	25	3	41	5

Table 3.3

Potential Secondary Fine Arts Teacher Participants, by Discipline (N = 81)

	District A (n = 7)	District B (n = 25)	District C (n = 3)	District D (n = 41)	District E (n = 5)
Dance	0	1	0	0	0
Music – Inst	2	8	3	10	1
Music - Voice	2	4	0	7	1
Theatre	1	5	0	9	0
Visual Arts	2	7	0	15	3

Research Design

This convergent mixed methods study will address fine arts teachers' perception of the Tennessee Fine Arts Student Growth Portfolio as a measure of teacher performance. Convergent design is a single-phase design that utilizes quantitative and qualitative data. In this design, "a researcher collects both quantitative and qualitative data, analyzes them separately, and then compares the results to see if the findings confirm or disconfirm each other" (Creswell & Creswell, 2018, p.217). The primary assumption of a convergent design is that qualitative and quantitative data provide "different types of information" (Creswell & Creswell, 2018, p. 217). With convergent design both quantitative and qualitative data are collected exploring the same constructs (i.e., fairness, utility, feasibility, accuracy; Creswell & Creswell, 2018).

Both qualitative and quantitative research methods have inherent limitations, but by utilizing a mixed methods approach, the limitations of each method can be offset (Creswell & Plano Clark, 2018). This method provides a more complete picture by validating findings, collecting richer data, and initiating new thinking methods (Creswell & Plano Clark, 2018; Johnson et al., 2007). In this study, the qualitative data help explain the results of the quantitative data (Creswell & Creswell, 2018). As such, the quantitative and qualitative data are equally weighted (Creswell & Plano Clark, 2018).

Participants

Participant Selection

Participants were selected based on criterion sampling for both the quantitative and qualitative phases of this study. Criterion sampling reviews all cases that meet predetermined criteria (Patton, 1990). The initial intent of this study was to use confirming and disconfirming purposeful sampling method to select interview participants that had both positive and negative

perceptions of portfolio evaluation. However, analysis of quantitative data revealed that the perceptions of those who agreed to be interviewed were generally negative. Because the initial interview sampling method was no longer valid, I employed the same criterion sampling method as the quantitative portion of the study.

Participants for the survey portion included secondary school fine arts teachers with a valid Tennessee professional educator license with an endorsement in their specific art form (i.e., dance, music, theatre, or visual art). To be eligible to participate in the study, participants were required to have a minimum of 4 years of classroom experience as a full-time teacher whose instructional responsibilities primarily (i.e., at least 75% of the day) pertain to teaching fine arts. Additionally, participants were required to have participated in the Tennessee Fine Arts Student Growth Portfolio for at least two academic years.

Survey Participants

Invitations to participate in the survey portion of the study were sent to all 81 fine arts teachers. In total, 39 fine arts teachers completed the survey portion of the study. Of the 39 survey submissions 11 were not valid because they did not teach high school or did not participate in the portfolio for the minimum number of years. The 28 eligible fine arts teachers were comprised of 1 dance, 8 instrumental music, 5 vocal music, 9 theatre, and 5 visual art teachers (see Table 3.4).

Table 3.4

Secondary Fine Arts Teacher Participants, by Discipline (N = 28)

	District A (n = 1)	District B (n = 4)	District C (n = 2)	District D (n = 18)	District E (n = 1)
Dance	0	1	0	0	0
Music – Inst	0	0	1	6	0
Music - Voice	0	0	0	3	1
Theatre	0	2	0	7	0
Visual Arts	1	1	1	2	0

The majority of survey participants were White females with a master's degree (see Table 3.5). Of the 28 participants only three teachers were people of color. Female teachers accounted for 57% (16) of participants with the remaining 43% (12) being male. Nearly half (46%) of participants had a fine arts-specific degree (i.e., Bachelor of Fine Arts, Master of Fine Arts). More than half (57%) of participants had a graduate degree (i.e., master's, Master of Fine Arts, Education Specialist). In regards to tenure, 79% of participants held tenure with only six non-tenured teachers participating in the study.

Table 3.5

Participant Gender, Race, Education Level, and Tenure Status (N = 28)

Participants	<i>n</i>	%
Gender		
Female	16	57
Male	12	43
Race		
Black/African American	3	11
White	25	89
Degree Type		
Bachelor's Degree	4	14
Bachelor of Fine Arts	8	29
Master's Degree	9	32
Master of Fine Arts	5	18
Education Specialist Degree	2	7
Doctorate	0	0
Tenure Status		
Non-tenure	6	21
Tenure	22	79

Participants' years of experience as fine arts teachers ranged from 4 years to more than 20 years, with a mean of 15.32 years of teaching experience (see Table 3.6). Three of the five participating districts have participated in the Tennessee Fine Arts Student Growth Portfolio for since the 2013-2014 academic year. The remaining two districts implemented the portfolio model in the 2014-2015 academic year. Most of the participants ($n = 20$) came from District D. Both District A and E had one participant with District B having three and District C having two

fine arts teachers participant. Of the 28 participants, eight have served as a peer reviewer for the state at least once with four serving as a peer reviewer two or more times.

Table 3.6

Participant Years of Experience, District Portfolio Use, and Peer Reviewer Status (N = 28)

Participant	Years of Experience	Years of District Portfolio Use	Peer Reviewer
A1	15-19	4-5	No
B1	4-9	6-7	Yes
B2	4-9	6-7	No
B3	20+	6-7	No
B4	20+	6-7	No
C1	20+	6-7	Yes
C2	15-19	6-7	Yes
D1	10-14	6-7	No
D2	4-9	6-7	Yes
D3	4-9	6-7	No
D4	4-9	6-7	No
D5	4-9	6-7	No
D6	20+	6-7	No
D7	4-9	6-7	Yes
D8	15-19	6-7	No
D9	15-19	6-7	No
D10	20+	6-7	No
D11	20+	6-7	No
D12	20+	6-7	No
D13	20+	6-7	No
D14	15-19	6-7	Yes
D15	20+	6-7	No
D16	15-19	6-7	Yes
D17	20+	6-7	No
D18	15-19	6-7	No
D19	20+	6-7	No
D20	15-19	6-7	Yes
E1	10-14	4-5	No

Note. Participants were labeled to match their respective school district letter. For example,

Participant A1 is a teacher from District A.

Interview Participants

The six individuals who took part in the interview portion of the study are identified using the following pseudonyms: Blanche, Dorothy, Miles, Rose, Sophia, and Stan. All of the interview participants were theatre teachers. Interview participants were comprised of four

females and two males. Years of experience range from 7 years to more than 21 years of teaching experience. Four of the six interview participants held a master's or Master of Fine Arts degree (see Table 3.7).

Table 3.7

Interview Participant Demographics (N = 6)

Participant	Years of Experience	Degree	Peer Reviewer
Blanche	15	Master's	Yes
Dorothy	21+	Master of Fine Arts	No
Miles	21+	Bachelor's	No
Rose	21+	Master's	No
Sophia	7	Bachelor of Fine Arts	Yes
Stan	21+	Master's	No

Instrumentation

This section discusses the process for designing and conducting the survey and semistructured interviews and provides sample questions for each instrument.

Survey

The survey was created using the Google Forms, a web-based platform. The survey consisted of two sections used to collect responses to the Perceived Value of Teacher Portfolios Questionnaire, and demographic information. The survey ended with a question asking if participants were willing to participate in a follow-up interview (see Appendix D). The survey took approximately 30 minutes to complete.

Perceived Value of Teacher Portfolios Questionnaire. A modified version of the Perceived Value of Teacher Portfolios Questionnaire, developed by Tucker et al. (2003), was used for this study (see Appendix F). The Perceived Value of Teacher Portfolios questionnaire was selected to address Research Questions 1 and 2. Research Question 1 relates to the construct of teacher perception of summative evaluation scores and Research Question 2 relates to the

construct of professional learning and how summative evaluation scores inform teachers' practice.

Tucker et al. (2003) developed the Perceived Value of Teacher Portfolios Questionnaire to explore the efficacy of teacher portfolios as an element of teacher evaluation and professional growth. Tucker et al. (2003) established internal consistency reliability for the total sample of participants as well as subsets (i.e., teachers, administrators) using the coefficient alpha method, which produces a measure of internal consistency reliability between 0 and 1.00. Internal consistency reliability indicates whether survey items within each of the four constructs (i.e., accuracy, fairness, feasibility, utility) measure the same knowledge area in a consistent way.

The questionnaire measured four constructs as outlined by the Joint Committees on Standards for Educational Evaluation (1998). The four constructs are:

1. Fairness: The evaluation system considers and protects the legal rights of those affected by evaluation.
2. Utility: The evaluation system provides timely and meaningful information in the evaluation process.
3. Feasibility: The evaluation system offers an efficient, easy to use, and viable means of evaluation.
4. Accuracy: The evaluation system produces sound information about performance (Joint Committee on Standards for Educational Evaluation, 1998, p. 7).

The response format for this 25-item survey consisted of 17 items as a Likert-type scale ranging from 4 ("strongly agree") to 1 ("strongly disagree") with the option of "don't know". The remaining survey items collected demographic information. Two sample questions from each of the four abovementioned constructs are provided. Sample fairness items include "The

portfolio is a fair means for me to demonstrate my performance of professional standards” and “The portfolio gives me a more prominent role in the evaluation process.” Sample usefulness items include “The portfolio promotes good teaching practices” and “Development of a portfolio encourages self-reflection about my work.” Sample feasibility items include “The time required by me to develop a portfolio is reasonable” and “The portfolio is a practical strategy for me to demonstrate my performance of professional standards.” Sample accuracy items include “The portfolio is a means to provide evidence of my fulfillment of professional standards not readily observable” and “The portfolio is an accurate reflection of my performance of professional standards.” The survey also included two general questions regarding the effect the portfolio had on professional development and the evaluation process in general. A sample general question is “To what degree do portfolios promote the professional development of teachers in your building?” The original survey also includes two open-ended questions about the advantages and disadvantages of the portfolio evaluation.

The modified Perceived Value of Teacher Portfolios Questionnaire consists of 18 Likert scale items ranging from 1 (“strongly disagree”) to 5 (“strongly agree”) and two open-ended questions. A five-point Likert scale was used because items requiring a “don’t know” option were eliminated. The original Perceived Value of Teacher Portfolios Questionnaire gauged the perception of a portfolio system that was evaluated by teachers’ administrators.

Demographic Information. The final portion of the survey was used to collect participants' demographic information. The survey collected the following demographic information: age, gender, race, ethnicity, years of experience, certification area(s), subject(s) taught, school size, number of fine arts teachers in school, and school location.

Semistructured Interviews

Semistructured interviews are designed to encourage the participant to speak on their terms by asking questions that are not overly specific (Patton, 1990) and allow for a conversation-style environment instead of a more structured interviewer/interviewee dynamic. Participants were asked 21 questions as outlined in the semistructured interview protocol (see Appendix G). The first set of questions were related to teachers' summative evaluation scores, specifically, about any changes in each sub-component (i.e., observation, achievement, student growth). These questions relate to Research Question 1. The next section of questions addressed the teachers' perception of the accuracy and acceptability of the portfolio system. For example, teachers were asked questions about their understanding of the portfolio model, the scoring process and its fairness. These questions related to Research Question 1. The third set of questions pertained to the utility of the portfolio system. For example, teachers were asked how they have used their summative evaluation scores to inform their practice as a teacher. These questions helped answer Research Question 2.

Procedure

This section outlines the procedures for participant recruitment, data collection, and data analysis.

Participant Recruitment

Upon institutional and local IRB approval, potential participants were identified through an initial email sent to fine arts coordinators or portfolio leads of districts that utilize the Tennessee Fine Arts Student Growth Portfolio (see Appendix H). An introductory email was sent to the fine arts coordinators or state portfolio leads at each participating school district (see Appendix I). The initial email request (a) described the study, (b) explained participation criteria,

and (c) asked them to either provide a list of fine arts teachers who meet the participation criteria or asked them to send the survey on my behalf. Based on their suggestions, potential participants received a recruitment email with information about the study (see Appendix J). Participants also received an informed consent form as part of the recruitment email (see Appendix K).

Teachers willing to participate in interviews were identified through their survey responses. Six participants were selected based on their academic discipline and availability to participate in a 30-minute follow-up interview. Upon selection, interview participants were sent an introductory email with the interview informed consent form (see Appendix L and M). Once the informed consent form was received, the interview was scheduled based on participant availability.

Data Collection

Survey. A link to the survey was emailed to all eligible secondary fine arts teachers. Participants were allotted 2 weeks to complete the data collection survey. A reminder email was sent after Day 5 and Day 10. The survey took 30 minutes to complete. Survey data were collected in August and September of 2021. Responses were downloaded from Google Forms after 14 days and stored on a password protected computer.

Semistructured Interviews. Interviews were conducted virtually using videoconferencing software (i.e., Zoom). Interviews were scheduled at a time that was convenient to the participants. Video interviews were recorded using videoconferencing software and were also transcribed verbatim using the videoconferencing software or the Rev Transcription application. Each interview lasted between 16 and 26 minutes. After transcription, interview transcripts were sent to participants to allow them to make any corrections to the

interview transcripts. Participants were given 5 days to make corrections. All six interview participants replied and approved their interview transcript without correction.

Data Analysis

Survey. The survey responses were analyzed based on the type of data collected. The Perceived Value of Teacher Portfolios Questionnaire, the summative evaluation score data, and demographic portion of the survey were analyzed using descriptive analysis. The two open-ended questions were coded as described below, using in vivo and focused coding (Saldaña, 2016).

Semistructured Interviews. Qualitative data collected from semistructured interviews was prepared and organized for analysis. This study hopes to identify teachers' perceptions; as such, pre-coding is not possible (Lincoln & Guba, 1985; Saldaña, 2016). For first cycle coding, I used the in vivo coding method as this method attunes the researcher to "participant perspectives and actions" (Saldaña, 2016, p. 73). Also known as "verbatim coding" and "literal coding," in vivo coding uses participant's own words as codes (Saldaña, 2016, p. 105). The codes produced by the first coding cycle were separated into categories (Saldaña, 2016). Next, interviews were recoded to pare down further the codes and categories (Saldaña, 2016). Second cycle coding was then used to help further organize and narrow the codes and themes of the first cycle coding process. Focused coding follows in vivo coding and "searches for the most frequent or significant codes to develop the most salient categories in the data" (Saldaña, 2016, p. 240). During both coding cycles, data were coded and recoded multiple times. Data were coded, and emerging categories and themes were documented (Saldaña, 2016). During coding, saliency analysis was also applied to the open-ended survey items and semistructured interviews. Saliency analysis allows inclusion of codes that are salient (i.e., important) but do not converge

into existing themes (Buetow, 2010). Using saliency analysis allows these non-recurrent codes to be accommodated (i.e., included) without overstressing the existing themes (Buetow, 2010).

Trustworthiness and Credibility

In qualitative research, the researcher is often considered the instrument of inquiry and is naturally invested in the researched topic (Brantlinger et al., 2005). As a result, steps must be taken to ensure that the researcher and their data are credible and trustworthy (Brantlinger et al., 2005). Simply put, qualitative researchers must be able to answer the question, "can the findings be trusted" (Korstjens & Moser, 2018, p. 121). Research is trustworthy when researchers "demonstrate that data analysis has been conducted in a precise, consistent, and exhaustive manner through recording, systematizing, and disclosing the methods of analysis with enough detail to enable the reader to determine whether the process is credible" (Nowell et al., 2017, p.1). Credibility "is the equivalent of internal validity in quantitative research and is concerned with the aspect of truth-value" (Korstjens & Moser, 2018, p. 121). Internal validity is defined as "the extent to which the observed results represent the truth in the population we are studying and, thus, are not due to methodological errors" (Patino & Ferreira, 2018, p. 183). Creswell and Creswell (2018) suggest using multiple strategies to ensure the accuracy of findings, and in turn, convince consumers of research that the results are accurate.

To enhance this study's trustworthiness and credibility, I used methodological triangulation, member checks, an audit trail, peer debriefing, thick description, particularizability, and researcher reflexivity. Methodological triangulation utilizes more than one method to study a phenomenon (Bekhet & Zauszniewski, 2012). This convergent mixed methods study utilized quantitative and qualitative data collection methods (i.e., survey and semistructured interviews). Member checks are the process of having participants review

interview transcripts for accuracy (Brantlinger et al., 2005). Interview participants were given 5 days to review interview transcripts and provide corrections. An audit trail is a detailed record of dates and times of study interviews (Brantlinger et al., 2005). Particularizability is the process of “documenting cases with thick description so that researchers can determine the degree of transferability to their own situations” (Brantlinger et al., 2005, p. 201). Lastly, researcher reflexivity, or a researcher's self-disclosure statement, allows the researcher to share their "assumptions, beliefs, values, and biases" with consumers of their research (Brantlinger et al., 2005), and is provided in the next section.

Researcher Reflexivity

I approach this research from the perspective of a secondary school theatre arts teacher who actively participates in the Tennessee Fine Arts Student Growth Portfolio Model. I am a Level 5 teacher and have been for the past four years. I have earned a five (highest possible score) on my portfolio every year except for my first year teaching in Tennessee. I have worked at the same high school and taught theatre exclusively for the past five years. I have served as my district's theatre arts lead teacher for Grades 6-12 and facilitated over 72 hours of professional development geared towards developing lesson plans and collecting artifacts for fine arts portfolio submissions. I served as a state-certified portfolio peer reviewer for 3 of those years.

I have worked as an actor, director, and theatre teaching artist for almost 20 years. In addition to my work as a classroom teacher, I have held various leadership positions in professional theatre organizations (e.g., Southeastern Theatre Conference, Theatre for Young Audiences/United States of America, and Tennessee Theatre Association), where my focus has been theatre for young audiences and theatre education. I have a dual bachelor's degree in theatre performance and theatre design, a Master of Fine Arts in theatre for youth, and a Master of Arts

in teaching. Before transitioning to my current post as a secondary school theatre teacher, I taught undergraduate theatre courses in Arizona and Tennessee. Further, I have worked with other alternative assessment models while a special education teacher in Kentucky, where I taught youth with profound and severe disabilities.

I understand the importance of reflective teaching practices that promote student achievement and professional growth. I understand that the Tennessee Fine Arts Student Growth Portfolio aims to foster student growth and serve as a tool for professional development (Wilson, 2020). I recognize that I am considered an authority to many fine arts teachers in my school district by virtue of my credentials, leadership position within the district, and organizational affiliations. As a researcher, I understand that I hold a unique position as a producer and reviewer of student growth portfolios. In my experience as the theatre arts lead teacher, I have perceived a disconnect between state growth portfolio scores and its accuracy as a measure of teacher performance. Additionally, I often hear teachers in my school district talk about the ineffectiveness of portfolio scores as a tool to inform professional growth. These factors have informed and driven me to investigate further the Tennessee Fine Arts Student Growth Portfolio Model as an accurate measure of teacher performance and an informer of professional growth.

Chapter IV

Findings

This convergent mixed methods study addressed fine arts teachers' perception of the Tennessee Fine Arts Student Growth Portfolio as a measure of teacher evaluation. Quantitative and qualitative data were gathered to address the study's research questions. A Likert scale questionnaire was used to collect quantitative data and recruit interview participants for the qualitative portion of the study. The questionnaire was analyzed using descriptive analysis. The qualitative data were analyzed using in vivo coding (Saldaña, 2016). Qualitative data were presented using direct quotes from the interviews. This chapter presents the findings gathered through the survey and interviews with participants and is organized according to research question. To maintain confidentiality, interview participants are identified by the following pseudonyms: Blanche, Dorothy, Miles, Rose, Sophia, and Stan.

Teacher Perceptions of the Soundness of Portfolio Evaluation

Research Question 1 sought to understand fine arts teachers' perceptions of the Tennessee Fine Arts Student Growth Portfolio as an element of teacher evaluation. Survey and interview responses were analyzed to answer this question. Quantitative data are presented first, followed by qualitative data that is used to help explain the quantitative findings.

Participants were asked questions about the portfolio model's fairness, utility, feasibility, accuracy, and two questions about their perception of the portfolio model in general. Frequency values for all questionnaire responses indicate that participants have a less than favorable view of the Tennessee Fine Arts Student Growth Portfolio as an element of teacher evaluation. Participants "disagree" that the Tennessee Fine Arts Student Growth Portfolio is a fair, useful, feasible, and accurate means of evaluating fine arts teachers.

Fairness

Fairness Quantitative Findings

Four of the 18 survey items addressed fairness. Frequencies for all fairness survey items can be found in Table 4.1.

Table 4.1

Frequency Ratings of Perceived Value for Fairness Survey Items (N = 28)

Survey Item	Strongly Disagree		Disagree		Neutral		Agree		Strongly Agree	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
Fairness										
The portfolio is a fair means for me to demonstrate my performance of professional standards	5	18	10	36	6	21	6	21	1	4
The portfolio gives me a more prominent role in the evaluation process.	3	10	7	25	7	25	10	36	1	4
The portfolio peer reviewer(s) were qualified to evaluate my teaching.	7	25	7	25	12	43	2	7	0	0
The scoring rubrics on which I was evaluated were fair.	6	21	8	29	8	29	6	21	0	0

Fairness Question 1 asked if the portfolio was a fair means to demonstrate performance of professional standards. The most frequent response ($n = 10$) was “disagree.” Additionally, several participants ($n = 5$) selected “strongly disagree.” These data indicate that of the 28 participants, most ($n = 15$) “disagree” or “strongly disagree” that the portfolio is a fair means to demonstrate their performance of professional standards.

Fairness Question 2 asked if the portfolio gives them a more prominent role in the evaluation process. Although the most frequent response ($n = 10$) was “agree,” it is noteworthy that both “neutral” and “disagree” were selected seven times by participants. Although several of the 28 participants indicated that they “agree” that the portfolio gives them a more prominent role in the evaluation process, a number of participants ($n = 10$) indicated that they “disagree” (n

= 7) or “strongly disagree” ($n = 3$) that the portfolio gives them a more prominent role in the evaluation process.

Fairness Question 3 asked whether they perceived that portfolio peer reviewers were qualified to evaluate their teaching. The most frequent response for this survey item ($n = 12$) was “neutral.” However, a total of 14 participants selected either “strongly disagree” ($n = 7$) or “disagree” ($n = 7$) indicating that half of the 28 respondents do not perceive that the peer reviewers were qualified to evaluate their teaching.

Fairness Question 4 asked participants if the scoring rubrics on which they were evaluated were fair. This survey item had an equal amount of participants select “disagree” ($n = 8$) and “neutral” ($n = 8$). Similar to Fairness Question 3, 14 participants selected “strongly disagree” ($n = 6$) or “disagree” ($n = 8$), indicating that half of the respondents do not agree that the scoring rubrics on which they were evaluated were fair.

Fairness Qualitative Findings

Survey participants also indicated concerns with manipulation and technology in the open-ended question portion of the survey (see Table 4.2).

Table 4.2

Survey Participant’s Perception of the Portfolio’s Fairness

Theme	Description	Sample Statements
Manipulation	Participants spoke to manipulation in the portfolio system	“Teachers can cheat and set themselves up for success”; “Easily manipulated, I can have a student perform bad and have them do it good 10 minutes later”; “easy for teachers to manipulate”; “I can control what is shown to the evaluator”; “Recycling the previous year’s submissions”
Technology	Participants spoke to issues of technology with the portfolio system	“The training we got on the portfolio was not sufficient”; “System is not user friendly”; “Platform keeps changing”; “Quality of training is highly questionable”; “System glitches caused completed work to be lost”

A portion of the interview was designed to lend further understanding to the fairness portion of the survey. Mirroring the survey, fairness interview questions asked if the portfolio

was a fair means of evaluating teacher performance of professional responsibilities. Five of the six interview participants responded “no.” Blanche, who responded “yes”, stated, “I believe that because we are not a multiple-choice type of class, that it’s a more fair [way of] showing what our kids are capable of and what are classes are capable of.” However, Sophia shared her perspective of a burden the portfolio placed on performing arts (i.e., dance, music, theatre) and fine arts teachers. Sophia said, “We [performing arts teachers] kind of got the short end of the stick [subject to unfortunate circumstances]” because of the additional work required outside of regular contract hours to create a portfolio collection. Miles stated that he “believe[d] the concept is [fair] and Sophia was talking about how portfolios had been used successfully in other areas and shared that “It’s not impossible to have a portfolio that is capable of fairly assessing [fine arts teachers].

Manipulation emerged as a theme regarding the fairness of the portfolio as an element of teacher evaluation. Because teachers have complete control over what they submit in their portfolio, “it can be easily manipulated and falsified” (Rose). Dorothy concurred by saying, “We can manipulate it to be what we want it to be rather than maybe what it is.” Sophia, who was a peer reviewer, shared that “fraud was rampant...I could tell that they had been submitting the same student work year after year after year.” Sophia went on to explain that she saw evidence of fraud in video collections where teachers submitted items in which dates were visible in video segments and noted: “I’m evaluating something in 2018 that has a 2016 or a 2015 date on it.” She explained, “there’s not enough of a motivation, other than just being an ethical person, to keep teachers from just submitting the same work year to year.”

Participants also shared accounts of how others manipulated their portfolios. Dorothy said that she knew of teachers who created portfolio collections out of order. Meaning they collected

the Point B (e.g., post assessment) evidence and then went back to complete the Point A (e.g., pre assessment) evidence saying, “people that just created it and said they did it when they didn’t really do it in the way that they said they did.” Rose stated, “People bragged about using the same information...even some of the same videos repeatedly” and said that she

“even heard of somebody having a kid who grew a beard really fast and asking him to shave their beard and do a really awful version of his performance that he knew [well].

And then when his beard grew back, he did the lead part that he’d done in a play to show how much he had grown [as a performer].

Technology was a second theme that emerged as a factor in the ability of the portfolio system to be a fair means of evaluating teachers. Specifically, comments about technology dealt with multiple changes in the portfolio platform and the lack of training provided to teachers.

Stan, Sophia, and Blanche all commented that the portfolio platform changed multiple times in the first several years of its use. Stan said that they “had three or four different portfolio platforms during the first three or four years.” Others agreed, saying, “We’ve had several different brands of portfolios” (Sophia) and “The platform keeps changing” (Blanche).

Participants also commented on the lack of platform-specific training they received. “In the beginning, it was really ambiguous. It was so unclear...we were kind of creating it as we went” (Dorothy). Stan relayed that they “had no consistent training to help us navigate [the change in platforms]” and that he “believe[d] that they [the state] did more harm than good by switching platforms every year for the first three or four years.”

Technology resources and skills needed to produce a portfolio were an additional theme.

Dorothy commented,

If you're being judged by someone from a county that has a bigger budget or more things to do, or whatever, and maybe our county's not where they are, that's not for them to judge me. Nor is it fair for someone who doesn't have the things [e.g., laptop computers, high-quality camera and sound equipment, access to technology integration coaches, budget to purchase project materials for students] that we might have in District D.

Video editing was also mentioned by several participants as a skillset needed to create a portfolio. Rose recounted that one of her colleagues "was hiring students to help her do her work because she was not that tech savvy" and that she had "talked to people who evaluated the portfolios and they were talking about the vast differences in teacher's technology... [and that] they didn't know how to video edit." Several participants acknowledged their own technological shortcomings. Dorothy simply said that she was "not tech savvy" whereas Miles and Stan spoke specifically about video editing: "[editing is] not a skillset I have (Miles) and "I didn't know how to edit" (Stan).

The issue of lack of technology (e.g., video editing) knowledge was echoed by Blanche and Sophia, who had served as peer reviewers. Blanche said, "From low-quality videos to poor sound quality, it was obvious that these teachers had no clue how to create a portfolio" and that "It was clear which teachers had technology support and which ones did not." Sophia's experience was similar and share that,

So frequently it was big video files where I didn't know what I was looking for, where I wasn't able to identify what the teacher was even going for...I'm searching for meaning in this poorly crafted, poorly conceived portfolio."

Fairness Summary

Quantitative data indicate that fine arts teachers have a negative perception of the Tennessee Fine Arts Student Growth Portfolio as a fair element of teacher evaluation. Qualitative data supports these findings and indicate that concerns over manipulation and technology issues contribute to fine arts teachers' negative perception of the portfolio. Specifically, fine arts teachers indicated that the portfolio was not fair because teachers can easily manipulate it. Further, qualitative data indicate that multiple changes to the portfolio platform and a lack of technology resources and skills needed to create a portfolio contributed to fine arts teachers' negative perception of the fairness of portfolio evaluation.

Utility

Utility Quantitative Findings

Five of the 18 survey items addressed utility. Utility Questions 1–3 will be addressed below. Utility Question 4 and 5 will be discussed as part of Research Question 2. Frequencies for all utility items can be found in Table 4.3.

Table 4.3

Frequency Ratings of Perceived Value for Utility Survey Items (N = 28)

Survey Item	Strongly Disagree		Disagree		Neutral		Agree		Strongly Agree	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
Utility										
The portfolio promotes good teaching practices.	7	25	7	25	8	29	4	14	2	7
Development of a portfolio encourages self-reflection about my work.	3	11	3	11	5	18	15	53	2	7
Development of a portfolio helps me to think differently about my work as a teacher.	3	11	6	21	7	25	12	43	0	0

Utility Question 1 asked participants if the portfolio promoted good teaching. The most frequent response ($n = 8$) was “neutral”; however, 14 fine arts teachers selected “strongly disagree” ($n = 7$) or “disagree” ($n = 7$) for this item. This indicates that half of the 28 participants do not agree that the portfolio promotes good teaching.

Utility Question 2 asked whether the development of a portfolio encouraged self-reflection about their work. The most frequent response ($n = 15$) was “agree.” Additionally, two participants selected “strongly agree” for this survey item. It is noteworthy that of all 18 survey items, Utility Question 2 was the only question to receive more than half of the responses on the positive side (i.e., “agree,” “strongly agree”) of the Likert scale. These data indicate that most ($n = 17$) of the 28 participants agree that the development of a portfolio encouraged them to be self-reflective about their work.

Utility Question 3 asked participants if the development of a portfolio helped them to think differently about their work as a teacher. The most frequent response ($n = 12$) was “agree.” However, seven respondents indicated that they were “neutral” with an additional nine respondents selecting either “strongly disagree” or “disagree.” These data indicate that although several fine arts teachers “disagree,” “strongly disagree,” or were “neutral” more respondents agree that the development of a portfolio helped them to think differently about their work as a teacher.

Utility Qualitative Findings

Survey participants also indicated issues with feedback in the open-ended question portion of the survey (see Table 4.4)

Table 4.4*Survey Participant's Perception of the Portfolio's Utility*

Theme	Description	Sample Statements
Feedback	Participants spoke to feedback in the portfolio system	"No feedback, only scores"; "There is no feedback, you get an arbitrary number"; "Feedback is so generic that it is completely useless"; "No actionable suggestions"; "Feedback given is not worth the amount of time necessary for producing a portfolio"

A portion of the interview was designed to lend further understanding to the utility portion of the survey. Mirroring the survey, utility interview questions asked participants if the portfolio was a useful way to demonstrate performance of professional responsibilities and if the development of a portfolio encouraged reflection. Blanche, Miles, and Stan responded in the affirmative that the portfolio was a useful way for teachers to demonstrate their professional responsibilities whereas Dorothy, Rose, and Sophia responded that it could be. However, all interview participants shared concerns specifically related to the depth and timeliness of feedback as well as the system for providing feedback as factors in the ability of the portfolio system to be a useful means of evaluating teachers.

The depth of feedback provided by peer reviewers was either limited to predetermined responses or not provided at all. The first platform used by the state allowed peer reviewers to provide feedback to teachers by selecting from a dropdown box of predetermined responses. Miles indicated that the predetermined responses were inadequate and actually showed a "lack of understanding" of the portfolio collection. Sophia shared that she had "not benefited from any of the feedback that she had ever received [from her peer reviewers]." Stan added that the portfolio feedback system was so broken that it was "impossible for a teacher to know what they were doing wrong." Blanche, too, agreed that the feedback given was not "workable feedback" and

frustration that when feedback was provided it was not done so in a “timely manner where it could be used” to inform decisions about curriculum and instruction.

Subsequent portfolio platforms did not allow peer reviewers to leave any feedback comments for teachers. Blanche, Dorothy, and Rose shared their frustration with the absence of feedback and a seemingly subjective score. “After all this effort, we get it [portfolio scores] back and they’re just numbers” (Rose). Blanche explained that receiving numerical scores without written feedback “doesn’t do anything” to improve her teaching process.

Reflection was a second theme that emerged as a factor in the ability of the portfolio system to be a useful means of evaluating teachers. Although all six interview participants “agreed” that the development of a portfolio did encourage them to be reflective, two participants voiced concerns with other elements of the process impeding the portfolio’s use as a reflective tool. Rose expressed that she thinks the portfolio can help teachers reflect and improve their teaching; however, the portfolio has become “that one more thing” [that teachers have to do in addition to their many other responsibilities]. Stan’s frustration was with the grading of the portfolio, noting, “If the portfolio is graded objectively I believe it would be very useful for reflection.” This implies that Stan does not have faith in the scores he has received and, therefore, does not use the portfolio to reflect on his teaching.

Utility Summary

Quantitative data indicate that fine arts teachers have a negative perception of the utility of the Tennessee Fine Arts Student Growth Portfolio as an element of teacher evaluation. Although survey participants indicated that the development of a portfolio encouraged self-reflection, these data indicate that completing a portfolio did not encourage them to change their instructional practices. Qualitative data supports these findings and indicates that the depth and

timeliness of feedback provided by peer reviewers contribute to fine arts teachers' negative perception of the utility of portfolio evaluation.

Feasibility

Feasibility Quantitative Findings

Three of the 18 survey items addressed feasibility. Frequencies for all feasibility items can be found in Table 4.5.

Table 4.5

Frequency Ratings of Perceived Value for Feasibility Survey Items (N = 28)

Survey Item	Strongly Disagree		Disagree		Neutral		Agree		Strongly Agree	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
Feasibility										
The time required to review portfolios by peer reviewers is reasonable.	5	18	11	39	8	29	4	14	0	0
The time required by me to develop a portfolio is reasonable.	8	29	11	39	5	18	4	14	0	0
The portfolio is a practical strategy for me to demonstrate my performance of professional standards.	8	29	5	18	8	29	6	21	1	3

Feasibility Question 1 asked whether the time required to review portfolios by peer reviewers was reasonable. The most frequent response ($n = 11$) was “disagree.” An additional five participants indicated, “strongly disagree.” Of the 28 participants, eight indicated “neutral” concerning the reasonableness of the time required to review portfolios by peer reviewers. Although a small number of respondents ($n = 4$) “agree” with the survey item, these data indicate that more than half ($n = 16$) of respondents “disagree” or “strongly disagree” that the time required to review portfolios by peer reviewers was reasonable.

Feasibility Question 2 asked fine arts teachers if the time required by them to develop a portfolio was reasonable. The most frequent response ($n = 11$) was “disagree” with an additional eight participants selecting “strongly disagree.” It is worth mentioning that of the 28 study

participants, only four indicated that they “agree” that the time required by them to develop a portfolio was reasonable. These data indicate that more than half ($n = 18$) of participants “disagree” or “strongly disagree” that the time required by them to develop a portfolio was reasonable.

Feasibility Question 3 asked participants if the portfolio was a practical strategy for them to demonstrate their performance of professional standards. The most frequent responses for this survey item were “strongly disagree” ($n = 8$) and “neutral” ($n = 8$). Seven participants selected responses on the positive side (i.e., “agree,” “strongly agree”) of the Likert scale. However, almost half of the 28 respondents ($n = 13$) indicated that they “disagree” or “strongly disagree” that the portfolio is a practical strategy for them to demonstrate their performance of professional standards.

Feasibility Qualitative Findings

Survey participants also expressed concerns with the time and burden of the portfolio in open-ended survey questions (see Table 4.6).

Table 4.6

Survey Participant’s Perception of the Portfolio’s Feasibility

Theme	Description	Sample Statements
Time	Participants spoke to the time requirements of the portfolio system	“Takes a TON of time”; “It eats teacher time that could be better spent in any number of ways”; “It’s very time consuming”; “Hours of preparation”; “Time requirement is large”; “Took an enormous amount of time outside of the school day”; “stayed late and worked weekends to complete”; “It takes time away from my classes progressing through the curriculum”
Efficiency	Participants spoke to the resulting efficiency of the portfolio system	“It was stressful and added more work”; “Feels like one more thing to have to stress about”; “More work on an already busy schedule”; “Stress of worrying whether or not people are going to see what I see”; “Stress is not worth the little extra on your overall score”; “Placed the burden of interpreting the score on me”; “took personal days off work to work on my portfolio”; “Our attention and energies turned from teaching to teaching to the test”;

A portion of the interview was designed to lend further understanding to the feasibility portion of the survey. Mirroring the survey, feasibility interview questions asked participants to what degree the development and feedback processes were feasible. Perceptions were mixed; however, all interview participants expressed concerns with the time associated with the Tennessee Fine Arts Student Growth Portfolio process.

Efficiency emerged as a factor in the ability of the portfolio to be a feasible means of evaluating teachers. Concerns with efficiency fell into two categories: overall time requirement and loss of instructional time. All six participants spoke to the significant amount of time they invested in the portfolio. Blanche, for example, mentioned “being responsible for drama club and keeping up with everything that is done in the theater” were just some of the additional responsibilities that many theatre teachers have to balance when creating portfolios. Dorothy related the already long hours that theatre teachers spend at school: “we come at 6:30 in the morning and we stay until the night, many days, to develop our shows...and to have the portfolio on top is just not good.” Rose shared a similar opinion and used hyperbole when she said that the portfolio “takes 12-and-a-half years and it just makes you want to jump off a bridge in the spring when it’s due.”

Of the six interview participants, Stan and Miles shared a different perspective of the feasibility of the portfolio as an element of teacher evaluation. Miles noted, “When I think of feasible, I think of the word ‘possible’...I think it is possible,” but went on to say that his challenge with the portfolio process was “how much time I invested, and the amount of time that I invested, actually took away from the time that I needed for other educational responsibilities.” Stan did not indicate time as being an issue he shared that “As I became more experienced in

doing the portfolio, it took less and less time to do it...time is really not an issue feasibility-wise...time was a minor thing.”

Sophia expressed an inherent inequity among performing arts and fine arts teachers and the portfolio system, as dance, music, and theatre teachers’ evidence collections consist of audio and video files. She explained, “I can’t take a static picture of a child acting and say this is [a student performing a monologue]” but a visual arts teacher could simply take a static picture and “say, this is Johnny’s skill with light and shadow at the beginning of the year.” Sophia further explained that, for a theatre teacher, she has to “take time out of my lesson plan to record all of my students in a given day.” After recording all of her students Sophia added that, she then has to “dump all that footage, and then I have to go back and I have to pull all those clips together to make [an evidence collection].”

Participants reported that the inefficiency of the portfolio system also extended to students. Sophia acknowledged the loss of “instructional time as a result of evidence collection” placed an undue burden on students. The most significant loss of instructional time occurred at the beginning of a project that would be included in the portfolio. For example, Sophia explained, “I have to take time out of my lesson plan to record all of my students” for a standard monologue project, “I spent [at least one] whole day of class recording everybody reading their monologue at the beginning of the project.”

Blanche and Rose both indicated that when you take a camera out to record the dynamics of the classroom change when you take out a camera for recording purposes. Rose noted, “I don’t think kids act as natural and are as comfortable when they’re being videotaped.” Blanche recounted having to do multiple takes of performances because “students became anxious when they saw the camera come out.” Dorothy spoke to instructional time being lost because of

“taping issues and attendance.” Blanche indicated that she had lost instructional time due to having a transient student population in a “military town.”

Feasibility Summary

Quantitative data indicate that fine arts teachers have a negative perception of the feasibility of the Tennessee Fine Arts Student Growth Portfolio as an element of teacher evaluation. Fine arts teachers indicated that the time required to create develop a portfolio and the time required for portfolio review were not reasonable. Qualitative data supports these findings and indicate that concerns over time and the efficiency of the portfolio process contribute to fine arts teachers’ negative perception of the feasibility of portfolio evaluation.

Accuracy

Accuracy Quantitative Findings

Four of the 18 survey items addressed accuracy. Frequencies for all accuracy items can be found in Table 4.7.

Table 4.7

Mode Ratings of Perceived Value for Accuracy Survey Items (N = 28)

Survey Item	Strongly Disagree		Disagree		Neutral		Agree		Strongly Agree	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
Accuracy										
The portfolio is a means to provide evidence of my fulfillment of professional standards not readily observable.	6	21	7	25	8	29	5	18	2	7
The peer reviewer(s) scores matched my self-score.	6	21	10	36	5	18	7	25	0	0
The portfolio is an accurate reflection of my performance of professional standards.	8	29	11	39	4	14	5	18	0	0
The portfolio provides a richer and more comprehensive picture of my performance of the professional standards than a stand-alone observation.	7	25	4	14	7	25	8	29	2	7

Accuracy Question 1 asked fine arts teachers if the portfolio was a means to provide evidence of their fulfillment of professional standards that were not readily observable (e.g.,

rehearsal process). Most ($n = 13$) of the 28 participants selected “disagree” ($n = 7$) or “strongly disagree” ($n = 6$). These data indicate that almost half of participants do not agree that the portfolio is a means to provide evidence of their fulfilment of professional standards not readily observable.

Accuracy Question 2 asked fine arts teachers if the peer reviewer(s) scores matched their self-score. It is noteworthy that a quarter ($n = 7$) of the 28 respondents “agree” with the survey item. However, slightly more than half ($n = 16$) respondents “disagree” or “strongly disagree” that the peer reviewer(s) scores matched their self-score.

Accuracy Question 3 asked fine arts teachers if the portfolio is an accurate reflection of their performance of professional standards. Although a small number ($n = 5$) of participants “agree” with this survey item, most ($n = 19$) of the 28 respondents “disagree” or “strongly disagree” that the portfolio is an accurate reflection of their performance of professional standards.

Accuracy Question 4 asked fine arts teachers if the portfolio provided a richer and more comprehensive picture of their performance of professional standards than a stand-alone observation. Although the most frequent response ($n = 8$) was “agree,” more respondents ($n = 11$) selected a negative response (i.e., “disagree,” “strongly disagree”), indicating that they do not agree that the portfolio provided a richer and more comprehensive picture of their performance of professional standards than a stand-alone observation.

Accuracy Qualitative Findings

Survey participants also expressed concerns of mistrust and peer reviewer integrity in open-ended survey questions (see Table 4.8).

Table 4.8*Survey Participant's Perception of the Portfolio's Accuracy*

Theme	Description	Sample Statements
Mistrust	Participants spoke to mistrust in the portfolio review process.	"Incompetent people who aren't good teachers themselves grading your work"; Not all evaluators took their job seriously"; "Strength of evaluators"; "Review process can be incredibly subjective"; "Quality of training for reviewers is highly questionable"; "Peer reviewers are not all qualified and their reviews are often rushed";
Peer Reviewer Integrity	Participants spoke to peer reviewer integrity	"Peer review process is terrible, it relies on teacher integrity"; "The review process is questionable"; "Some bragged on the pay check and how fast the work was"

A portion of the interview was designed to lend further understanding to the accuracy portion of the survey. Mirroring the survey, accuracy interview questions asked participants about their understanding of the portfolio's purpose as an evaluative tool, if it provides evidence of performance standards not readily observable, and if they believe it is an accurate and reliable measure of teacher performance.

Some participants responded in the affirmative, that they believed there was a common understanding of the purpose of the portfolio. Stan said, "I think everybody realizes that the portfolio is designed to allow us to have a portion of our LOE [level of overall effectiveness] score that is based on our effect on students [growth]." Dorothy and Miles shared that the purpose of the evaluation process was to help teachers grow. Blanche added that evaluation also serves "to make sure we are being held accountable for the job we're doing." Other participants indicated that they did not believe that there was a common understanding of the purpose of the portfolio. "I don't think it's ever been really crystal clear what the exact purpose was" (Rose). Sophia indicated that there was "really inconsistent communication about the purpose of the portfolio [especially about] how it was designed to be something that teachers could benefit from year to year."

Mistrust in the review process emerged as a theme regarding the accuracy of the portfolio as an element of teacher evaluation. Participants' mistrust fell into three categories: peer reviewer qualifications, scoring subjectivity in the peer reviewer process, and the integrity of peer reviewers. Unlike classroom observations performed by licensed administrators, Rose expressed concern "that we don't know who's evaluating us and what their qualifications are." While attending a portfolio scoring session facilitated by the state, Blanche shared, "I was a scorer [peer reviewer] and I know for a fact there were new teachers scoring...I don't want a brand new teacher scoring my portfolio." Similarly, Rose also expressed doubt in peer reviewer qualifications stating, "I [personally] know some of the people who judged our work and [I am not impressed with] the quality of their teaching."

Interview participants also expressed mistrust in the peer reviewer process because of its subjectivity. Stan said, "The evaluation of the portfolio is incredibly subjective" adding that he and other teachers in his district had submitted the same project with new students and "received different scores" from one year to the next for the same type of project. Dorothy concurred that scores were "skewed" from year to year when similar projects were submitted.

Blanche pointed out that subjectivity might be a result of a "discrepancy in expectation." Blanche shared,

If you have somebody grading [your portfolio] who has been teaching advanced theatre for the majority of their career, and they're looking at a Theatre I [student], they may not see the same things that you see because they're expecting that kid to be like their advanced kid by the time they get to the end of that project.

Others echoed Blanche's sentiment. "There were times when I would receive a score that I didn't feel like that was a fair score, and I didn't feel like it accurately reflected the quality of teaching

that I had done” (Miles). Speaking from the perspective of a peer reviewer, Blanche said, “There’s too much unknown as a scorer, you can see what they presented, but because they may see something or know something that you don’t know as a scorer, you’re not seeing the growth that they see.” Sophia shared,

The results can vary so much based on who you’re randomly assigned [to score your portfolio]. I could look at the art teacher across the hall’s portfolio, and he got a good evaluator who was just cranking it through [scoring portfolios quickly] and assigning 5s and I got someone who’s cranking it through [scoring portfolios quickly] and assigning 2s, and think I’m a better teacher than the art teacher. [Then I think] why am I getting [the scores I am getting]...it’s very demoralizing.

Interview participants also mentioned peer reviewer integrity as a concerning factor of the portfolio review process. Dorothy spoke of hearing “rumors and murmurings of some people who say, ‘well I graded all of mine in an hour, and I just gave them all a 1, [or] ‘I just gave everybody a 3,’ and they didn’t even read [the portfolio].” Rose also shared her doubts of peer reviewer integrity, as peer reviewer compensation is based on the number of portfolio collections they score. Rose recalled hearing people talk about “how many portfolios they’ve checked out [scored] in a day” as they bragged about how the peer reviewer “got [their] \$1,000” for scoring 12 collections that day.

Accuracy Summary

Quantitative data indicate that fine arts teachers have a negative perception of the Tennessee Fine Arts Student Growth Portfolio as an accurate element of teacher evaluation. Responses to the accuracy survey items indicated that although some fine arts teachers indicated that some elements of accuracy (e.g., self-score) were present, the portfolio, as a whole was not

an accurate reflection of their performance. Qualitative data supports these findings and indicates that concerns of mistrust and peer reviewer integrity contribute to fine arts teachers' negative perception of the portfolio. Specifically, fine arts teachers indicated that the perceived lack of qualifications and scorer subjectivity of peer reviewers contributed to fine arts teachers' negative perception of the accuracy of portfolio evaluation.

Teachers' Perceptions by Academic Discipline and Tenure Status

Questionnaire responses were also aggregated by academic discipline and tenure status to determine any variations. When aggregated by academic discipline, frequency ratings for questionnaire responses indicate that fine arts teachers have a negative perception of the Tennessee Fine Arts Student Growth Portfolio as an element of teacher evaluation (see Appendix N). Noteworthy differences among academic disciplines, however, occurred in the attributes of fairness and accuracy and some variation occurred between tenured and non-tenured teachers.

These differences occurred among instrumental music and theatre teachers. Theatre teachers most frequently selected "disagree" across all four fairness survey items whereas instrumental music teachers most frequently selected "neutral" across the same four survey items; vocal music and visual arts teachers were more equally distributed across the four fairness survey items. For Accuracy Item 3, which asked if the portfolio was an accurate reflection of their performance of professional standards, theatre teachers were the only group that did not have any participants respond positively (i.e., "agree" or "strongly agree"). Additionally, theatre teachers ($n = 9$) responded either "disagree" or "strongly disagree" more than any other subgroup. They selected "strongly disagree" ($n = 36$) and "disagree" ($n = 53$) 89 times whereas instrumental music teachers ($n = 8$) selected "neutral" 55 times across the 16 survey items covering the attributes of sound evaluation.

Findings were also aggregated by participants' tenure status (see Appendix O). Frequency ratings for all questionnaire responses indicate that tenured teachers ($n = 22$) have a negative perception of the Tennessee Fine Arts Student Growth Portfolio as an element of teacher evaluation, whereas non-tenured teachers ($n = 6$), hold a more "neutral" view of the portfolio as an element of teacher evaluation. Tenured teachers selected "strongly disagree" ($n = 91$) and "disagree" ($n = 105$) 196 times across the 16 survey items about the attributes of sound evaluation. The six non-tenured teachers selected "neutral" 47 times across the 16 survey items about the attributes of sound evaluation.

Unintended Consequences of Portfolio Evaluation

In addition to themes that emerged under the four attributes of sound evaluation, interview participants made three comments about the Tennessee Fine Arts Student Growth Portfolio model that are salient but do not converge into existing themes. Using saliency analysis allows these non-recurrent codes to be accommodated (i.e., included) without overstretching existing themes (Buetow, 2010). These non-recurrent comments are related to unintended consequences of portfolio evaluation and include the accuracy of the portfolio, morale of fine arts teachers across participating districts, and the potential unintended consequences of the portfolio model

In this instance, one participant shared Miles shared his concern about the portfolio was not always an accurate and reliable measure of teacher performance because, "teachers actually get to hide what they don't do well, and they get to really highlight what they do well." He acknowledged that it was beneficial that teachers are able to put their "best foot forward [make a strong first impression]" however, because of their ability to submit the evidence collections they want, "sometimes they still did not reveal what really needed to be improved."

Sophia pointed out that the burden of the portfolio had an impact on her entire fine arts department. Sophia recounted that teachers “lost planning and instruction time from evidence collection” in addition to negatively affecting morale. Sophia continued, “We got to a point in my arts department where we would just say GLADiS [one of the portfolio platforms used] like you would use the F-word [a curse word]...it was so bad that we had to laugh at it...several teachers I know quit their jobs over this.”

Several interview participants indicated that one of the purposes of the Tennessee Fine Arts Student Growth Portfolio was to give fine arts teacher more control over their level of overall effectiveness scores, especially in districts with poor academic records on state level assessments. Rose shared, “I think that sometimes the people who most needed the boost from the portfolio were the people that were hurt by the portfolio.”

Summary

The purpose of Research Question 1 was to understand fine arts teachers’ perceptions of the Tennessee Fine Arts Student Growth Portfolio as an element of teacher evaluation. Responses to survey items indicated that fine arts teachers had a negative perception of the Tennessee Fine Arts Student Growth Portfolio as a fair, useful, feasible, or accurate element of teacher evaluation. Interview responses support these findings and revealed that issues with evidence manipulation, technology, timeliness and depth of feedback, efficiency, time commitment, and the mistrust and integrity of peer reviewers contributed to fine arts teachers’ negative perception of the portfolio system.

Evaluations Informing Professional Development

Research Question 2 sought to understand how fine arts teachers use their summative evaluation scores to inform their teaching practice and professional development. Survey and

interview responses were analyzed to answer this question. Quantitative data are presented first, followed by qualitative data that is used to help explain the quantitative findings.

Utility to Support Professional Development

Four of the 18 survey items addressed teachers' perception of the Tennessee Fine Arts Student Growth Portfolio in regard to utility as a means of professional development. Two professional development-related items fell under the utility attribute, which was presented within the Research Question 1 findings. The final two items were listed as professional development questions. Frequencies for the two utility items can be found in Table 4.9. Frequencies for the two professional development survey items can be found in Table 4.10.

Table 4.9

Frequency Ratings of Perceived Value for Utility Survey Items (N = 28)

Survey Item	Strongly Disagree		Disagree		Neutral		Agree		Strongly Agree	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
Utility										
Development of a portfolio encourages me to change instructional practices.	5	18	7	25	6	21	10	36	0	0
Development of a portfolio assists me in identifying areas of strength and weakness.	5	18	7	25	9	32	6	21	1	4

Utility Quantitative Findings

Utility Question 4 asked fine arts teachers if the development of a portfolio encouraged them to change instructional practices. The most frequent response ($n = 10$) was “agree.” Twelve participants indicated that they either “strongly disagree” or “disagree” that the development of a portfolio encouraged them to change their instructional practices. These data indicate that although the most frequent response was “agree” more participants “disagree” or “strongly

disagree” that the development of a portfolio encouraged them to change their instructional practices.

Utility Question 5 asked fine arts teachers if the development of a portfolio assisted them in identifying areas of strength and weakness. The most frequent response ($n = 9$) was “neutral.” However, most ($n = 12$) respondents rated this survey item as “strongly disagree” or “disagree.” These data indicate that although the most frequent selection was “neutral,” more fine arts teachers “disagree” or “strongly disagree” that the development of a portfolio assisted them in identifying areas of strength and weakness.

Table 4.10

Frequency Ratings of Perceived Value for Professional Development Items (N = 28)

Survey Item	Strongly Disagree		Disagree		Neutral		Agree		Strongly Agree	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
Professional Questions										
Preparing a portfolio promoted my professional development.	6	21	16	57	3	11	3	11	0	0
The feedback provided by peer reviewer(s) contributed to my professional development.	18	64	7	25	3	11	0	0	0	0

Professional Development Quantitative Findings

Professional Development Question 1 asked fine arts teachers if preparing a portfolio promoted their professional development. Most of the 28 participants selected “strongly disagree” ($n = 6$) or “disagree” ($n = 16$). These data indicate that participants had a negative perception of the portfolio’s ability to promote their professional development.

Professional Development Question 2 asked fine arts teachers if the feedback provided by the peer reviewer(s) contributed to their professional development. Similar to the first professional development survey question, most of the 28 participants selected “strongly

disagree” ($n = 18$) or “disagree” ($n = 7$). These data indicate that the feedback provided by peer reviewer(s) did not contribute to their professional development.

Utility to Support Professional Development Qualitative Findings

A portion of the interview was designed to lend further understanding to the remaining utility and professional development portions of the survey. As a peer reviewer, Sophia explained a struggle she had when evaluating a monologue under the original portfolio management system:

The things that would've been truly valuable for the submitter to hear from me, I was not able to say...there is no way for me to be able to say, “look, you're evaluating this kid on their character's accent when they're not even making eye contact...they're not attempting to do any kind of gesture...there's no posture”...But in this kind of portfolio system, I can [only] say random phrases that aren't going to be able to be usable [applicable] for feedback for a teacher [to improve their practice].

Participants also indicated that the time it took to receive feedback was a significant factor in their perception of the utility of the portfolio system. Participants reported that there were years that portfolio scores were not released until the middle of the fall term. Blanche shared that “a lot of times we'll pick things early on in the first semester to use [in our portfolio]” and begin the collection process early. Because of the delay in receiving feedback, the scores and any potential written feedback is useless. Dorothy added, “It's horrible, we wait and we wait and we wait, and there have been years when we didn't get it [a score] back.”

Blanche explained that, unlike the prompt observation feedback that teachers receive from their administrators, the delay and lack of timely feedback for the portfolio “puts the burden back on me” in terms of selecting projects for the next year's portfolio. Dorothy noted that there

was a shift in her perception of the portfolio system. Dorothy shared that, after initial implementation of the portfolio, “I think it made me a better teacher...but then it became burdensome...it just became such a burden and really affected me.” Dorothy reported being stressed and extremely overwhelmed at the idea of completing a portfolio stating there were times that she would “just cry” to her colleagues. Dorothy’s perception of the portfolio continued to decline over the years and she became discouraged when she “realized that other counties weren’t doing it at all” and she “almost felt like we were being penalized, given more and more and more work, but busy work.”

Professional Development by Academic Discipline and Tenure Status

An examination of survey data aggregated by discipline indicate similar findings (see Appendix P). As indicated previously, instrumental music teachers selected “neutral” ($n = 5$) more than any other discipline. However, it is noteworthy that for Professional Development Question 1 no participants selected “strongly agree.” Additionally, for Professional Development Question 2, no participants selected “agree” or “strongly agree.” Among all 18 survey items, Professional Development Question 2 was the only survey item that did not have a positive response (i.e., “agree,” “strongly agree”).

Findings were also aggregated by participants’ tenure status (see Appendix Q). Survey responses indicate that most tenured and non-tenured teachers have a negative perception of the Tennessee Fine Arts Student Growth Portfolio to inform their professional development. Although non-tenured teachers again selected “neutral” ($n = 5$) more than their tenured counterparts, the remaining responded negatively (i.e., “disagree,” “strongly disagree”). In contrast, most tenured teachers selected “disagree” or “strongly disagree” for both professional development questions. Of note, responses to Professional Development Question 2 were

negative by tenured teachers. Tenured teachers selected “strongly disagree” ($n = 17$) and “disagree” ($n = 5$) 22 times indicating that none of them believed that the feedback provided to them by peer reviewer(s) contributed to their professional development.

Changes in Level of Overall Effectiveness Scores

Two survey items asked participants to self-report changes in their level of overall effectiveness scores during the first 2 years of portfolio implementation (see Appendix R). Aggregated changes in level of overall effectiveness by direction of change can be found in Table 4.11. Of the 28 survey participants, five participants elected not to respond to the survey items about level of overall effectiveness scores. Of note, two participants reported that their level of overall effectiveness decreased both the first and second years of portfolio implementation. Three participants reported that their level of overall effectiveness only decreased after the second year of implementation. Additionally, eight participants reported that their level of overall effectiveness stayed the same during the first two years of portfolio implementation. These data indicate that 16 of the 23 participants’ level of overall effectiveness scores either stayed the same or increased after the second year of portfolio implementation.

Table 4.11

Changes in Level of Overall Effectiveness Scores Year 1–Year 2 (N = 23)

Direction of Change	<i>n</i>
Decrease to Decrease	2
Decrease to Stay the Same	1
Decreased to Increase	3
Increase to Decrease	1
Increase to Stay the Same	2
Increase to Increase	1
Stay the Same to Decrease	3
Stay the Same to Stay the Same	8
Stay the Same to Increase	2

Of the interview participants, Blanche, Dorothy, Rose, and Sophia all reported that their level of overall effectiveness scores stayed the same after two years of portfolio implementation. Miles reported that during the first two years of portfolio implementation his score, “actually went up.” Stan reported that his level of overall effectiveness score “might be one number off in either direction every time” but that his score “stayed the same for the most part.” These findings support those of the quantitative survey data that lower level of overall effectiveness scores did not appear to be negatively associated with participation in the Tennessee Fine Arts Student Growth Portfolio.

Summary

The purpose of Research Question 2 was to understand how fine arts teachers use their summative evaluation scores to inform their teaching practice and professional development. Responses to survey items indicate that fine arts teachers overwhelmingly “disagree” that the Tennessee Fine Arts Student Growth Portfolio contributes to changes in their instructional practices or informs their professional development. Qualitative interview data support these findings and revealed issues with the depth and timeliness of feedback as barriers to fine arts teachers’ using their portfolio scores to inform their professional practice or development.

Finally, the secondary purpose of this study was to understand how fine arts teachers’ level of overall effectiveness scores have changed since the implementation of the Tennessee Fine Arts Student Growth Portfolio as an element of teacher evaluation. Based on participants’ self-reported level of overall effectiveness scores these data indicate that most ($n = 16$) participants’ scores either stayed the same or increased after two years of portfolio implementation. These quantitative findings are supported by interview data in which five of the

six interview participants reported their level of overall effectiveness scores stayed the same or increased after implementation.

Chapter V

Discussion and Recommendations

The primary purposes of this study are to describe (a) fine arts teachers' perceptions of the portfolio model as a summative evaluation measure and (b) to determine how fine arts teachers use their summative evaluation scores to inform their teaching practice. An additional purpose is to understand how teachers' level of overall effectiveness scores have changed since implementing the Tennessee Fine Arts Student Growth Portfolio Model. This chapter provides a discussion of the findings as they relate to the literature, the limitations and delimitations of the study, and implications for future practice and research.

Discussion

Interpretation of findings revealed three major themes about teachers' perceptions of the Tennessee Fine Arts Student Growth Portfolio as an element of teacher evaluation. First, fine arts teachers reported that there was not a common understanding of the purpose of the student growth portfolio. Second, fine arts teachers overwhelmingly chose negative responses regarding the portfolio as a fair, useful, feasible, or accurate element of teacher evaluation. Lastly, fine arts teachers indicated that the portfolio did not inform their professional learning as other elements of teacher evaluation (e.g., teacher observation).

Purpose of Teacher Evaluation

Understanding the purposes of teacher evaluation, and specifically the Tennessee Fine Arts Student Growth Portfolio, is important to understanding teachers' perceptions of the evaluation system. The two primary purposes of teacher evaluation are quality assurance and professional development (Danielson & McGreal, 2000). The findings of the present study show that fine arts teachers understand the primary purpose of teacher evaluation in general, however,

participants reported that there is not a common understanding of the purpose of the Tennessee Fine Arts Student Growth Portfolio. In particular, participants attributed this lack of common understanding to inconsistent messaging from the state, which may relate to teachers' perceptions about the utility of the Tennessee Fine Arts Student Growth Portfolio.

Attributes of Sound Evaluation

The current study found that fine arts teachers had negative perceptions of the Tennessee Fine Arts Student Growth Portfolio as an element of teacher evaluation. With regard to the attributes of sound evaluation, findings from the current study deviated slightly from the findings of previous studies that utilized the Perceived Value of Teacher Portfolios questionnaire (Tucker et al., 2003). Past studies, such as Tucker et al. (2003) and Attinello et al. (2006), included administrators (i.e., evaluators) as participants in their studies; both found that teachers and administrators had positive perceptions of portfolio-based evaluation across all four attributes of sound evaluation. Quantitative findings from Attinello et al. (2006) and Tucker et al. (2003) showed that teachers had more positive perceptions of portfolio evaluation, while each study's qualitative findings found that participants had concerns with the portfolio as a fair, useful, feasible, and accurate means of teacher evaluation. Tucker et al. (2003) cited concerns of accuracy and diminished feasibility because of the time requirement of completing the portfolio process. Attinello et al. (2006), too, found that the time required to complete a portfolio was a disadvantage. Additionally, their participants voiced concerns related to the accuracy of portfolio-based evaluation noting, "ineffective teachers may be compilers of great portfolios and that instead of focusing on reflection and changing behaviors in the classroom, teachers might be perceived as doing tremendous jobs simply because their portfolio was glitzy, yet full of fluff" (Attinello et al., 2006, p. 146).

Conversely, Denison (2008) only included teachers in their study and found that survey data were all rated toward the midpoint of the scale, with qualitative data indicating a much more negative view of the portfolio as a whole, especially for the attributes of accuracy and feasibility. Teachers, for example, noted that portfolios could be falsified and that the time required to complete the portfolio process was not worth the work (Dennison, 2008). Regarding the element of accuracy, Denison (2008) aligns with the findings of Attinello et al. (2006) and Tucker et al. (2003). Overall, however, the present findings align with the qualitative findings of Attinello et al. (2006), Tucker et al. (2003), and Denison (2008) in that fine teachers have a negative perception of portfolio-based teacher evaluation.

Utility to Support Professional Growth

The second research question aimed to understand how fine arts teachers use their evaluation scores to inform their professional practices and development. Across the four attributes of sound evaluation, fine arts teachers' responses indicated that they do not have a positive perception of the portfolio system as an element of teacher evaluation that would inform their instructional practices and professional learning. Although participants indicated several issues explaining their perception of the portfolio system, fine arts teachers' lack of faith in the portfolio evaluation system stem from two major concerns: mistrust in the peer reviewers and the feedback received, if any. This is concerning, as the perception of an evaluation's effectiveness contributes to teachers' motivation to alter their instructional practices (Derrington & Martinez, 2019).

Although peer reviewers who are certified teachers in a fine arts discipline score portfolios, participants shared concern about peer reviewer qualifications. As found in Norris et al. (2017), Robinson (2017), and Tuytens and Devos (2010), participants' perceptions of an

evaluation system's effectiveness are dependent on the confidence they have in their evaluator. Interview participants expressed concerns with the peer reviewers' years of experience and the reviewer's instructional practices. Interestingly, participants identified being scored by a fine arts teacher as a strength. This idea is supported by research that suggests that feedback to encourage teacher growth might best be provided by a content-specific evaluator (Lochmiller, 2016).

The depth and timeliness of feedback were factors in the negative perceptions of the portfolio as an element of teacher evaluation. Interview participants indicated that feedback, if received, was not timely, was nonspecific, and consisted of predetermined responses. This corresponds with the literature, which indicates that effective feedback is actionable, personalized, and timely (Wiggins, 2012). These factors may have contributed to the negative perception that fine arts teachers have about the portfolio as an element of teacher evaluation.

Variations by Academic Discipline

It is noteworthy to mention that of all fine arts disciplines theatre teachers have participated in the portfolio evaluation system the longest. The Tennessee Fine Arts Student Growth Portfolio was piloted by theatre teachers during the 2011-2012 academic year and implemented state-wide during the 2012-2013 academic year. Theatre teachers were the only discipline to participate in the portfolio process until visual arts teachers were added in 2015-2016 followed by music during the 2017-2018 academic year. In addition to utilizing the portfolio model the longest, theatre teachers in districts who implemented the portfolio system early also experienced all three platform changes. These factors may explain why theatre teachers had a more negative perception of the Tennessee Fine Arts Student Growth Portfolio compared to other fine arts disciplines.

Delimitations

Delimitations are attributes that limit the scope of the study placed on the study by the researcher (Simon, 2011). While not positive or negative, delimitations are justifications for the choices of the researcher (Theofanidis & Fountouki, 2018). The delimitations of the study are the problem being studied and the participant eligibility. The researcher works as a secondary school fine arts teacher whose school district participated in the Tennessee Fine Arts Student Growth Portfolio Model. Because of this, the researcher has a vested interest in the portfolio model and its use as an accurate measure of teacher performance. The criteria for participating in the study were also a delimitation because the researcher required educators to teach at the high school level and to have participated in the Tennessee Fine Arts Student Growth Portfolio model for at least 2 years. I chose high school teachers because I am a high school teacher and I wanted to ensure that fine arts teachers participating in the study had a set schedule in which they saw their students consistently as part of a credit-bearing class; this is not always possible for elementary and middle school fine arts teachers. I also wanted to ensure teachers had experienced at least one full cycle (i.e., planning, creation, submission, feedback, reflection) using the portfolio model.

Limitations

Limitations are out of the control of the researcher and may serve as potential weaknesses of a study (Simon, 2011). Study limitations often include research design, data collection and analysis, and study results (Theofanidis & Fountouki, 2018). Limitations should be disclosed by the researcher (Ross & Bibler Zaidi, 2019). The limitations of the study include population, instrumentation, and the recency of evaluation data due to the COVID-19 pandemic.

First, population is a limitation because only six Tennessee school districts used the Tennessee Fine Arts Student Growth Portfolio to evaluate fine arts teachers during the most recent two years that portfolios were submitted. Second, the Tennessee Fine Arts Student Growth Portfolio was introduced on a rolling schedule in which certain disciplines were added over time. Theatre was the first discipline added in 2013-2014, followed by visual arts in 2015-2016, and music in 2016-2017. Additionally, the number of districts participating in the Tennessee Fine Arts Student Growth Portfolio Model has changed from year to year. At the height of its use during the 2017-2018 academic year, 26 districts across the state utilized the portfolio model. Because of this, not all fine arts teachers were eligible to participate in the study. Due to the small sample size, participants are not representative of the general population of teachers, even within the participating counties. Therefore, the findings of this study may not be generalized outside the context of this study.

Another limitation of this study was the instrumentation (i.e., surveys and interviews). Low response rates are a common obstacle for survey research (Coughlan et al., 2013). Response rates for this study may have been low because the researcher faced delays in access to email addresses for four of the five participating districts. In those four districts, the point of contact for the portfolio system also held other district-level leadership positions, which resulted in the extended wait times for responses to email inquiries. Secondly, invitations were sent to potential participants at the beginning of the 2021-2022 academic year. The busyness of beginning of the year activities may have impacted response rates. Finally, the researcher relied on self-reported survey and interview data. Self-reported data are subject to participant bias; in short, participants may modify their responses to fit what they think the researcher wants to discover (Marshall & Rossman, 2016). Individuals may have selective memories, get the timing of events wrong (i.e.,

telescoping), may attribute adverse events to outside forces or positive events to their actions, or exaggerate the significance of actions or events (Connelly, 2013). As a result of the issues mentioned above, systematic bias (i.e., the divergence between self-reported and the true value of data) may occur (Bauhoff, 2011).

Finally, COVID-19 pandemic occurred while this study was being conducted and may impact its external validity. COVID-19 has caused much disruption to education since March 2020. Outcomes of research during historical events may be a threat to external validity because of “specific combinations of stimulus conditions at that time” (Campbell & Stanley, 1963, p. 20). Because of the pandemic’s disruption to daily instruction, for example, the Tennessee General Assembly suspended state testing and teacher evaluations. In turn, Tennessee fine arts teachers did not submit portfolios for the 2019-2020 or the 2020-2021 academic year. Because of this, the most recent data collected from participants were from the 2018-2019 academic school year. Additionally, District F, which employed 170 of the 251 potential participants did not allow out of district research because of COVID-19. This study includes a discussion of the context of the COVID-19 pandemic under which it was conducted, which may increase external validity (Fell et al., 2020).

Implications for Practice and Research

The current study examined fine arts teachers’ perceptions of the Tennessee Fine Arts Student Growth Portfolio as an element of teacher evaluation. Based on the findings of this study, implications for practice include providing adequate and ongoing training for teachers and peer reviewers, and to utilize the portfolio process as part of a multidimensional evaluation system that includes multiple classroom visitations and regular feedback and support. Implications for research include replicating the present study; exploring how teachers

experienced changes in portfolio platform technology; exploring portfolio-based evaluation, teacher morale, and attrition; and exploring the interrater reliability and validity of the Tennessee Fine Arts Student Growth Portfolio.

Implications for Practice

The first implication for practice is related to the frequent mention of technology issues, particularly the changes in portfolio platforms. Teachers and peer reviewers should be provided with adequate and ongoing training that provides them with the requisite knowledge to successfully complete a portfolio. Additionally, training should be provided on technology (e.g., portfolio platform) and the skills needed (e.g., video editing) to complete the portfolio process; a teachers' portfolio score should not be affected by their access to technology nor their technology skills. Trainings could be provided by state and local personnel and include academic coaches, technology integration coaches, curriculum consulting teachers (e.g., fine arts), and lead teachers. The expectation should be set that all of the aforementioned positions contribute to the success of the student growth portfolio model. Finally, peer reviewers should be provided with high-quality training and support that provides them with a thorough understanding of their role in the evaluation process.

A second implication for practice is related to the access of equipment needed to complete the Tennessee Fine Arts Student Growth Portfolio. Participants mentioned knowing teachers who still had “big desktop computers” and other outdated technology (e.g., still cameras). In addition to providing access to physical equipment (e.g., video recording cameras, microphones, scanners), teachers may also require computer software to assist in the planning, creation, and curation of evidence collections for their portfolio. Lastly, training should be provided for any equipment or software needed to complete the portfolio process. Providing

teachers with the tools and training that they need to properly complete a portfolio is imperative for the evaluation process to be effective.

Another implication of practice would be to build time for portfolio development into the workday. Both quantitative and qualitative data both revealed that time, especially time spent during out-of-contract hours (e.g., after school, weekends) contributed to the negative perception of the Tennessee Fine Arts Student Growth Portfolio as an element of teacher evaluation. Unlike other elements of teacher evaluation (e.g., observations, end-of-course exams), which take place during the normal school day, fine arts teachers report spending time outside contract hours to curate a portfolio collection. The creation of a portfolio collection entails selecting student samples from dozens of student work samples, editing video from performance based projects, and matching Point A and Point B samples from student work. Building in time for fine arts teachers, especially performing arts teachers (i.e., dance, music, theatre) during the workday would reduce the amount of out-of-contract time fine arts teachers spend on portfolio-based teacher evaluation and increase their time spent on instructional responsibilities.

Although participants shared concerns about the accuracy, utility, feasibility, and fairness of the Tennessee Fine Arts Student Growth Portfolio, especially in regard to its fairness and accuracy, research supports the use of portfolios as part of a multiple data source evaluation system (Stronge, 1997, 2002; Tucker et al., 2003; Wolf, 1996; Wolf & Dietz, 1998). Utilizing the portfolio process as part of a multidimensional evaluation system that includes multiple classroom visits paired with regular feedback and support is an important part of a comprehensive evaluation system. In addition to classroom visits performed by administrators, utilizing curriculum consulting teachers and content-specific lead teachers, who are intimately

trained in the portfolio process, to perform classroom visits would provide teachers with another level of support and feedback as they develop, collect, and prepare their portfolio.

Additionally, participants reported lack of confidence in peer reviewer qualifications as a contributing factor of their negative perception of the portfolio system. Although all peer reviewers are certified in their discipline, and should be teachers with at least a level of overall effectiveness score of 3, participants reported new teachers and low performing teachers serving as peer reviewers. In addition to ensuring that peer reviewers have previously participated in the portfolio process and earned a suitable score on their portfolio, novice teachers, or teachers with less than three years of experience as a full time fine arts teacher, should not be allowed to serve as peer reviewers. All potential peer reviewers should also have an average observation score of at least a 4 or 5. Limiting peer reviewers to those who are experienced with consistent observation and level of overall effectiveness scores of at least a 4 or 5 might help improve the negative perception fine arts teachers' hold of the peer review process.

A final implication of practice would be for school districts to work with their fine arts teachers to develop an alternative growth measure to represent fine arts teachers' individual growth measure. The Tennessee Department of Education (2021d) instructs each school district to utilize TVASS, a student growth portfolio, or an alternative growth measure approved by the state giving each school district the agency to develop an appropriate measure of student growth for their teachers in non-tested subjects. It is recommended that school districts engage in dialogue with their fine arts teachers, assess their local resources (e.g., technology, content-specific consulting teachers, lead teachers), and collaboratively develop their own fine arts specific alternative growth measure and seek state approval. The alternative growth measure should not only be representative of individual fine arts teachers' impact on student growth but

also allow for release time so that the evaluation process does not create additional out-of-contract work for those in performing arts.

Implications for Research

Future research opportunities include replicating this study to include middle school fine arts teachers. Although not every middle school has a full fine arts department, many offer music and visual arts classes. Including middle school teachers would increase the number of potential participants. Researchers might also explore including elementary school fine arts teachers too. Although the inclusion of elementary school teachers would increase the sample size, researchers should be mindful of the structure of elementary school fine arts teachers' schedules as many do not see their students on a regular basis.

In addition to replicating this study to include middle school teachers to increase the potential sample size, replicating this study to determine the perception of teachers of other portfolio models utilized by the state of Tennessee. In addition to fine arts, Tennessee currently utilizes student growth portfolios as an element of teacher evaluation for prekindergarten-2nd grade, physical education, and world language teachers. Expanding this study to include other portfolio models would help determine if the negative perception of student growth portfolios is an issue only pertaining to fine arts teachers or if the issues lies with the student growth portfolio model itself. Additionally, expanding the study to include a larger sample size that includes a wider range of grade levels, urban and rural population centers, and socioeconomic standing would improve reliability and generalizability of the findings.

Another implication for research is to explore teachers' experiences with changes in portfolio platform technology and the time needed to learn new platforms. The state of Tennessee has changed portfolio platforms three times since its inception. Multiple participants

mentioned issues with the changes in portfolio platform and the “glitchy” nature of portfolio platforms.

A study exploring the impact of portfolio-based teacher evaluation on teacher morale and attrition would also be helpful. Specifically, future research could explore the characteristics of portfolio-based evaluation systems and how those characteristics relate to teacher morale and attrition. Participants mentioned that the portfolio was stressful and burdensome. Teachers even reported that they knew fine arts teachers who quit because they did not want to do a portfolio. Studies have shown that higher rates of burnout and low morale can contribute to a reduction in retention rates (Pratt & Booker, 2014). Research may discover that certain aspects of portfolio-based evaluation may contribute to low teacher morale and attrition.

Finally, future research extending from these findings would include a study to explore the interrater reliability and validity of the Tennessee Fine Arts Student Growth Portfolio to include validating the rubrics used to evaluate portfolio collections. Multiple participants mentioned the subjectivity of peer reviewers in regard to scoring their portfolios. The literature supports the concern teachers have pertaining to interrater reliability and teacher evaluation (Gillespie et al., 1996; Gitomer & Bell, 2016).

References

- Airasian, P. W., & Gullickson, A. R. (1997). *Teacher self-evaluation tool kit*. Corwin Press.
- Andrade, H. G. (2005). Teaching with rubrics: The good, the bad, and the ugly. *College Teaching* 53(1), 27-30. <https://www.jstor.org/stable/27559213>
- American Recovery and Reinvestment Act of 2009, Public Law 111-5 (2009). Authenticated U.S. Government Information: 123 Statute 115. <https://www.congress.gov/111/plaws/publ5/PLAW-111publ5.pdf>
- Attinello, J. R. (2004). *Teachers' and administrators' perceptions of the value of portfolio-based teacher appraisals for evaluation and professional growth* (UMI No. 3120835) [Doctoral dissertation, Indiana University of Pennsylvania]. ProQuest Dissertations and Theses Global.
- Attinello, J. R., Lare, D., & Waters, F. (2006). The value of teacher portfolios for evaluation and professional growth. *National Association of Secondary School Principals*, 90(2), 132-152.
- Boston Latin School (2020). Boston Latin School History. https://www.bls.org/apps/pages/index.jsp?uREC_ID=206116&type=d
- Ballafkih, A. H., & Van Middelkoop, D. (2019). Beliefs about student achievement held by teachers at Dutch universities of applied sciences. *International Journal of Higher Education* 8(5), p. 45-55. <https://doi.org/10.5430/ijhe.v8n5p45>
- Ballou, D., & Springer, M. G. (2015). Using student test scores to measure teacher performance: Some problems in the design and implementation of evaluation systems. *Educational Researcher*, 44(2), 77-86.
- Bandura, A. (1977). *Social learning theory*. Prentice Hall.

- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Prentice-Hall.
- Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist*, 28(2), 117-148.
- Bandura, A. (1994). Self-efficacy. In V. S. Ramachaudran (ed.), *Encyclopedia of human behavior* (Vol. 4.), (pp. 71-81). Academic Press.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. W. H. Freeman.
- Bandura, A., Ross, D., & Ross, S. A. (1961). Transmission of aggression through the imitation of aggressive models. *Journal of Abnormal and Social Psychology*, 63, 575-582
- Barr, A. (1948). The measurement and predication of teaching efficiency: A summary of investigations. *Journal of Experimental Education*, 16(4), 203-283.
- Bauhoff, S. (2011). Systematic self-report bias in health data: Impact on estimating cross-sectional and treatment effects. *Health Services and Outcomes Research Methodology*, 11(1-2), 44-53.
- Beck, J. & Weiland, L. (2001). Teacher portfolios: Pathways to teacher empowerment. *Science Scope*, 24(6), 60-63.
- Bekhet, A. K., & Zauszniewski, J. A. (2012). Methodological triangulation: An approach to understanding data. *Nurse Researcher*, 20(2), 40–43.
<https://doi.org/10.7748/nr2012.11.20.2.40.c9442>
- Birgin, O. & Baki, A. (2007). The use of portfolio to assess student's performance. *Journal of Turkish Science Education*, 4(2), 75-90. <https://files.eric.ed.gov/fulltext/ED504219.pdf>

- Bratcher, R. D. (1998). *Teacher portfolios: The impact on teaching performance and student achievement as perceived by selected teachers in Conroe ISD*. (UMI No. 9830866)
[Doctoral dissertation, Texas A&M]. ProQuest Dissertations and Theses Global.
- Bresciani, M. J., Oakleaf, M., Kolkhorst, F., Nebeker, C., Barlow, J. (2009). Examining design and inter-rater reliability of a rubric measuring research quality across multiple disciplines. *Practical Assessment, Research, and Evaluation*, 14(12).
<https://doi.org/10.7275/1w3h-7k62>
- Brimley, Jr., V., Verstegen, D. A., & Garfield, R. R. (2016). *Financing education in a climate of change* (12th ed.). Pearson.
- Brookhart, S. M. (2013). *How to create and use rubrics for formative assessment and grading*. Alexandria, VA: Association for Supervision & Curriculum Development.
- Bruce, R. F., & Hoehn, L. (1980). *Supervisory practice in Georgia and Ohio*. Paper presented at the Annual Meeting of the Council of Professors of Instructional Supervision, Hollywood, FL.
- Bryant, L. H., & Chittum, J. R. (2013). ePortfolio effectiveness: A(n ill-fated) search for empirical support. *International Journal of ePortfolio*, 3(2), 189-198.
- Buetow, S. (2010). Thematic analysis and its reconceptualization as 'saliency analysis.' *Journal of Health Services, Research, & Policy*, 15(2), 123-125. DOI:10.1258/jhsrp.2009.009081
- Burke, K., Fogarty, R., & Belgrad, S. (1994). *The mindful school: The portfolio connection*. Skylight.
- Campbell, D. M., Cignetti, P. B., Melenyzer, B. J., Nettles, D. H., & Wyman, R. M. (1997). *How to develop a professional portfolio: A manual for teachers*. Pearson.

- Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Rand McNally.
- Center on Great Teachers and Leaders. (2015). *Databases on state teacher and principal evaluation polices*. <http://resource.tqsource.org/stateevaldb/Compare50States.aspx>.
- Chang, C. & Wu, B. (2012). Is teacher assessment reliable or valid for high school students under a web-based portfolio environment? *Journal of Educational Technology & Society*, 15(4), 265-278.
- Chappuis, J., Stiggins, R. J., Chappuis, S., & Arter, J. A. (2012). *Classroom assessment for student learning: Doing it right – using it well* (2nd ed.). Pearson.
- Clancy, M. & Gardner, J. (2017). Using digital portfolios to develop non-traditional domains in special education settings. *International Journal of ePortfolio*, 7(1), 93-100.
- Close, K., Amrein-Beardsley, A., & Collins, C. (2019). Mapping America's teacher evaluation plans under ESSA. *Phi Delta Kappan*, 101(2), 22-26.
- Coleman, E. (1945). The supervisory visit. *Educational Leadership*, 2(4), 164-167.
- Connelly, L. M. (2013). Limitation section. *Medical-Surgical Nursing*, 22(5), 325-336.
- Connally, K., & Tooley, M. (2016). Beyond ratings: Re-envisioning state teacher evaluation systems as tools for professional growth. *New America*.
https://static.newamerica.org/attachments/12740-beyond-ratings-2/NA_BeyondRatingsPaper.87353c50fc4a4411881b576a50535336e.pdf
- Coughlan, M., Cronin, P., & Ryan, F. (2013). Survey research: Process and limitations. *International Journal of Therapy and Rehabilitation*, 16(1), 9-15.
<https://doi.org/10.12968/ijtr.2009.16.1.37935>

- Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). Sage.
- Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). Sage.
- Cubberley, E. P. (1929). *Public school administration: A statement of the fundamental principles underlying the organization and administration of public education* (3rd ed.). Houghton Mifflin Company.
- Danielson, C. (2011). Evaluations that help teachers learn. *The Effective Educator*, 68(4), p. 35-39.
- Danielson, C., & Abrutyn, L. (1997). *Introduction to using portfolios in the classroom*. Association for Supervision and Curriculum Development.
- Danielson, C., & McGreal, T. L. (2000). *Teacher evaluation to enhance professional practice*. Association for Supervision and Curriculum Development.
- Darling-Hammond, L., Amrein-Beardsley, A., Haertel, E., & Rothstein, J. (2012). Evaluating teacher evaluation. *The Phi Delta Kappan*, 93(6), 8-15.
- deMontmollin, J. S. (2018). *A case study of fine arts teachers' perceptions of eportfolio evaluation as professional development*. (Publication No. 13426177) [Doctoral dissertation, Concordia University-Portland]. ProQuest Dissertations and Theses Global.
- de Oliveira Fernandez, A. P., Holanda Ramos, M. F., Costa e Silva, S. S., Furtado Nina, K. C., & Ramos Pontes, F. A. (2016). Overview of research on teacher self-efficacy in social cognitive perspective. *Anales de Psicologia*, 32(3), 793-802.
<http://dx.doi.org/10.6018/analesps.32.3.220171>

- Denison, T. J. (2008). *An assessment of the perceived value of teacher portfolios in the early years enhancement (EYE) program by primary and secondary teachers in Utah's schools* [Doctoral dissertation, Utah State University]. *All Graduate Theses and Dissertations*, 175. <https://digitalcommons.usu.edu/edt/175>
- Derrington, M. L., & Martinez, J. A. (2019). Exploring teachers' evaluation perceptions: A snapshot. *National Association of Secondary School Principals Bulletin*, 103(1), p. 32-50.
- Dewey, J. (1938). *Experience and education*. Macmillan.
- Dewey, J. (1981). *The philosophy of John Dewey* (J. McDermott, Ed.). University of Chicago Press.
- Doherty, K. M., & Jacobs, S. (2015). State of the states: Evaluating teaching, leading and learning. *National Council on Teacher Quality*.
<https://www.nctq.org/dmsView/StateofStates2015>
- Donaldson, M. L., & Papay, J. P. (2014). Teacher evaluation reform: Policy lessons for school principals. *Principal's Research Review*, 9(5), 1-8. https://cepa.uconn.edu/wp-content/uploads/sites/399/2015/11/DonaldsonPapay.PRR_Sept_2014.pdf
- English, D. & Lachlan-Haché, L. (2015). Uncommon measures: Using teacher portfolios in educator evaluation. *American Institutes for Research*.
<https://www.air.org/sites/default/files/Uncommon-Measures-Teacher-Portfolios-Nov-2015.pdf>
- Eskelson, T. C. (2020). How and why formal education originated in the emergence of civilization. *Journal of Education and Learning*, 9(2). 29-47.

- Etikan, I., Alkassim, R., & Abubakar, S. (2015). Comparison of snowball sampling and sequential sampling technique. *Biometrics & Biostatistics International Journal*, 3(1), <http://dx.doi.org/10.15406/bbij.2016.03.00055>
- Fell, M. J., Pagel, L., Chien-fei, C., Goldberg, M. H., Herberz, M., Huebner, G. M., Sareen, S., Hahnel, U. J. J. (2020). Validity of energy social research during and after COVID-19: Challenges, considerations, and responses. *Energy Research & Social Science*, 68, 1-7. <https://doi.org/10.1016/j.erss.2020.101646>
- Firestone, W. A., & Donaldson, M. L. (2019). Teacher evaluation as data use: What recent research suggests. *Educational Assessment, Evaluation and Accountability*, 31, 289-314. <https://doi.org/10.1007/s11092-019-09300-z>
- Freeman, J. J. (1998). *The teaching portfolio as a vehicle for professional growth* (UMI No. 9923833). [Doctoral dissertation, University of New Hampshire]. ProQuest Dissertations and Theses Global.
- Fromer, K., & Michelli, N. M. (2015). Teacher evaluation in the current era: Implications for teacher education policy and pedagogy. *Counterpoints*, 455, 3-17.
- Garet, M. S., Wayne, A. J., Brown, S., Rickles, J., Song, M., & Manzeske, D. (2017). *Impact of providing performance feedback to teachers and principals*. Institute of Education Sciences.
- Gates, K., Hansen, D., & Tuttle, L. (2015). Teacher evaluation in the arts disciplines: Three state perspectives. *Arts Education Policy Review*, 116, 162-175.
- Gillespie, C. S., Ford, K. L., Gillespie, R. D., & Leavell, A. G. (1996). Portfolio assessment: Some questions, some answers, some recommendations. *Journal of Adolescent & Adult Literacy*, 39(6), 480-491.

- Gitomer, D. H., & Bell, C. A. (2016). *Handbook of research on teaching*. American Educational Research Association.
- Goldberg, G. L. (2011). Portfolio assessment as a tool for teacher evaluation. Presentation made to the Maryland Council for Educator Effectiveness. Retrieved from http://archives.marylandpublicschools.org/NR/rdonlyres/69B8F5FC-6A4A-468F-8886-E089A2605AD8/28669/Portfolio_Assessment_Presentation_Outline_051611.pdf
- Green, E. (2014). *Building a better teacher: How teaching works (and how to teach it to everyone)*. W. W. Norton & Company.
- Hamilton, J. (2010). Delaware and Tennessee win first race to the top grants. U.S. Department of Education. <https://www.ed.gov/news/press-releases/delaware-and-tennessee-win-first-race-top-grants>
- Hanover Research. (2012). Best practices for including multiple measures in teacher evaluations. <https://www.shaker.org/Downloads/BestPracMultTeachEval.pdf>
- Hart, H., Healey, K., & Spote, S. E. (2014). Measuring up. *The Phi Delta Kappan*, 95(8), 62-66.
- Haladyna, T. M. (1997). *Writing test items to evaluate higher order thinking*. Allyn & Bacon.
- Henderson, J. L., Davis, R. R., & Day, T. M. (1995). Developing portfolios to document student performance and accomplishments. *NACTA Journal*, 39(4), 18-21.
- Hibler, W. D., & Snyder, J. A. (2015). Teaching matters: Observations on teacher evaluations. *Schools: Studies in Education*, 12(1), 33-47. <https://www.jstor.org/stable/10.1086/680693>
- Hinkle, D. E., Wiersma, W., & Jurs, S. G. (2003). *Applied statistics for the behavioral sciences* (5th ed.). Houghton Mifflin Company.
- Humphry, S. M., & Heldsinger, S. A. (2014). Common structural design features of rubrics may represent a threat to validity. *Educational Researcher*, 43(5), 253-263.

- IBM Corp. Released 2020. IBM SPSS Statistics for Windows, Version 27.0. IBM Corp.
- Johnson, R. B., Onwuegbuzie, A. J., & Turner, L. A. (2007). Toward a definition of mixed methods research. *Journal of Mixed Methods Research, 1*(2), 112-133.
- Joint Committee on Standards for Educational Evaluation. (1988). *Personnel evaluation standards*. Sage.
- Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review, 2*(2), 130-144.
<https://doi.org/10/1016/j.edurev/2007.05.002>
- Katz, M. (1987). *Reconstructing American education*. Harvard University Press.
- Kearley, I. (1997). *The use of teacher portfolios in teacher evaluation* (UMI No. 9955726). [Master's thesis, Memorial University of Newfoundland]. ProQuest Dissertations and Theses Global.
- Korstjens, I. & Moser, A. (2018). Series: Practical guidance to qualitative research. Part 4: Trustworthiness and publishing, *European Journal of General Practice, 24*(1), 120-124.
<http://dx.doi.org/10.1080/13814788.2017.1375092>
- Kratz, H. E. (1896). Characteristics of the best teacher as recognized by children. *The Pedagogical Seminary 3*(3), 413-460. DOI:
<http://dx.doi.org/10.1080/08919402.1896.9943604>
- Livingston, K. (2017). The complexity of learning and teaching: Challenges for teacher education. *European Journal of Teacher Education, 40*(2), 141-143.
<http://dx.doi.org/10.1080/02619768.2017.1296535>
- Lochmiller, C. R. (2016). Examining administrators' instructional feedback to high school math and science teachers. *Educational Administration Quarterly, 52*, p. 75-109.

- Marshall, C., & Rossman, G. B. (2016). *Designing qualitative research* (6th Ed.). Sage.
- Martinez, J. F., Schweig, J., & Goldschmidt, P. (2016). Approaches for combining multiple measures of teacher performance: Reliability, validity, and implications for evaluation policy. *Educational Evaluation and Policy Analysis*, 38(4), 738-756.
- Marzano, R. J. (2012). The two purposes of teacher evaluation. *Educational Leadership*, 30(3).
<http://www.ascd.org/publications/educational-leadership/nov12/vol70/num03/The-Two-Purposes-of-Teacher-Evaluation.aspx>
- Marzano, R. J., Frontier, T., & Livingston, D. (2011). *Effective supervision: Supporting the art and science of teaching*. Association for Supervision and Curriculum Development.
- Marzano, R.J. & Toth, M.D. (2013). *Teacher evaluation that makes a difference: A new model for teacher growth and student achievement*. Association for Supervision and Curriculum Development.
- McCullough, M., English, B., Angus, M. H., & Gill, B. (2015). *Alternative student growth measures for teacher evaluation: Implementation experiences of early-adopting districts*. Regional Educational Laboratory Mid-Atlantic.
- McNelly, T. A. (2002). Evaluations that ensure growth: Teacher portfolios. *Principal Leadership (Middle School Ed.)*, 3(4), 55-59.
- McNergney, R., & Imig, S. (2003). Teacher evaluation. In J. W. Guthrie (Ed.) *Encyclopedia of Education* (p. 2453-2457). Macmillan.
- Medley, D. M. (1979). The Effectiveness of teachers. In P. L. Peterson & H. J. Walberg (Eds.), *Research on teaching: Concepts, findings, and implications*, McCutchan.

- Mehrens, W. (1989). Combining evaluation data from multiple sources. In J. Millman & L. Darling-Hammond (Eds.), *The new handbook of teacher evaluation: Assessment of elementary and secondary school teachers* (p. 322-336), SAGE.
- Melograno, V. J. (2000). Designing a portfolio system for k-12 physical education: A step-by-step process. *Measurement in Physical Education & Exercise Science*, 4(2), 97-116).
- Moskal, B. M. (2002). Recommendations for developing classroom performance assessments and scoring rubrics. *Practical Assessment, Research, and Evaluation*, 8(14), 1-5.
<https://doi.org/10.7275/jz85-rj16>
- Moskal, B. M. & Leydens, J. A. (2000). Scoring rubric development: Validity and reliability. *Practical Assessment, Research & Evaluation*, 7(10), 1-6. <https://doi.org/10.7275/qrm-gg74>
- NCTQ (2016). All states evaluation timeline. National Council on Teacher Quality.
https://www.nctq.org/dmsView/Evaluation_Timeline_Brief_AllStates
- National Council on Teacher Quality (2019). Teacher & principal evaluation policy. National Council on Teacher Quality.
https://www.nctq.org/dmsView/NCTQ_State_of_the_States_2019_print_ready
- NCTQ (2021). Measures of student growth national results. *State Teacher Policy Database*.
<https://www.nctq.org/yearbook/national/Measures-of-Student-Growth-95>
- Nitko, A. J., & Brookhart, S. M. (2010). *Educational assessment of student* (6th Ed.). Pearson.
- Norris, J., van der Mars, H., Kulinna, P., Amrein-Beardsley, A., Kwon, J., & Hodges, M. (2017). Physical education teacher perceptions of teacher evaluation. *The Physical Educator*, 74, 41-62.

- Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic analysis: Striving to meet the trustworthiness criteria. *International Journal of Qualitative Methods*, 16, 1-13, SAGE.
- Painter, B. (2001). Using teacher portfolios. *Educational Leadership*, 58(5), 31-34.
- Panitz, B. (1996). The student portfolio: A powerful assessment tool. *ASEE Prism*, 5(7), 24-29.
- Papay, J. P. (2012). Refocusing the debate: Assessing the purposes and tools of teacher evaluation. *Harvard Educational Review*, 82(1), 123-143.
- Parkes, K. A., Rohwer, D., & Davison, D. (2015). Measuring student music growth with blind-reviewed portfolio: A pilot study. *Bulletin of the Council for Research in Music Education*, 203, 23-44.
- Patino, C. M., & Carvalho Ferreira, J. (2018). Internal and external validity: Can you apply research study results to your patients? *Jornal Brasileiro de Pneumologia*, 44(3), 183. <https://dx.doi.org/10.1590%2FS1806-37562018000000164>
- Patton, M. Q. (1990). *Qualitative evaluation and research methods*. SAGE.
- Perlman, C. C. (2003). Performance assessment: Designing appropriate performance tasks and scoring rubrics. In J.E. Wall & G.R. Walz (Eds.), *Measuring Up: Assessment Issues for Teachers, Counselors, and Administrators* (pp. 497-506). Caps Press.
- Popham, W. J. (1997). What's wrong-and what's right-with rubrics. *Educational Leadership*, 55(2), 72-75.
- Popham, W. J. (2000). *Modern educational measurement: Practical guidelines for educational leaders* (3rd ed.). Allyn and Bacon.
- Popham, W. J. (2013a). Can classroom assessment of student growth be credibly used to evaluate teachers? *The English Journal*, 103(1), 34-39.

- Popham, W. J. (2013b). On serving two masters: Formative and summative teacher evaluation. *Principal Leadership, 13*, 18-22.
- Pratt, T., & Booker, L. (2014). *Teacher retention in Tennessee: Are we keeping our best teachers? Policy brief*. Retrieved from <https://eric.ed.gov/?id=ED560632>
- Reddy, Y. M., & Andrade, H. (2010). A review of rubric use in higher education. *Assessment and Evaluation in Higher Education, 35*(4), 435-448.
- Rezaei, A. R., & Lovorn, M. (2010). Reliability and validity of rubrics for assessment through writing. *Assessing Writing, 15*(1), 18-39.
- Rhodes, T., Chen, H. L., Watson, C. E., & Garrison, W. (2014). Editorial: A call for more rigorous eportfolio research. *International Journal of ePortfolio, 4*(1), 1-5.
- Riley, R. (2010). PDF. Tennessee's Race to the top application: Key provisions. Offices of Research and Education Accountability. Nashville.
https://comptroller.tn.gov/content/dam/cot/orea/advanced-search/orea-reports-2010/2010_OREA_RTKeyProvisions.pdf
- Ritter, G. W., & Barnett, J. H. (2016). Learning on the job: Teacher evaluation can foster real growth. *The Phi Delta Kappan, 97*(7), 48-52.
- Robinson, M. (2017). Music teachers' perceptions of high stakes teacher evaluation. *Arts Education Policy Review, 1-12*. <http://doi.org/10.1080/10632913.2017.1373380>
- Robinson, S. B., (n.d.). Teacher Evaluation: Why it matters and how we can do better. *Frontline Education*. <https://www.frontlineeducation.com/teacher-evaluation/>
- Rolheiser, C., Bower, B., & Stevahn, L. (2000). *The portfolios organizer: Succeeding with student portfolios in your classroom*. Association for Supervision and Curriculum Development.

Ross, E., & Walsh, K. (2019). State of the state 2019: Teacher and principal evaluation policy.

Washington, DC: National Council on Teacher Quality.

https://www.nctq.org/dmsView/NCTQ_State_of_the_States_2019_print_ready

Ross, P. T., & Bibler Zaidi, N. K. (2019). Limited by our limitations. *Perspectives on Medical*

Education, 8(4), 261-264. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6684501/>

Rothstein, J. (2016). Can value-added models identify teachers' impacts? *The Institute for*

Research on Labor Employment. [https://irle.berkeley.edu/can-value-added-models-](https://irle.berkeley.edu/can-value-added-models-identify-teachers-impacts/)

[identify-teachers-impacts/](https://irle.berkeley.edu/can-value-added-models-identify-teachers-impacts/)

Sadler, D. R. (2009). Indeterminacy in the use of preset criteria for assessment and grading in

higher education. *Assessment and Evaluation in Higher Education*, 34, 159-179.

Saldaña, J. (2016). The coding manual for qualitative researchers. SAGE

SCORE (2010). The state of education in Tennessee – 2010. State Collaborative on Reforming

Education. <https://tnscore.org/wp-content/uploads/2017/02/2010-11-State-of-Education-in-TN.pdf>

Scriven, M. (1967). The methodology of evaluation. In R. W. Tyler, R.W. Gagne, & M. Scriven

(Eds.), *American Educational Research Association Monograph Series on Curriculum Evaluation 1: Perspectives on Curriculum Evaluation*. Rand McNally.

Seely, A. (1996). *Professional guide: Portfolio assessment*. Hawker Brownlow Education.

Silveira, J. M. (2013). Portfolios and assessment in music class. *National Association for Music*

Education, 15-24.

Simon, M. K. (2011). *Dissertation and scholarly research: Recipes for success* (2nd ed.).

CreateSpace.

- Skaalvik, E. M., & Skaalvik, S. (2007). Dimensions of teacher self-efficacy and relations with strain factors, perceived collective teacher efficacy, and teacher burnout. *Journal of Educational Psychology, 99*, 611-625. <https://www.doi.org/10.1037/0022-0663.99.3.611>
- Slater, T. F. (1996). Portfolio assessment strategies for grading first-year university physics students in the USA. *Physics Education, 31*, 82-86.
- Smith, E. C., Starratt, G. K., McCrink, C. L., & Whitford, H. (2020). Teacher evaluation feedback and instructional practice self-efficacy in secondary school teachers. *Educational Administration Quarterly, 56*(4), 671-701.
- Steinberg, M. P., & Donaldson, M. L. (2016). The new educational accountability: Understanding the landscape of teacher evaluation in the post-NCLB era. *Education Finance and Policy, 11*(3): 340–359. https://doi.org/10.1162/EDFP_a_00186
- Stoller, A. (2015). Taylorism and the logic of learning outcomes. *Curriculum Studies, 47*(3), 317-333. <http://dx.doi.org/10.1080/00220272.2015.1018328>
- Stronge, J., H. (1997). *Evaluating teaching: A guide to current thinking and best practice*. Corwin.
- Stronge, J., H. (2002). *Qualities of effective teachers*. Association for Supervision and Curriculum Development.
- Sullivan, S. (2016). High-stakes accountability and the demise of formative supervision of instruction: The case of New York City. In J. Glanz & S. Zepeda (Eds.), *Supervision: New perspectives for theory and practice* (pp. 145-163). Roman & Littlefield.
- Sutherland, S. D. (1998). *A systematic examination of portfolio assessment to promote teacher professional growth* (UMI No. MQ38768). [Master's thesis, University of Ottawa]. ProQuest Dissertations and Theses Global.

Szafran, R. (2017). The miscalculation of interrater reliability: A case study involving the AAC&U value rubrics. *Practical Assessments, Research, and Evaluation*, 22(11), 1-7.

<https://doi.org/10.7275/y36w-hg55>

Taylor, F. W. (1911). *The principles of scientific management*. Reprinted 2007, NuVision.

Tennessee Department of Education (2016). *Student growth portfolio model general teacher guidebook*. https://team-tn.org/wp-content/uploads/2016/07/General-2016-17-Student-Growth-Portfolio-Model-Guidebook_Update-12.16.16.pdf

Tennessee Department of Education (2020a). *Growth measures*. <https://team-tn.org/growth-measures/>

Tennessee Department of Education (2020b). *TEAM student growth portfolio resource guide*. <https://team-tn.org/wp-content/uploads/2013/08/Fine-Arts-TEAM-Portfolio-Resource-Guide.pdf>

Tennessee Department of Education (2021a). *Inservice Guidelines*.

<https://www.tn.gov/education/lea-operations/inservice-guidelines.html>

Tennessee Department of Education (2021b). *Portfolio Resources*. <https://team-tn.org/portfolio-resources/>

Tennessee Department of Education (2021c). *Teacher Evaluation*. <https://team-tn.org/teacher-evaluation-2-2/>

Tennessee Department of Education (2021d). *Alternative Growth Measures*. <https://team-tn.org/alternative-measures-of-growth-to-portfolios/>

Tennessee First to the Top Act of 2010, Tenn. Code Ann. §§ 46-1 (2010).

<https://publications.tnsosfiles.com/acts/106/pub/pc0002EOS.pdf>

Tennessee State Board of Education (2018). Teacher and administrator evaluation policy 5.201.

https://www.tn.gov/content/dam/tn/stateboardofeducation/documents/5.201_Teacher_Principal_Evaluation_Policy_7-28-17.pdf

Tennessee Teacher Tenure Law, Tenn. Code Ann. § 49-5-503 (2012).

<https://codes.findlaw.com/tn/title-49-education/tn-code-sect-49-5-504.html>

Thorndike, E. L. (1910). *Educational psychology*. Frank H. Evory & Co. (Original work published in 1903).

Tobin, T. J., Mandernach, B. J., & Taylor, A. H. (2015). *Evaluating online teaching: Implementing best practices*. Jossey-Bass.

Toch, T., & Rothman, R. (2008). *Rush to judgement: Teacher evaluation in public education*. Education Sector.

Theofanidis, D., & Fountouki, A. (2019). Limitations and delimitations in the research process. *Perioperative Nursing*, 7(3), 155-162. <https://doi.org/10.5281/zenodo.2552022>

Tracy, S. J. (1995). How historical concepts of supervision relate to supervisory practices today. *The Clearing House*, 68(5), 320-325.

Tschannen-Moran, M., & Hoy, A. W. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education*, 17, 783-805. [https://www.doi.org/10.1016/S0742-051X\(01\)00036-1](https://www.doi.org/10.1016/S0742-051X(01)00036-1)

Tucker, P. D., & Stronge, J. H. (2005). *Linking teacher evaluation and student learning*. Association for Supervision and Curriculum Development.

Tucker, P. D., Stronge, J. H. & Gareis, C. R. (2002). *Handbook on teacher portfolios for evaluation and professional development*. Eye on Education.

- Tucker, P. D., Stronge, J. H., Gareis, C. R., & Beers, C. S. (2003). The efficacy of portfolios for teacher evaluation and professional development: Do they make a difference? *Education Administration Quarterly*, 39(5), 572-602.
<https://www.doi.org/10.1177/0013161X03257304>
- Tuytens, M., & Devos, G. (2010). The influence of school leadership on teachers' perception of teacher evaluation policy. *Educational Studies*, 36, 521-536.
- United States Department of Education (2012). Elementary and Secondary Education Act flexibility. www.ed.gov/esea/flexibility
- United States Department of Education (2019). Opportunities and responsibilities for state and local report cards under the Elementary and Secondary Education Act of 1965, as amended by the Every Student Succeeds Act.
<https://www2.ed.gov/policy/elsec/leg/essa/report-card-guidance-final.pdf>
- United States Department of Education (2020) Definition of student growth.
<https://www.ed.gov/race-top/district-competition/definitions#:~:text=Student%20Growth%3A%20The%20change%20in,points%20in%20time%2C%20defined%20as%20E2%80%94>
- VERBI Software. (2019). MAXQDA 2020 [computer software]. Berlin, Germany: VERBI Software. Available from maxqda.com.
- Walvoord, B. E., (2010). *Assessment clear and simple: A practical guide for institutions, departments, and general education*. Wiley.
- Wetzel, W. A. (1929). Scientific supervision and curriculum-building. *The School Review*, 37(2), 117-123. <https://www.jstor.org/stable/1078902>

- Wetzel, W. A., (1932). Characteristics of pupil leaders. *The School Review*, 40(7), 532-534.
<http://www.jstor.com/stable/1080749>
- Wiggins, G. (2012). Seven keys to effective feedback. *Educational Leadership*, 70(1), 10-16.
- Williams, E. (1981). Audiovisual Interview Portfolio (U.S. Patent No. 4,255,872). U.S. Patent and Trademark Office. [http://patft.uspto.gov/netacgi/nph-Parser?Sect1=PTO2&Sect2=HITOFF&u=%2Fnethtml%2FPTO%2Fsearch-adv.htm&r=35&f=G&l=50&d=PTXT&p=1&S1=\(\(audiovisual+AND+portfolio\)+AND+Williams\)&OS=audiovisual+AND+portfolio+AND+Williams&RS=\(\(audiovisual+AND+portfolio\)+AND+Williams\)](http://patft.uspto.gov/netacgi/nph-Parser?Sect1=PTO2&Sect2=HITOFF&u=%2Fnethtml%2FPTO%2Fsearch-adv.htm&r=35&f=G&l=50&d=PTXT&p=1&S1=((audiovisual+AND+portfolio)+AND+Williams)&OS=audiovisual+AND+portfolio+AND+Williams&RS=((audiovisual+AND+portfolio)+AND+Williams))
- Wilson, J. P. (2020). *Student growth portfolios for teacher evaluation*. Tennessee Comptroller of the Treasury Office of Research and Education Accountability.
<https://comptroller.tn.gov/content/dam/cot/orea/advanced-search/2020/PortfolioWebsite.pdf>
- Winsor, P., & Ellefson, B. (1995). Professional portfolios in teacher education: An exploration of their value and potential. *The Teacher Educator*, 31(1), 68-91.
- Wolf, K. P. (1996). Developing an effective portfolio. *Educational Leadership* 53(6), 34-37.
- Wolf, K., & Dietz, M. (1998). Teaching portfolios: Purposes and possibilities. *Teacher Education Quarterly*, 25(1), 9-22.
- Wolfe, E. W., Kao, C., & Ranney, M. (1998). Cognitive differences in proficient and nonproficient essay scores. *Written Communication*, 15(4), 465-492.
<http://dx.doi.org/10.1177/0741088398015004002>

Appendix A

Fine Arts Scoring Rubric Indicators

The indicators on the scoring rubric are meant to serve as a guide for peer reviewers. Fine arts rubrics have seven level indicators to guide peer reviewers as they score student growth portfolios. The following are indicator definitions provided by the state to peer reviewers:

- Level 1: Insufficiently (or Below Expectations) – Student/group is unable to demonstrate expectations of the standard at grade/course/development level.
- Level 2: Inconsistently – Student/group demonstrates part-but not all-of the expectations of the standard at grade/development level.
- Level 3: Sufficiently (or At Expectations) – Student/group demonstrates expectations of the standard at grade/development level.
- Level 4: Exceptionally – Student/group meets or exceeds expectations of the standard at grade/development level with few inconsistencies and/or errors.
- Level 5: Masterfully (or Above Expectations) – Student/group exceeds expectations of the standard at grade/development level with no inconsistencies or errors.
- Level 6: Demonstrates one level of growth from the Level 5 criteria descriptor for this particular grade-or course-level expectation.
- Level 7: Demonstrates two levels of growth from the Level 5 criteria descriptor for this particular grade-or course-level expectation. (Tennessee Department of Education, 2020b, p. 8-9)

Appendix B

Tennessee Fine Arts Student Growth Portfolio Structure

Structure of the Fine Arts Portfolio

- Portfolios contain four evidence collections (i.e., student work artifacts)
- Each evidence collection must contain evidence from two points in time
 - Point A: a standards-aligned task administered to collect student work artifacts
 - Point B: a standards-aligned task administered after instruction to collect student work artifacts
- Teachers will self-score each collection at Point A and Point B based on the Tennessee Fine Arts Student Growth Portfolio standards-based rubric
- Types of Sampling
 - Differentiated sample: demonstrates growth for multiple learning levels (i.e., emerging, proficient, and advanced)
 - Group sample: available only in dance, theatre, media arts, and music group samples represent growth for an ensemble or group of three or more students without specific regard to individual learning levels (i.e., emerging, proficient, advanced; Tennessee Department of Education, 2020b, p. 5-6)

District Flexibility Guidelines

It is highly recommended that portfolios represent at least two of the four domains (i.e., Perform/Present, Create, Connect, Respond). Additionally, it is recommended that at least two of the four collections contain evidence of growth for three levels of learners (i.e., differentiated samples). However, individual districts may have additional guidelines or set expectations for the number of domains represented or types of samples (differentiated vs. grouped) in a fine arts teacher's portfolio. Teachers should reach out to the district portfolio lead for additional guidance (Tennessee Department of Education, 2020b, p. 7).

Sample Portfolio Guide from Participating School District

Fine Arts teachers will submit student growth portfolios utilizing the following parameters:

- Teachers will submit four evidence collections that contain student work artifacts from two points in time.
- Teachers may select any of the four domains for their evidence collections, although it is recommended that at least two of the four domains are represented.
- Teachers may select any combination of sampling while composing their evidence collections, although it is recommended that at least two of the four collections contain evidence of growth for three levels of learners (i.e., differentiated samples).

Appendix C

Research Matrix

Research Question (s)	Constructs or Variables	Instrumentation	Data Collection (e.g., timing, frequency)	Data Analysis Method
RQ 1: What are the perceptions of Fine Arts teachers regarding the use of portfolios in teacher evaluation? (survey)	Fairness*, Feasibility*, Accuracy*, Utility*	Survey (all) based on Tucker et al. (2003).	once	Descriptive analysis
	Fairness, Feasibility, Accuracy, Utility	Semistructured Interviews (subset)	once, post-survey analysis	Thematic analysis (Saldaña, 2016)
RQ 2: How do student growth portfolios contribute to the professional learning of fine arts teachers? (survey + focus group)	Utility*	Survey (all) based on Tucker et al. (2003).	Once	Descriptive analysis
	Fairness*, Feasibility*, Accuracy*, Utility*	Semistructured Interviews (subset)	once, post-survey analysis	Thematic analysis (Saldaña, 2016)

*Constructs come from the Joint Committee on Standards for Educational Evaluation, 1998, p. 7

Appendix D

Survey

Thank you for your willingness to participate in this research study on teachers' perception of the Tennessee Fine Arts Student Growth Portfolio. This survey will collect responses to a Perceived Value of Teacher Portfolios Questionnaire, allow you to briefly answer questions about your Level of Overall Effectiveness, and collect demographic data. Please be sure to complete each part and click submit at the end.

All responses will be kept confidential. Only the researcher will have access to your responses.

Informed Consent

Please read the Informed Consent Form.

I have read, understand, and agree to the informed consent form and choose participate in this study

Yes

No

Perceived Value of Teacher Portfolios Questionnaire

This portion of the survey will ask you about various aspects of participating in the Tennessee Fine Arts Student Growth Portfolio Model. Please consider each question carefully and select the answer that best fits your response.

Please complete this survey based off of your 2018-2019 student growth portfolio score. All responses will be kept confidential. Only the researcher will have access to your responses to this survey.

Fairness

The portfolio is a fair means for me to demonstrate my performance of professional standards.

Strongly Agree

Agree

Neutral

Disagree

Strongly Disagree

The portfolio gives me a more prominent role in the evaluation process.

Strongly Agree

Agree

Neutral

Disagree

Strongly Disagree

The portfolio peer reviewer(s) were qualified to evaluate my teaching.

Strongly Agree

Agree
 Neutral
 Disagree
 Strongly Disagree

The scoring rubrics on which I was evaluated were fair.

Strongly Agree
 Agree
 Neutral
 Disagree
 Strongly Disagree

Utility

The portfolio promotes good teaching practices.

Strongly Agree
 Agree
 Neutral
 Disagree
 Strongly Disagree

Development of a portfolio encourages self-reflection about my work.

Strongly Agree
 Agree
 Neutral
 Disagree
 Strongly Disagree

Development of a portfolio helps me to think differently about my work as a teacher.

Strongly Agree
 Agree
 Neutral
 Disagree
 Strongly Disagree

Development of a portfolio encourages me to change instructional practices.

Strongly Agree
 Agree
 Neutral
 Disagree
 Strongly Disagree

Development of a portfolio assists me in identifying areas of strength and weakness.

Strongly Agree
 Agree
 Neutral
 Disagree

Strongly Disagree

Feasibility

The time required to review portfolios by peer reviewers is reasonable.

Strongly Agree
 Agree
 Neutral
 Disagree
 Strongly Disagree

The time required by me to develop a portfolio is reasonable.

Strongly Agree
 Agree
 Neutral
 Disagree
 Strongly Disagree

The portfolio is a practical strategy for me to demonstrate my performance of professional standards.

Strongly Agree
 Agree
 Neutral
 Disagree
 Strongly Disagree

Accuracy

The portfolio is a means to provide evidence of my fulfillment of professional standards not readily observable.

Strongly Agree
 Agree
 Neutral
 Disagree
 Strongly Disagree

The peer reviewer(s) scores matched my self-scores.

Strongly Agree
 Agree
 Neutral
 Disagree
 Strongly Disagree

The portfolio is an accurate reflection of my performance of professional standards.

Strongly Agree
 Agree
 Neutral
 Disagree

Strongly Disagree

The portfolio provides a richer and more comprehensive picture of my performance of the professional standards than a stand-alone observation.

Strongly Agree

Agree

Neutral

Disagree

Strongly Disagree

General

Preparing a portfolio promoted my professional development.

Strongly Agree

Agree

Neutral

Disagree

Strongly Disagree

The feedback provided by peer reviewer(s) contributed to my professional development.

Strongly Agree

Agree

Neutral

Disagree

Strongly Disagree

Open Ended Questions

From your perspective, what are the major advantages of portfolios in the evaluation process?
(Please list the top three reasons.)

From your perspective, what are the major disadvantages of portfolios in the evaluation process?
(Please list the top three reasons.)

LOE Data

Your answers will help me understand the complexities of the TEAM Summative Evaluation process and the impact the Tennessee Fine Arts Student Growth Portfolio Model.

Remember, data about your level of overall effectiveness can be located by logging into TNCompass.

What year(s) has your school/district participated in the Tennessee Fine Arts Student Growth Portfolio Model? (select all that apply)

2012-2013

2013-2014

2014-2015

2015-2016

2016-2017

2017-2018
2018-2019

After completing the Tennessee Fine Arts Student Growth Portfolio for the first time, did your LOE score.

Increase (go up)
Decrease (go down)
Stay the same (no change)

After completing the Tennessee Fine Arts Student Growth Portfolio for a second year, did your LOE score.

Increase (go up)
Decrease (go down)
Stay the same (no change)

Demographic Information

This portion of the survey will collect information about you, your work experience, and about teacher evaluations. Again, only the researcher will have access to this information and all submissions are kept confidential.

About You

Please answer these questions about yourself

What is your age?

18-24
25-34
35-44
45-54
Above 54

What would best describe you?

Asian
Black/African American
Native American
Native Hawaiian or Pacific Islander
White
Two or More
Other
Prefer not to say

Gender

Female
Male
Non-Binary/Third Gender
Prefer not to say

Education Level (highest degree earned)

- Bachelor's Degree
- Bachelor of Fine Arts
- Master's Degree
- Master of Fine Arts
- Education Specialist Degree (EdS)
- Doctorate (EdD/PhD)

About Your Work Experience

Please fill out these questions about your current work experience as a fine arts teacher.

How many years have you been teaching? (Including this year)

Do You Hold Tenure?

- Yes
- No

Name of School District

- Cumberland
- Madison
- Marion
- Montgomery
- Shelby
- Warren

What subject(s) do you teach? Check all that apply.

- Dance
- General Music
- Instrumental Music
- Vocal Music
- Theatre
- Visual Arts
- Media Arts (to include Film)
- Other: _____

What grade(s) do you currently teach?

- Elementary School
- Middle School
- High School
- I teach across multiple grades (i.e., elementary and middle school or middle and high school)

About Teacher Evaluation

Please answer these questions about teacher evaluation.

How many years has your school district participated in the Tennessee Fine Arts Student Growth Portfolio Model?

2-3 years

4-5 years

6-7 years

Have you ever served as a peer reviewer for the Tennessee Fine Arts Student Growth Portfolio Model?

Yes

No

Submitting Survey

Please answer the following questions and then click submit

Would you be interested in participating in a 20-30 minute interview as part of this study?

Interviews will be conducted face-to-face or virtually via Zoom.

Yes

No

Maybe

Contact Information

Please provide the following information.

Last Name

First Name

Email

Phone

Appendix E

Permission to Use Teacher Survey: Perceived Value of Teacher Portfolios

On Jul 7, 2021, at 7:56 AM, Stronge, James H <jhstro@wm.edu> wrote:

Dear Riley,

Please accept this email as permission to use the survey entitled, "Teacher Survey: Perceived Value of Teacher Portfolios," in your dissertation study at Austin Peay State University. You have permission to modify the survey as needed.

It has been so long since the survey was developed, I don't recall the specifics for development of the instrument. Consequently, I don't have available validation information. I'm sorry I can't help with this aspect of your study. Also, I am not aware of other studies that used the instrument.

If feasible, I would like to request an e-copy of the study once you complete it.

Best wishes for success!

James Stronge

Appendix F

Teacher Survey: Perceived Value of Teacher Portfolios (Tucker et al., 2003)

Directions: Please indicate your opinion of the following statements about the use of portfolios as part of a comprehensive teacher evaluation system by placing a check in the appropriate box. The scale used includes: strongly agree (SA), agree (A), disagree (D), strongly disagree (SD), and don't know (DK). All surveys are to be completed anonymously and will be handled in a confidential manner.

Background Information:

School _____ Probationary _____ Tenured _____
 Years of experience w/School District _____ Total years of teaching experience _____

Fairness	SA	A	D	SD	DK
The portfolio is a fair means for me to demonstrate my performance of professional standards.					
The portfolio gives me a more prominent role in the evaluation process.					
The portfolio promotes two-way communication between my administrator and me.					
The portfolio improves my collegial relationships with administrators.					

Usefulness	SA	A	D	SD	DK
The portfolio promotes good teaching practices.					
Development of a portfolio encourages self-reflection about my work.					
Development of a portfolio helps me to think differently about my work as a teacher.					
Development of a portfolio encourages me to change instructional practices.					
Development of a portfolio assists me in identifying areas of strength and weakness.					

Feasibility	SA	A	D	SD	DK
The time required to review portfolios by administrators is reasonable.					
The time required by me to develop a portfolio is reasonable.					
The portfolio is a practical strategy for me to demonstrate my performance of professional standards.					
The portfolio offers additional substance for discussion at the evaluation review conferences.					

Please answer the remaining questions on the back of this sheet.

Teacher Survey: Perceived Value of Portfolios

Accuracy

- The portfolio is a means to provide evidence of my fulfillment of professional standards not readily observable.
- The portfolio helps the principal to know me better.
- The portfolio is an accurate reflection of my performance of professional standards.
- The portfolio provides a richer and more comprehensive picture of my performance of the professional standards.

SA	A	D	SD	DK

General

1. To what degree did your principal use portfolios in evaluating your performance under the new evaluation system?

A great deal _____ **Somewhat** _____ **A little** _____ **Not at all** _____

Comments:

2. To what degree did preparing a portfolio promote your professional development?

A great deal _____ **Somewhat** _____ **A little** _____ **Not at all** _____

Comments:

3. To what degree did the former observation-only evaluation system contribute to your professional growth?

A great deal _____ **Somewhat** _____ **A little** _____ **Not at all** _____

Comments:

4. What are the major advantages of portfolios in the evaluation process from your perspective?

5. What are the major disadvantages of portfolios in the evaluation process from your perspective?

Please return to _____ by _____
(name) (date)

Appendix G

Semistructured Interview Protocol

Script:

Thank you for agreeing to be interviewed for this study. The purpose of this study is to investigate fine arts teachers' perceptions of the Tennessee Fine Arts Student Growth Portfolio as an element of teacher evaluation. This interview, in addition to the survey you have already completed will help me answer my research questions.

I want to ensure I capture what you say accurately so I would like to record our interview with your permission. Do I have your permission to record this interview session?

If at any time you would like clarification of a question, or if you would prefer not to answer the question, please let me know. For the purpose of this interview, please think about your interaction with the Tennessee Fine Arts Student Growth Portfolio and your most recent LOE score that used the portfolio as your student growth measure.

Can you please state your name, the subject you teach, and the county in which you currently teach? Your name and the district you work for will be replaced with pseudonyms during data analysis. I would like to begin by gathering some base information about your LOE scores.

1. How has your LOE score changed since the implementation of the portfolio?
 - a. Has it gone up, down, or stayed the same?
2. Has there been a significant change in your observation scores or your school's achievement measure?
 - a. If so, can you explain what has changed?
3. Have you taught at the same school the entire time data were collected for this study?
 - a. If not, where did you teach before starting at your current school?
 - b. If you have changed schools, how have your observation scores changed?

*The next several questions are about your perception of the **accuracy** of the fine arts portfolio system.*

4. Do you believe fine arts teachers share a common understanding of the purpose of the Tennessee Fine Arts Student Growth Portfolio Model? Why or why not? (RQ1)
 - a. From your perspective, what is the purpose of teacher evaluation? Of the student growth portfolio?
5. Can you describe the types of training provided to teachers by the state and your school district during the implementation of the student growth portfolio model? (RQ1)
 - a. How has this training benefited you as a consumer of the student growth portfolio model? (RQ1)

- i. If there has been no training, what types of training do you feel would be beneficial?
6. Do you believe portfolios provide evidence of performance standards not typically observable in the classroom? In what ways? (RQ1)
 7. Do you believe portfolios are an accurate and reliable measure of teacher performance? How so? (RQ1)

*The next several questions are about your **perception** of the fairness of the fine arts portfolio system.*

8. Do you believe portfolios are a fair means of evaluating teacher performance of professional responsibilities? How so? (RQ1)
9. Do you believe portfolios promote collegiality among teachers? In what ways? (RQ1)
 - a. Prompt for more detail as needed.

*The next several questions are about your **perception** of the utility of the fine arts portfolio system.*

10. Do you believe portfolios are a useful way for teachers to demonstrate their performance of professional responsibilities? How so? (RQ2)
11. Do you believe the development of a portfolio encourage reflection by teachers? In what ways? (RQ2)
 - a. Prompt for more detail as needed.

*The next several questions are about your **perception** of the feasibility of the fine arts portfolio system.*

12. Do you believe the development of a portfolio is a feasible endeavor for teachers as part of the evaluation process? Why or why not? (RQ1)
13. Do you believe the portfolio review process is a feasible part of the evaluation process? (RQ1)
 - a. Prompt for more detail as needed.

*The next several questions are about your **perception** of the fine arts portfolio system in general.*

14. What do you think are some of the advantages of using portfolios as part of a comprehensive teacher evaluation system? (RQ1 & 2)
15. What do you believe are some disadvantages of using portfolios as part of a comprehensive evaluation system?

16. Is there any other information about your experience with the Tennessee Fine Arts Student Growth Portfolio system you would like to share? (RQ2)

Thank you for taking the time to participate in the interview and share information with me about your perception of the Tennessee Fine Arts Student Growth Portfolio Model as a measure of teacher evaluation. I appreciate you, your time, and what you do as a fine arts teacher.

Appendix H

APSU IRB Approval



Date: 08/09/2021

Re 21-030:

TITLE OF PROJECT: Fine Arts Teachers' Perceptions of Student Growth Portfolios as an Element of Teacher Evaluation.

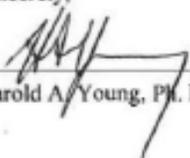
Dear Dr. Prosser and Mr. Braem,

We appreciate your cooperation with the human research review process. This letter is to inform you that study 21-030 the application has been reviewed on an expedited level. It is my pleasure to tell you that your application is approved.

This approval is subject to APSU Policies and Procedures governing human subject research. The IRB reserves the right to withdraw approval if unresolved issues are raised during the review period. Any changes or deviations from the approved protocol must be submitted in writing to the IRB for further review and approval before continuing.

This approval is for one calendar year and a closed study report or request for continuing review is required on or before the expiration date, 08/09/2022. If you have any questions or require further information, you can contact me by phone (931-221-7059) or email (youngh@apsu.edu).

Sincerely,



Harold A. Young, Ph. D. Chair, APIRB

Appendix I

Introductory Email (Fine Arts Coordinator/State Portfolio Lead)

Dear Fine Arts Coordinator/State Portfolio Lead:

My name is Riley Braem and I am a doctoral candidate at Austin Peay State University. I am writing this email to tell you about an opportunity for your fine arts teachers to participate in a research study regarding their perception of the Tennessee Fine Arts Student Growth Portfolio as an element of teacher evaluation. The primary purposes of this study are to describe (a) fine arts teachers' perceptions of the portfolio model as a summative evaluation measure and (b) to determine how fine arts teachers use their summative evaluation scores to inform their teaching practice.

The study has been reviewed and received ethics clearance through Austin Peay State University Institutional Review Board and the [*insert name of the school system here*].

This study will be conducted during the fall of 2021. During that time, I will collect and analyze data related to teachers' perceptions of the Tennessee Fine Arts Student Growth Portfolio. To be eligible to participate, teachers must meet the following criteria:

- hold a valid Tennessee teaching license with a fine arts endorsement (i.e., dance, music, theatre, or visual art),
- have participated in the Tennessee Fine Arts Student Growth Portfolio at least twice.

Would you please provide the contact information of all fine arts teachers in your school district that meet the aforementioned criteria?

Please feel free to reach out to me anytime if you have any questions about the research.

Thank you for your time and assistance.

Sincerely,

W. Riley Braem, MFA
EdD. Candidate, Educational Leadership
Austin Peay State University
Phone: 931.561.8755
Pronouns: he/him/his

Appendix J

Introductory Email (Survey)

Dear Prospective Study Participant:

My name is Riley Braem and I am a doctoral candidate at Austin Peay State University. I am writing this email to invite you to participate in a research study regarding your perception of the Tennessee Fine Arts Student Growth Portfolio as an element of teacher evaluation. The primary purposes of this study are to describe (a) fine arts teachers' perceptions of the portfolio model as a summative evaluation measure and (b) to determine how fine arts teachers use their summative evaluation scores to inform their teaching practice.

The study has been reviewed and received ethics clearance through Austin Peay State University Institutional Review Board and the *[insert name of the school system here]*.

This study will be conducted during the fall of 2021. During that time, I will collect and analyze data related to teachers' perceptions of the Tennessee Fine Arts Student Growth Portfolio. To be eligible to participate, you must meet the following criteria:

- hold a valid Tennessee teaching license with a fine arts endorsement (i.e., dance, music, theatre, or visual art),
- have participated in the Tennessee Fine Arts Student Growth Portfolio at least twice.

If you meet the criteria above and agree to participate, you will be asked to complete a 37-item online survey using Google Forms, which is expected to take approximately 30 minutes to complete. Your survey responses will be anonymous unless you indicate you would be willing to participate in a follow-up interview, in which case you would be asked to provide your contact information.

Your participation in this study is entirely voluntary: you choose whether to participate. If you choose to participate in the study, you can stop your participation at any time. Participation or non-participation will have no bearing on your employment or teacher evaluations.

By participating in this study, you will provide insights into the perceptions of fine arts and the use of student growth portfolios as an element of teacher evaluation. These findings could be used to inform state and local policy makers on improving the Tennessee Fine Arts Student Growth Portfolio as an element of teacher evaluation.

If you are interested in participating, please use this link to complete the survey

<https://forms.gle/KuN6VhDFo7cZex4w5>

Please feel free to reach out to me anytime if you have any questions about the research.

Thank you for your time and consideration.

Sincerely,

W. Riley Braem, MFA

Edd. Candidate, Educational Leadership

Austin Peay State University

Phone: 931.561.8755

Pronouns: he/him/his

Appendix K

Informed Consent (Survey)

INFORMED CONSENT STATEMENT

Fine Arts Teachers' Perceptions of Student Growth Portfolios as an Element of Teacher Evaluation

INTRODUCTION

The Department of Education Specialties at Austin Peay State University supports the practice of protection for human subjects participating in research. The following information is provided to help you decide whether you wish to participate in the present study. You retain the right to refuse to sign this form and not participate in this study. You should be aware that even if you consent to participate in this study, you may withdraw from this study at any time without consequence. If you choose to withdraw from this study, it will not affect your relationship with this department, the services it may provide to you, or Austin Peay State University.

PURPOSE

The primary purposes of this study are to describe (a) fine arts teachers' perceptions of the portfolio model as a summative evaluation measure and (b) to determine how fine arts teachers use their summative evaluation scores to inform their teaching practice. An additional purpose is to understand how teachers' level of overall effectiveness scores have changed since implementing the Tennessee Fine Arts Student Growth Portfolio Model.

PROCEDURES

You are being asked to participate in a survey and, potentially a follow-up interview related to your experience with the Tennessee Fine Arts Student Growth Portfolio Model. After providing your digital signature, you will be taken to the survey. At the end of the survey you will be asked to indicate, by providing your contact information, if you would be willing to participate in a follow-up interview. The link will be open for two weeks. A reminder email will be sent after 5 and 10 days. The survey is expected to take approximately 30 minutes to complete. The follow-up interview will last approximately 30 minutes.

RISKS

The risks associated with participation in this study are no greater than those encountered in daily life.

BENEFITS

A benefit of this study would be to determine the perceived accuracy of the Tennessee Fine Arts Student Growth Portfolio System as an element of teacher evaluation. These findings could be used to inform school districts and state policymakers and their decisions surrounding the evaluation model.

COMPENSATION

Participants will not receive compensation.

PARTICIPANT CONFIDENTIALITY

Any study records that identify you will be kept confidential to the extent possible by law. The records from your participation may be reviewed by people responsible for making sure that research is done properly, including members of the Austin Peay State University Institutional Review Board. Otherwise, records that identify you will be available only to people working on the study, unless you give permission for other people to see the records.

REFUSAL TO SIGN CONSENT

You are not required to sign this Consent and you may refuse to do so without affecting your right to participate in any programs or events of Austin Peay State University or any services you are receiving or may receive from Austin Peay State University. However, if you refuse to sign, you cannot participate in this study.

CANCELLING THIS CONSENT

You may withdraw your consent to participate in this study at any time. If you choose to withdraw from the study before data collection is completed, any collected data will be destroyed and not used.

QUESTIONS ABOUT PARTICIPATION

If you have any questions about the procedures, you may direct them to the principal investigator, W. Riley Braem.

CONSENT

I have read the above information and received a copy of this form. I have had the opportunity to ask questions regarding my participation in this study. I agree to take part in this study as a research participant.

By my digital signature I affirm that I am at least 18 years old.

 Print Participant's Name

Date

 Participant's Signature

Date

RESEARCHER CONTACT INFORMATION

Primary Investigator: W. Riley Braem
 Email: Walter.Braem@cmcss.net
 Phone: 931.561.8755

Faculty Advisor: Dr. Sherri Prosser
 Email: prossers@apsu.edu
 Phone: 931.221.7516

IRB Contact Information

Dr. Harold Young, Chair
Beth Hoilman, IRB Assistant
 irb@apsu.edu
 931.221.7881

Appendix L

Interview Recruitment Email

Dear Prospective Study Participant:

My name is Riley Braem and I am a doctoral candidate at Austin Peay State University. I am writing this email to invite you to participate in a research study regarding your perception of the Tennessee Fine Arts Student Growth Portfolio as an element of teacher evaluation. The purposes of this study are to describe (a) fine arts teachers' perceptions of the portfolio model as a summative evaluation measure and (b) to determine how fine arts teachers use their summative evaluation scores to inform their teaching practice. An additional purpose is to understand how teachers' level of overall effectiveness scores have changed since implementing the Tennessee Fine Arts Student Growth Portfolio Model.

The study has been reviewed and received ethics clearance through Austin Peay State University Institutional Review Board and the *[insert name of the school system here]*.

This study will be conducted during the fall of 2021. During that time, I will collect and analyze data related to teachers' perceptions of the Tennessee Fine Arts Student Growth Portfolio. As part of your survey response, you indicated you would be willing to participate in a follow-up, one-on-one interview. This interview will be conducted via Zoom. Interviews will be audio recorded, and is expected to take approximately 30 minutes to complete. You will have the opportunity to review the audio transcript for accuracy.

All data will be de-identified prior to being analyzed. If you choose to participate in the interview, your name and contact information will not be included in the verbatim transcription of the audio recording. Instead, you will be assigned a participant number, which will match the participant number of your survey. The specifics for confidentiality and data storage are detailed in the informed consent form.

Your participation in this study is entirely voluntary: you choose whether to participate. If you choose to participate in the study, you can stop your participation at any time. Participation or non-participation will have no bearing on your employment or teacher evaluations.

By participating in this study, you will provide insights into the perceptions of fine arts and the use of student growth portfolios as an element of teacher evaluation. These finding could be used to inform state and local policy makers on improving the Tennessee Fine Arts Student Growth Portfolio as an element of teacher evaluation.

If you are interested in participating, please review the attached informed consent form and return it with your signature.

Please feel free to reach out to me anytime if you have any questions about the research.
Thank you for your time and consideration.

Sincerely,
W. Riley Braem, MFA
EdD. Candidate, Educational Leadership
Austin Peay State University
Phone: 931.561.8755
Pronouns: he/him/his

Appendix M

Informed Consent (Interview)

INFORMED CONSENT STATEMENT

Fine Arts Teachers' Perceptions of Student Growth Portfolios as an Element of Teacher Evaluation

INTRODUCTION

The Department of Education Specialties at Austin Peay State University supports the practice of protection for human subjects participating in research. The following information is provided to help you decide whether you wish to participate in the present study. You retain the right to refuse to sign this form and not participate in this study. You should be aware that even if you consent to participate in this study, you may withdraw from this study at any time without consequence. If you choose to withdraw from this study, it will not affect your relationship with this department, the services it may provide to you, or Austin Peay State University.

PURPOSE

The primary purposes of this study are to describe (a) fine arts teachers' perceptions of the portfolio model as a summative evaluation measure and (b) to determine how fine arts teachers use their summative evaluation scores to inform their teaching practice. An additional purpose is to understand how teachers' level of overall effectiveness scores have changed since implementing the Tennessee Fine Arts Student Growth Portfolio Model.

PROCEDURES

You are being asked to participate in an interview related to your perception of the Tennessee Fine Arts Student Growth Portfolio as an element of teacher evaluation. The interview is expected to last approximately 30 minutes and will take place at a time convenient to you. The interview will be conducted via Zoom and be audio recorded. You will have an opportunity to review the interview transcript when it is available. The transcript will be emailed to you and you will have three days to review and respond with corrections.

RISKS

The risks associated with participation in this study are no greater than those encountered in daily life.

BENEFITS

A benefit of this study would be to determine the perceived accuracy of the Tennessee Fine Arts Student Growth Portfolio System as an element of teacher evaluation. These findings could be used to inform school districts and state policymakers and their decisions surrounding the evaluation model.

COMPENSATION

Participants will not receive compensation.

PARTICIPANT CONFIDENTIALITY

Any study records that identify you will be kept confidential to the extent possible by law. The records from your participation may be reviewed by people responsible for making sure that research

is done properly, including members of the Austin Peay State University Institutional Review Board. Otherwise, records that identify you will be available only to people working on the study, unless you give permission for other people to see the records.

REFUSAL TO SIGN CONSENT

You are not required to sign this Consent and you may refuse to do so without affecting your right to participate in any programs or events of Austin Peay State University or any services you are receiving or may receive from Austin Peay State University. However, if you refuse to sign, you cannot participate in this study.

CANCELLING THIS CONSENT

You may withdraw your consent to participate in this study at any time. If you choose to withdraw from the study before data collection is completed, any collected data will be destroyed and not used.

QUESTIONS ABOUT PARTICIPATION

If you have any questions about the procedures, you may direct them to the principal investigator, W. Riley Braem.

CONSENT

I have read the above information and received a copy of this form. I have had the opportunity to ask questions regarding my participation in this study. I agree to take part in this study as a research participant.

Print Participant's Name

Date

Participant's Signature

Date

RESEARCHER CONTACT INFORMATION

Primary Investigator: W. Riley Braem
 Email: Walter.Braem@cmcss.net
 Phone: 931.561.8755

Faculty Advisor: Dr. Sherri Prosser
 Email: prossers@apsu.edu
 Phone: 931.221.7516

IRB Contact Information

Dr. Harold Young, Chair
Beth Hoilman, IRB Assistant

irb@apsu.edu
 (931) 221-7881

Appendix N

Frequency Tables for Perceived Value Survey, by Discipline

Table N.1

Frequency Ratings of Perceived Value for Fairness Survey Items, by Discipline (N = 28)

Survey Item	Strongly Disagree <i>n</i>	Disagree <i>n</i>	Neutral <i>n</i>	Agree <i>n</i>	Strongly Agree <i>n</i>
Fairness					
The portfolio is a fair means for me to demonstrate my performance of professional standards					
Dance	0	1	0	0	0
Music-Instrumental	1	1	4	2	0
Music-Vocal	1	3	0	1	0
Theatre	2	4	2	1	0
Visual Art	1	1	0	2	1
Total	5	10	6	6	1
The portfolio gives me a more prominent role in the evaluation process.					
Dance	0	0	1	0	0
Music-Instrumental	0	2	3	3	0
Music-Vocal	1	1	1	2	0
Theatre	1	4	1	3	0
Visual Art	1	0	1	2	1
Total	3	7	7	10	1
The portfolio peer reviewer(s) were qualified to evaluate my teaching.					
Dance	0	0	1	0	0
Music-Instrumental	1	1	6	0	0
Music-Vocal	2	1	1	1	0
Theatre	3	4	2	0	0
Visual Art	1	1	2	1	0
Total	7	7	12	2	0
The scoring rubrics on which I was evaluated were fair.					
Dance	0	0	1	0	0
Music-Instrumental	1	1	3	3	0
Music-Vocal	2	0	1	2	0
Theatre	2	5	2	0	0
Visual Art	1	2	1	1	0
Total	6	8	8	6	0

Table N.2*Frequency Ratings of Perceived Value for Utility Survey Items, by Discipline (N = 28)*

Survey Item	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
	<i>n</i>	<i>n</i>	<i>n</i>	<i>n</i>	<i>n</i>
Utility					
The portfolio promotes good teaching practices.					
Dance	0	0	1	0	0
Music-Instrumental	1	1	4	2	0
Music-Vocal	2	1	2	0	0
Theatre	3	4	2	0	0
Visual Art	1	1	0	1	2
Total	7	7	9	3	2
Development of a portfolio encourages self-reflection about my work.					
Dance	0	0	1	0	0
Music-Instrumental	0	1	2	5	0
Music-Vocal	1	0	0	4	0
Theatre	1	1	2	5	0
Visual Art	1	1	0	1	2
Total	3	3	5	15	2
Development of a portfolio helps me think differently about my work as a teacher.					
Dance	0	0	1	0	0
Music-Instrumental	0	1	4	3	0
Music-Vocal	1	1	1	2	0
Theatre	1	2	1	5	0
Visual Art	1	2	0	2	0
Total	3	6	7	12	0
Development of a portfolio encourages me to change instructional practices.					
Dance	0	0	1	0	0
Music-Instrumental	1	1	2	4	0
Music-Vocal	1	2	2	0	0
Theatre	2	3	0	4	0
Visual Art	1	1	1	2	0
Total	5	7	6	10	0
Development of a portfolio assists me in identifying areas of strength and weakness.					
Dance	0	0	1	0	0
Music-Instrumental	0	3	4	1	0
Music-Vocal	1	2	1	1	0
Theatre	3	1	2	3	0
Visual Art	1	1	1	1	1
Total	5	7	9	6	1

Table N.3*Frequency Ratings of Perceived Value for Feasibility Survey Items, by Discipline (N = 28)*

Survey Item	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
	<i>n</i>	<i>n</i>	<i>n</i>	<i>n</i>	<i>n</i>
Feasibility					
The time required to review portfolios by peer reviewers is reasonable.					
Dance	0	1	0	0	0
Music-Instrumental	1	1	6	0	0
Music-Vocal	2	1	1	1	0
Theatre	2	4	1	2	0
Visual Art	0	4	0	1	0
Total	5	11	8	4	0
The time required by me to develop a portfolio is reasonable.					
Dance	0	1	0	0	0
Music-Instrumental	2	1	3	2	0
Music-Vocal	3	0	0	2	0
Theatre	2	6	1	0	0
Visual Art	1	3	1	0	0
Total	8	11	5	4	0
The portfolio is a practical strategy for me to demonstrate my performance of professional standards.					
Dance	0	1	0	0	0
Music-Instrumental	2	0	3	3	0
Music-Vocal	2	1	2	0	0
Theatre	3	2	2	2	0
Visual Art	1	1	1	1	1
Total	8	5	8	6	1

Table N.4*Frequency Ratings of Perceived Value for Accuracy Survey Items, by Discipline (N = 28)*

Survey Item	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
	<i>n</i>	<i>n</i>	<i>n</i>	<i>n</i>	<i>n</i>
Accuracy					
The portfolio is a means to provide evidence of my fulfillment of professional standards not readily observable.					
Dance	0	1	0	0	0
Music-Instrumental	2	2	3	1	0
Music-Vocal	1	0	2	2	0
Theatre	2	3	3	1	0
Visual Art	1	1	0	1	2
Total	6	7	8	5	2
The peer reviewer(s) scores matched my self-score.					
Dance	0	1	0	0	0
Music-Instrumental	1	1	3	3	0
Music-Vocal	1	2	1	1	0
Theatre	3	4	0	1	0
Visual Art	1	2	0	2	0
Total	6	10	4	7	0
The portfolio is an accurate reflection of my performance of professional standards.					
Dance	0	1	0	0	0
Music-Instrumental	2	1	2	3	0
Music-Vocal	2	3	0	0	0
Theatre	3	5	1	0	0
Visual Art	1	1	1	2	0
Total	8	11	4	5	0
The portfolio provides a richer and more comprehensive picture of my performance of the professional standards than a stand-alone observation.					
Dance	0	1	0	0	0
Music-Instrumental	1	1	3	3	0
Music-Vocal	1	1	1	2	0
Theatre	3	1	3	2	0
Visual Art	2	0	0	1	2
Total	7	4	7	8	2

Appendix O

Frequency Ratings by Tenure Status

Table O

Frequency Ratings of Perceived Value for All Survey Items, by Tenure Status (N = 28)

Survey Item	Strongly Disagree		Disagree		Neutral		Agree		Strongly Agree	
	T	NT	T	NT	T	NT	T	NT	T	NT
Fairness										
The portfolio is a fair means for me to demonstrate my performance of professional standards	5	0	9	0	3	3	4	3	1	0
The portfolio gives me a more prominent role in the evaluation process.	3	0	7	0	4	2	7	4	1	0
The portfolio peer reviewer(s) were qualified to evaluate my teaching.	7	0	7	0	6	6	2	0	0	0
The scoring rubrics on which I was evaluated were fair.	6	0	8	0	5	3	3	3	0	0
Utility										
The portfolio promotes good teaching practices.	7	0	7	0	5	3	1	3	2	0
Development of a portfolio encourages self-reflection about my work.	3	0	3	0	2	3	12	3	2	0
Development of a portfolio helps me to think differently about my work as a teacher.	3	0	6	0	3	4	10	2	0	0
Development of a portfolio encourages me to change instructional practices.	5	0	7	0	3	3	7	3	0	0
Development of a portfolio assists me in identifying areas of strength and weakness.	5	0	6	1	5	4	5	1	1	0
Feasibility										
The time required to review portfolios by peer reviewers is reasonable.	5	0	9	2	4	4	4	0	0	0
The time required by me to develop a portfolio is reasonable.	7	1	10	1	3	2	2	2	0	0
The portfolio is a practical strategy for me to demonstrate my performance of professional standards.	8	0	3	1	6	2	3	3	1	0
Accuracy										
The portfolio is a means to provide evidence of my fulfillment of professional standards not readily observable.	6	0	5	2	5	3	4	1	2	0
The peer reviewer(s) scores matched my self-score.	6	0	8	2	3	2	5	2	0	0
The portfolio is an accurate reflection of my performance of professional standards.	8	0	9	2	3	1	2	3	0	0
The portfolio provides a richer and more comprehensive picture of my performance of the professional standards than a stand-alone observation.	7	0	2	2	5	2	6	2	2	0

Note. T = tenure ($n = 22$), NT = Non-tenure ($n = 6$)

Appendix P

Professional Development Items, by Discipline

Table P

Professional Development Items, by Discipline (N = 28)

Survey Item	Strongly Disagree <i>n</i>	Disagree <i>n</i>	Neutral <i>n</i>	Agree <i>n</i>	Strongly Agree <i>n</i>
General Questions					
Preparing a portfolio promoted my professional development.					
Dance	0	1	0	0	0
Music-Instrumental	1	5	2	0	0
Music-Vocal	2	3	0	0	0
Theatre	2	4	1	2	0
Visual Art	1	3	0	1	0
Total	6	16	3	3	0
The feedback provided by peer reviewer(s) contributed to my professional development.					
Dance	0	1	0	0	0
Music-Instrumental	4	1	3	0	0
Music-Vocal	4	1	0	0	0
Theatre	7	2	0	0	0
Visual Art	3	2	0	0	0
Total	18	7	3	0	0

Appendix Q

Professional Development Items, by Tenure Status

Table Q

Professional Development Items, by Tenure Status (N = 28)

Survey Item	Strongly Disagree		Disagree		Neutral		Agree		Strongly Agree	
	T	NT	T	NT	T	NT	T	NT	T	NT
General Questions										
Preparing a portfolio promoted my professional development.	6	0	12	4	1	2	3	0	0	0
The feedback provided by peer reviewer(s) contributed to my professional development.	17	1	5	2	0	3	0	0	0	0

Note. T = tenure ($n = 22$), NT = Non-tenure ($n = 6$)

Appendix R

Self-Reported Changes in Level of Overall Effectiveness, Year 1-Year 2

Table R

Self-Reported Changes in Level of Overall Effectiveness Year 1-Year 2 (N = 28)

Participant	LOE Year 1	LOE Year 2
A1	Decreased	Decreased
B1	Decreased	Stayed the Same
B2	Increased	Stayed the Same
B3	No Response	No Response
B4	Stayed the Same	Stayed the Same
C1	Decreased	Increased
C2	Increased	Decreased
D1	No Response	No Response
D2	Stay the Same	Increase
D3	No Response	No Response
D4	Stay the Same	Stay the Same
D5	Stay the Same	Stay the Same
D6	Stay the Same	Decrease
D7	Stay the Same	Stay the Same
D8	Decrease	Increase
D9	Increase	Stay the Same
D10	No Response	No Response
D11	Stay the Same	Stay the Same
D12	No Response	No Response
D13	Decrease	Decrease
D14	Increase	Increase
D15	Stay the Same	Decrease
D16	Decrease	Increase
D17	Stay the Same	Stay the Same
D18	Stay the Same	Decrease
D19	Stay the Same	Increase
D20	Stay the Same	Stay the Same
E1	Stayed the Same	Stayed the Same

Note. Level of overall effectiveness (LOE) is a combination of teacher observation and student growth and achievement scores.