Evaluating the Effectiveness of the Texas Teacher Evaluation and Support System (T-TESS) for Fine Arts Educators

Dissertation-in-Practice Manuscript

Submitted to National University

Sanford College of Education

in Partial Fulfillment of the

Requirements for the Degree of

DOCTOR OF EDUCATION

by

BRIANNA MICHELLE MILLER

San Diego, California

October 2025

Abstract

The problem addressed by this study was the ineffectiveness of current fine arts teacher evaluation practices. The purpose of this action research was to explore perceptions held by evaluators and fine arts teachers regarding the effectiveness of the Texas Teacher Evaluation and Support System (T-TESS) in terms of how well the tool achieves validity and alignment with expected performance goals in the evaluation of fine arts educators. Eisner's theories of educational connoisseurship and criticism, which highlight the connection between content-specific understanding and quality evaluative feedback, served as the conceptual framework. Purposive sampling was used to identify ten fine arts educators and three evaluators who had completed at least one T-TESS evaluation cycle; participants were recruited through email and social media outreach. Participants completed a five-item questionnaire via Qualitrics and a semi-structured interview via Zoom. Clarke and Braun's six-step thematic analysis process was used to analyze the transcribed data; NVivo was used throughout coding. The findings indicated that while participants appreciated T-TESS' structure and focus on best teaching practices, the tool was perceived as ineffective for evaluating fine arts educators due to its inability to address the subjective aspects of fine arts instruction. Participants recommended supplementary arts-specific rubrics, the involvement of content experts, and evaluator participation in fine arts training to increase the perceived accuracy and helpfulness of evaluative feedback. This study contributes to practice by advocating for equity in teacher evaluations and highlighting the need for systems supporting the professional growth of fine arts educators. The findings offer pathways for the design and implementation of more authentic and effective evaluation tools by offering practice-based solutions and emphasizing evaluator connoisseurship.

Acknowledgements

Thousands of seemingly small moments, passing conversations, and encouragements to challenge the status quo all influenced my academic journey, long before I ever began writing this dissertation. Therefore, it would be impossible to thank every individual who helped me arrive at this achievement, though I would be remiss if I did not acknowledge the following:

Thank you to my family for instilling in me a lifelong passion for learning. Your unwavering support in all I have ever set out to do means the world.

I wish to thank the many teachers and leaders who have molded and inspired me throughout my education and teaching career. I would like to extend the sincerest thanks to the teachers and evaluators who were so kind as to volunteer their time and perspectives to this study. Thank you from the bottom of my heart; none of this would be possible without you.

Thank you, Dr. Joshua Herron, for seeing me through the start of this journey. Many thanks go to Dr. Michael Shriner and Dr. Suzanne Dunn for their guidance, encouragement, and patience. I have truly grown as a researcher and writer thanks to your investment in my success.

To my lovely husband, Jeffery—thank you for being my biggest advocate and the best life partner ever. I love you with all my heart. Thank you for helping me flesh out my ideas when I was stuck, for allowing me time to write, and for being by my side throughout this journey.

To my son, Jeffery IV—this very sentence is proof of all the wonder that exists in the world. I never thought I would be writing a dissertation, and I never thought I would be a parent, yet this journey saw me through pregnancy and the first year and a half of your life. Never forget the power of hard work, and do not be afraid to take advantage of all the world has to offer you. I love you forever.

Table of Contents

List of Tables

Section 1: Foundation

Though the practice of evaluating teachers dates back to colonial times (Jewell, 2017), many U.S. states are still modifying and adopting new evaluation systems to increase focus on the continuous improvement of teachers (Fresko & Levy-Feldman, 2023). For example, Texas piloted the Texas Teacher Evaluation and Support System (T-TESS) during the 2014-2015 schoolyear and began to require all Texas schools utilize T-TESS (or a similarly aligned model) to guide teacher evaluations beginning in the 2016-2017 schoolyear (Paufler et al., 2020; TEA, 2022c; TEA, 2022d). Yet, this relatively new evaluation system is far from perfect.

T-TESS is plagued by many of the same faults that have been associated with numerous other teacher evaluation systems. For example, the usage of student growth measures (also known as value-added measures, or VAMs) as an element of evaluation has been contested in several cases. VAMs do not account for individual student differences, nor do they account for the virtually infinite non-teacher-dependent components that could skew one's evaluation (Close & Amrein-Beardsley, 2018). Likewise, when any sort of standardized evaluation is used to evaluate teachers, those who teach non-tested subjects are often assigned the scores of another educator who teaches a tested subject (Gates et al., 2015). For example, a dance teacher (for whose class there is generally no standardized test) might be assigned the same evaluation scores as a peer who teaches a tested content area (such as mathematics or English). While student exam scores in a given area can be reflective of the effectiveness of the instructor of the corresponding content area, exam scores in tested subjects have little to do with the effectiveness of instructors in non-tested areas, meaning that teachers of non-tested subjects are often assigned performance evaluation scores that are not truly reflective of their work (Potter, 2021b).

These issues are only exacerbated for fine arts educators, who are often negatively impacted during evaluations. Fine arts teachers are typically evaluated by administrators who know little about the best practices for such creative and subjective content areas as the fine arts (Fresko & Levy-Feldman, 2023; Lin, 2013; Tutt, 2018). This is suboptimal, given the call for evaluators to have a firm grasp on the content areas of the educators they are assessing (Tutt, 2018)—an evaluator can provide more applicable feedback and can better facilitate an educator's professional development if they are viewing the educator's teachings from an informed perspective (Eisner & Day, 2004). Conventional methods of evaluation also lack the breadth needed to accurately reflect fine arts teachers' efforts, often forcing evaluators to rely on personal judgments, which may incorrectly skew these teachers' evaluation scores (Yang, 2023).

Research focusing on these inadequacies strongly indicates a need to reassess how public-school leaders understand the impact that evaluation systems like T-TESS have on fine arts teachers (Gates et al., 2015; Lin, 2013; Potter, 2021b; Tutt, 2018). While a number of studies have documented administrator perspectives on T-TESS and potential outcomes related to T-TESS (Krimbill et al., 2019; Paufler et al., 2020), there is a paucity of research focused on understanding how T-TESS is perceived by and affects those working in the field of fine arts in public education. This gap represents a research-worthy opportunity.

To address this gap, this study explored perceptions held by evaluators and fine arts teachers regarding the effectiveness and impact of T-TESS on the evaluation of fine arts educators. Though previous studies covered ways in which one group (educators or evaluators) views the evaluation process, this study was unique in its effort to encapsulate the perspectives of both fine arts educators and their evaluators. This study addressed a gap in current research and

contributed to future research efforts aimed at making the evaluation process more equitable for fine arts educators (Jewell 2017; Reid, 2018).

It was anticipated that this study would aid in the establishment of more genuine fine arts instruction in the public-school sector (Potter, 2021b) and reduce biases shown to fine arts teachers (Close & Amrein-Beardsley, 2018). This study may provide teachers in other school districts, evaluators, and arts advocates with new insights into possible initiatives or steps that could be piloted in their school districts to increase the effectiveness of teacher evaluations in the fine arts. Finally, this study has the potential to spark new dialogue around evaluation practices in the fine arts, which may serve to drive future educational practice forward in a positive manner.

Statement of the Problem

The problem addressed by this study was the ineffectiveness of current fine arts teacher evaluation practices (Gates et al., 2015, Potter, 2021b, Tutt, 2018). As a result of current evaluation practices, fine arts educators frequently receive observation scores that are not indicative of their true teaching capabilities; this is largely due to the fact that evaluators often lack the content knowledge needed to fairly evaluate fine arts teachers (Lin, 2013; Tutt, 2018). This is a critical issue in states like Texas, where observations count for 70% percent of one's evaluation score (TEA, 2015). Districts that employ fine arts educators must find ways to aid evaluators and fine arts educators in navigating the evaluation process if they wish to boost the effectiveness and quality of their evaluations.

Continuation of current evaluation practices may lead to disingenuous teaching and teacher burnout in the fine arts, resulting from the feeling that current teaching efforts of fine arts teachers are not being fully understood (Potter, 2021b). Failure to address this issue may

also lead to negative biases shown towards fine arts teachers, who often have a more difficult time demonstrating objective student growth, as well as administrators skewing observation scores to give the illusion of success with mandated evaluation systems (Close & Amrein-Beardsley, 2018; Derrington & Campbell, 2018).

Following the recommendations of Reid (2018) and Krimbill et al. (2019) regarding needed future research in the field of teacher evaluations, this qualitative study called for evaluators and fine arts teachers to share perspectives on the impact and benefits of T-TESS to better understand how this system impacts teacher evaluations. Continued examination of the long-term effects and impact of teacher evaluation systems was necessary to understand how all involved parties understood the evaluation process (Derrington & Campbell, 2018). Therefore, this study aimed to establish an initial set of perspectives, which may be compared to future reflections as conditions surrounding fine arts educator evaluations evolve over time.

Purpose of the Study

The purpose of this action research was to explore perceptions held by evaluators and fine arts teachers regarding the effectiveness of T-TESS in terms of how well the tool achieves validity and alignment with expected performance goals in the evaluation of fine arts educators. Validity was operationally defined as the extent to which performance evaluation rankings were perceived as fair and were supported by underlying reasoning. While the performance evaluations themselves typically incorporate the use of observational evidence, this study aimed to focus on individual perspectives regarding the validity and merit of T-TESS, rather than dissecting the actual data recorded in each evaluation. Therefore, defining validity in terms of evaluator and fine arts educator perspectives on fairness and reasoning made sense for this study. While the target population for this study comprised all fine arts

educators and evaluators working for Texas school districts utilizing T-TESS, the sampling frame comprised fine arts educators and evaluators who work for or have worked for a Texas school district and had at least one year's worth of experience with the T-TESS evaluation system. Focusing on this sampling frame allowed this study to concentrate on the specific population impacted by T-TESS. The researcher utilized purposive sampling to select at least ten fine arts educators and three evaluators from the sampling frame for the study; purposive sampling allowed for identification of individuals who were likely to be rich sources of information (Bedford, 2020). The specified number of participants aligned with Bloomberg's (2023) notion that sample size for qualitative studies should be determined on a case-by-case basis and so as to simultaneously be manageable and allow the researcher to provide a detailed description of the specified phenomenon. However, in alignment with the recommendation of Mertler (2019) regarding the use of purposive sampling in action research, all those who were affected by the phenomenon at hand were given the opportunity to participate, though the starting study sample size was twelve participants (ten educators and two evaluators). This study used a brief a pre-screening questionnaire; participants who met all inclusion criteria completed a questionnaire and one-on-one, semi-structured interview to provide data regarding the impact and benefit of T-TESS. Qualtrics was used to administer the questionnaires, and all interviews were hosted via Zoom. Participants were able to complete the questionnaire any time after completing the T-TESS cycle and after they were confirmed to have met all inclusion criteria, as determined by the pre-screening questionnaire. Interviews were scheduled within two weeks of a participant's completion of the questionnaire. This allowed the researcher time to familiarize herself with responses to the questionnaires to best determine any additional areas of focus needed for the individual

interviews, while accommodating any potential scheduling conflicts that may have been posed by trying to adhere to a shorter window of time for scheduling interviews. Throughout the data collection and analysis process, the researcher maintained a reflective journal to describe conversations with participants, personal reflections as regarded the research questions, and general thoughts that arose in relation to the study. The researcher employed thematic analysis to compile perspectives from each group of participants, with the intent of understanding how fine arts teachers and their evaluators perceived how T-TESS impacts such areas as teacher morale, accuracy of evaluations, and current teaching practices. This study provided insight that may be used in the future by the Texas Education Agency, Texas school districts using T-TESS, and other interested parties to motivate positive and lasting changes in the field of fine arts teacher evaluations. The qualitative research tool NVivo was used to aid in data organization and coding. In embracing a qualitative methodology, this study aimed to make sense of participant feelings and perspectives to craft a holistic picture of the phenomenon at hand (Creswell & Poth, 2018). This study also strove to uphold one of the key aims of action research—to improve practice by moving from habitual, instinctive practice (allowing ineffective fine arts evaluation practices to persist) to informed praxis (engaging in the evaluation process in a more effective manner (Kemmis et al., 2014; Whitehead, 2008)). In this way, an anticipated outcome of this study was inspiration for meaningful and lasting improvement to the way fine arts educators are evaluated in Texas.

Research Questions

RQ1

How do public-school evaluators of fine arts educators perceive the effectiveness and impact of T-TESS in evaluating fine arts educators?

RQ2

How do fine arts educators perceive the effectiveness and impact of T-TESS in evaluating their performance?

RQ3

Reflecting on their experiences with T-TESS, what additional supports and practices, if any, do fine arts educators and evaluators identify as beneficial for improving the long-term use and effectiveness of T-TESS in assessing fine arts educators?

Conceptual Framework

The conceptual framework for this study was based on Elliot Eisner's concepts of educational connoisseurship and criticism. An effective evaluation system presupposes one understands the subject matter that is to be evaluated (Eisner & Day, 2004); therefore, one must ensure evaluators and educators are armed with a detailed knowledge of the fine arts and their best practices so that evaluations may more accurately gauge the quality of fine arts teachers (Anderson, 2023). On a more conceptual level, this notion of using knowledge to inform evaluative practices served to link the design and implementation of T-TESS to the work of Eisner. As Nordin and Wahlström (2019) stated, one must be a connoisseur (someone who appreciates what they are observing) in order to provide meaningful feedback. Evaluation tools provide specific best practices and criteria to look for to ensure all users are using the same metrics when providing this evaluative feedback (Steinberg & Kraft, 2017). However, best practices in the arts often present themselves differently than they would in more traditional content area like mathematics and science, making it difficult for non-connoisseurs of the arts to conduct thorough evaluations of fine arts educators (Tutt, 2018). Therefore, it follows that it would beneficial for non-connoisseurs (such as evaluators with limited arts knowledge) to begin

to think like connoisseurs of the fine arts when it comes to evaluating fine arts educators so as to better identify and provide feedback on these best practices.

Educational connoisseurship can help shape the design of an evaluation tool, as well as the evaluation metrics used within a tool: the tenets held in high regard for educational connoisseurship within a specific area (such as the notion that students demonstrate the ability to listen and self-adjust their vocals in a choir class) should go on to inform the criteria deemed worthy of inclusion in the design of a tool. Similarly, these key aspects can also shape the way in which evaluation metrics are designed and executed, with metrics designed to capture and measure the specific aspects deemed worthy of notice by a connoisseur (or an evaluator acting in the place of a connoisseur). In simplest terms, Eisner's theory suggests T-TESS may be able to be used in a more meaningful manner in the evaluation of fine arts teachers if users of this evaluation system are able to apply a more arts-minded lens to their use of the tool. This idea aligns with the continued call for evaluations to embrace educators as masters within their content areas (Garet et al., 2020; Willey, 2019), as well as for evaluators to have a stronger grasp of the content areas taught by those they must evaluate (Tutt, 2018).

Effective, formalized evaluations are considered crucial from an instructional design (ID) perspective, since improved evaluation practices allow stakeholders and instructional designers to obtain the data needed to make informed decisions related to the overall ID process (Parkhurst & Preskill, 2014). This study explored teacher and evaluator perceptions surrounding an evaluative tool; it was anticipated that analysis of these perceptions would indicate whether or not there was a perceived sense of balance between connoisseurship and criticism in the evaluation process. While embracing an appropriate level of connoisseurship would ensure both evaluators and fine arts educators have a united idea of what tenets are most important in

each fine arts classroom, an appropriate appreciation for the role of criticism in the evaluation process is also important. More specifically, the criticism aspect of the evaluation process associated with this tool calls for evaluators to provide productive, growth-oriented feedback to fine arts educators in writing (using concrete examples from the observation); constructive criticism will also be discussed in detail during an individual, face-to-face meeting between the evaluator and fine arts educator within a week of the observation. While the provision of feedback is required by Texas' overall state evaluation system, it is also important to note that meaningful criticism is a necessity in ensuring the continued growth of professionals like teachers. This notion of balance between connoisseurship and criticism is highly relevant to the field of ID in that Eisner's concept provides a noteworthy perspective from which to build and execute meaningful evaluations, especially when it comes to fine arts education.

Eisner's concept of educational connoisseurship and criticism may be described as a form of qualitative inquiry with applications for the arts and educational research (Uhrmacher et al., 2017). An educator of the arts and a researcher in the field of curriculum studies, Eisner believed schools should organize their curriculum around what he called expressive activities, or educational encounters that allow students to come away with differing understandings of a topic due the students' diverse interests, perspectives, and personal histories. Due to the inherently individual nature of students' learning processes, Eisner argued that the impact of an expressive activity could not be predicted, but rather, would need to be evaluated after a lesson's completion (Donmoyer, 2014). Eisner further asserted that standardized testing leads to uniformity in education and student outcomes, which he claimed was an issue of significant importance in the arts, where students must be encouraged to embrace creative and diverse perspectives (Eisner, 2005). These ideas surrounding expressive activities and the effect of

uniformity on arts education ultimately led to Eisner's conceptualization of educational connoisseurship and criticism.

Educational connoisseurship refers to the idea that professional, educated judgments must be used to determine quality when conducting teacher evaluations (Eisner & Day, 2004; Nordin & Wahlström, 2019). However, one is not considered a connoisseur by virtue of their position within a school, and it can take time to develop the awareness necessary to be a connoisseur of a given subject (Eisner, 2003). One practices connoisseurship by possessing a deep knowledge of the subject matter being taught, and it is necessary for one to be able to accurately take in and understand what is happening, appreciate the subtle nuances within a given encounter, and to be genuinely curious about the phenomenon at hand (Eisner, 2003; Nordin & Wahlström, 2019). A connoisseur must be informed about the qualities of a topic or phenomenon and must possess the acute senses of perception and memory necessary to adapt the awareness needed to make judgment; one should never evaluate a phenomenon of which they do not have adequate awareness (Zerull, 1990). As Eisner (2003) indicated, in the realm of educational evaluation, a connoisseur might address curriculum quality, quality of teaching, student engagement, or a myriad of other components; Eisner was also careful to note that connoisseurship applies both to evaluators and teachers in the sense that an appreciation for quality and educational nuances must be expected of both parties. Regardless of the educational component being assessed, and regardless of the individual in question, educational connoisseurship may be thought to consist of three components: the ability to distinguish between qualities, appreciation resulting from being knowledgeable about a given phenomenon, and evaluation (Nordin & Wahlström, 2019). This third aspect relates to the notion of educational values, and in turn, the concept of educational criticism.

Though the phrase educational criticism may carry negative connotations (Eisner, 2003), educational criticism may actually be thought of as the public side of connoisseurship in the sense that criticism aims to describe, interpret, and evaluate an educational phenomenon in such a way as to provide greater understanding or a new perception (Eisner, 2005; Zerull, 1990). In practicing educational criticism, one begins by describing what they see before taking the initiative to interpret the potential meanings of the phenomenon in question; it is important to note a single event can have a variety of interpretations depending on the different aspects of the description upon which one focuses. The final step of educational criticism involves evaluating the educational significance of a given phenomenon and offering possible alternative interpretations, which may help guide future advice and educational actions (Nordin & Wahlström, 2019). Eisner (2003) stressed that a key component of criticism is the need to edify or instruct. In other words, one cannot hope to provide meaningful evaluative feedback through connoisseurship alone: one must also practice educational criticism if they wish to make the evaluation process an effective one. Likewise, one cannot hope to develop (in the case of an educator) or aid in another's professional development (in the case of an evaluator) without embracing the notion that the teacher in question is capable of learning and growing through continued practice and effort. This belief in the capacity for continued improvement can be fostered through continued connoisseurship, making the concept of educational connoisseurship and criticism somewhat iterative in nature (Eisner, 2004).

This study explored the way in which fine arts educators and their evaluators perceived the effectiveness of T-TESS in terms of how accurately the tool gauges the quality of fine arts teachers. Assuming Eisner's theories regarding connoisseurship and criticism are accurate

(that a balanced connection between educational connoisseurship and criticism improve the evaluation process), a large part of the way in which T-TESS is perceived hinges on how effectively Eisner's concepts are utilized through the design and usage of the evaluation system. This study may help further extend Eisner's work in the field of educational connoisseurship and criticism by examining the perceived effectiveness and impact associated with an evaluation tool based on the relationship between improved knowledge (connoisseurship) and improved assessments of teaching quality (criticism). Similarly, participant perceptions surrounding needs related to the effective and prolonged use of T-TESS in the future (as addressed by the third research question, above) were disseminated through this study. This would allow stakeholders to approach the continuous improvement and usage of the tool from an informed perspective. This cycle of feedback being used to fuel future improvements relates to Eisner's emphasis on reflective practices (Eisner, 2004), as well as to the iterative nature of ID (Parkhurst & Preskill, 2014). These perceptions also contribute to the field of ID by providing insight into how Eisner's concepts of connoisseurship and criticism may be used to inform the design and implementation of evaluative tools in the field of fine arts education.

Definitions of Key Terms

Evaluator

An evaluator is a public-school official, such as a principal, who is tasked with assessing a variety of teacher responsibilities, including instruction, classroom management, school contributions, content knowledge, and collaboration (Reid, 2018).

Evaluator Perceptions

In this study, evaluator perceptions refers to the feelings, thoughts, beliefs, and judgments evaluators have regarding T-TESS concerning the specific areas of measurement of teacher quality, educator morale, and student learning outcomes (Paufler et al., 2020).

Fine Arts

Fine arts should be understood as pertaining the fields of visual and performing arts, including such disciplines as music, dance, theatre, and visual art (Gaunt & Treacy, 2020).

Fine Arts Educator

A fine arts educator is a teacher who oversees instruction in any of the following areas: music, dance, theatre, or visual arts (Gaunt & Treacy, 2020).

Fine Arts Educator Perceptions

In this study, fine arts educator perceptions refers to the feelings, thoughts, pedagogical beliefs, and judgments fine arts educators have regarding T-TESS concerning the specific areas of self-perception, student learning experience, and effect of teaching practices (Matthews & Koner, 2022; Potter, 2021b).

Student Learning Experience

In this study, student learning experience refers to intentional student-centered teaching approaches that allow students to make progress towards and achieve outlined academic goals in a given content area (Potter, 2021b).

T-TESS

T-TESS is the acronym referring to the Texas Teacher Evaluation and Support System, a statewide evaluation system aiming to support educators in the professional development and growth (TEA, 2015).

Teacher Evaluation

Teacher evaluation refers to the processes by which the quality and performance of a teacher are assessed over the course of a given schoolyear (Jewell, 2017).

Review of the Literature

The purpose of this study was to inspire lasting change to the fine arts teacher evaluation system via an exploration of the perspectives of both fine arts educators and their evaluators regarding the manner in which T-TESS achieves validity and alignment with expected performance goals. Following the recommendations of such authors as Krimbill et al. (2019), Paufler (2018), and Reid (2018) regarding needed research into the evaluation process, this study aimed to cultivate a better understanding of the way in which evaluators and fine arts teachers comprehend the impacts and benefits of the teacher evaluation process. Specifically, this study sought to address the need to understand the perceived effectiveness of T-TESS when it comes to ensuring accuracy and quality of fine arts teacher evaluations. As such, a literature review of sources published between 2018 and 2023 was conducted on teacher evaluations in general, as well as on fine arts teacher evaluations, to capture a holistic view of the evaluation process. Using National University's NavigatorSearch and Boolean operators (such as "and," "or," and "not"), scholarly sources related to the teacher evaluation process and fine arts teacher evaluations were identified. Initial keywords, including *teacher evaluation* and *fine arts teacher evaluation*, were utilized to identify initial overarching themes. From here, key phrases were used to further refine and elaborate upon these themes. Key phrases included *performing arts teacher evaluation*, *teacher evaluation effect on morale*, *T-TESS*, *Texas Essential Knowledge and Skills*, and *evaluation systems created by educators and evaluators*.

Current Teacher Evaluation Practices in America

There is much debate across the country concerning the appropriate model that should be used to evaluate public-school teachers (Jewell, 2015). Even in the last fifteen years, much has changed regarding the landscape surrounding evaluations. In 2009, the educational reform Race to the Top (RTTT) encouraged implementing evaluation systems that focused on identifying the contribution of individual teachers to a child's education, as assessed through such assessment measures as standardized test scores (Gitomer & Marshall, 2023). Barely ten years later, the 2016 Every Student Succeeds Act eliminated the federal mandate to link teacher evaluations with student test scores, essentially reversing the impetus of Race to the Top; still, many states continue to rely on teacher evaluation systems that utilize student test scores and value-added models (VAMs (Paufler, 2018)). Amid this changing climate in evaluation recommendations and requirements, numerous lawsuits have arisen regarding such mistakes in the current teacher evaluation system as the misuse of VAMs, a lack of transparency in evaluations, and unreliable ratings (Close & Amrein-Beardsley, 2018; Geiger et al., 2020).

Teacher evaluation systems are currently utilized in all 50 states, though there is variance between states regarding the design of evaluation systems (Steinberg & Kraft, 2017). As of 2019, 42 states required school districts to utilize an evaluation system with three or more performance measures, with the remaining nine states utilizing a two-tiered system. Additionally, 34 states require teacher evaluations to include an objective measure of student growth, and 75% of these states insist standardized test scores must account for some of this data (National Council on Teacher Quality, 2020). In other words, despite objections to current evaluation systems and calls to better understand stakeholders' thoughts on the impact and understanding of evaluation systems (Close & Amrein-Beardsley, 2018; Geiger et al., 2020; Krimbill et al., 2019; Paufler,

2018; Reid, 2018), the use of teacher evaluation systems is highly ingrained in American educational policy and does not seem to be fading any time soon.

Design of Teacher Evaluation Systems. Four features generally help shape the design of an evaluation system: choice in performance measures, the way in which performance measures are placed on a scale, the system used to assign weights to performance measures, and thresholds for performance ratings (Steinberg & Kraft, 2017). Given that each state and district has their own needs and preferences when it comes to designing evaluations, it is important that designers and pertinent stakeholders begin by taking the time to analyze the role and responsibilities of their teachers so as to identify measures of performance to be included in the evaluation tool (Li et al., 2011); these measures may include (but are not limited to) such aspects as in-class observations of a teacher's instruction, student feedback on teacher effectiveness, or a teacher's demonstrated connection to measurable student growth. For example, the designers of the Texas Teacher Evaluation and Support System (T-TESS) identified the establishment and attainment of professional development goals, in-class observation, and analysis of student growth as crucial for inclusion in the state's preferred evaluation tool (TEA, 2015). When it comes to selecting specific performance measures, the designer must make such choices as which frameworks and protocols to use during observations, or how to convert responses on student surveys to criterion-based measures; in Texas, evaluators are given a specific rubric to complete during observations to ensure all evaluators follow similar protocols (Steinberg & Kraft, 2017).

A second feature shaping the design of a teacher evaluation system is the way in which performance measures are placed on a common scale. In determining an appropriate scale for performance measures, designers must consider such factors as the point range that should be

utilized in the scale and how points from different measures (such as observations and student surveys) should be placed on the scale (Steinberg & Kraft, 2017). In some cases, the scale range used in performance measures may stray away from the usage of obvious numerical points, instead relying on a scale made of different category descriptors. In Texas, the scale used in teacher observations involves five performance ratings: Improvement Needed, Developing, Proficient, Accomplished, and Distinguished. These five levels are used to assess teachers across sixteen different dimensions, such as "Classroom Culture" and "Content Knowledge and Expertise." Teachers must meet all the criteria listed under a given level in order to be ranked at that point on the scale; in other words, a teacher who clearly demonstrates all the criteria but one under the highest "Distinguished" category for a given dimension would be rated as "Accomplished" (the next highest scale value) in that dimension due to failure to meet all specified criteria for the topmost scale value (TEA, 2022e). While each level of the T-TESS observation scale is technically given a number (with one being assigned to the lowest level, "Improvement Needed" and five assigned to the highest level, "Distinguished"), these numbers are only used to electronically factor one's average score across all the dimensions, but these numbers do not appear on the T-TESS rubric and are never used in conversation by evaluators or educators.

The third feature that must be considered in designing an evaluation system is the use of performance weight measures; this aspect calls for designers to determine the way in which each aspect of a teacher's summative assessment should be weighed relative to one another (Steinberg & Kraft, 2017). T-TESS places the greatest emphasis on the observation cycle, with measures related to professional development comprising only a tenth of one's summative performance assessment score (TEA, 2015). The final feature to be considered, performance rating thresholds,

refers to the designer's responsibility in determining where to place the various level demarcations determining teacher ratings. For example, in New York City, earning 74% or more of available points earns a teacher a "proficient" rating, while teachers in Philadelphia only need to earn 50% of available points to receive this same summative evaluation rating (Steinberg & Kraft, 2017). Alternatively, T-TESS continues to use the five aforementioned levels to assign thresholds without publicizing details pertaining to the mathematical calculations underlying the ranking of teachers; that is to say, a teacher that averages "Needs Improvement" across the three performance measures used in evaluation (after the various measures have been weighed appropriately) stays in this bottom threshold. Of course, it is important to keep in mind that the design of evaluation systems varies from state to state. Though a lack of transparency often prevents teachers and external stakeholders from understanding exactly how teacher performance evaluations are designed (Close & Amrein-Beardsely, 2018), many published evaluative systems can still be broken down and understood in terms of the four factors emphasized by Steinberg and Kraft (2017).

Teacher Evaluation Practices in Texas. Like many other states, Texas has adopted its own unique evaluation system to aid in the process of teacher development and professional goal-setting. The Texas Teacher Evaluation and Support System (T-TESS) was developed by a committee of principals, teachers, and various representatives from educator and higher education associations. Facilitated by the Texas Comprehensive Center and the Texas Education Agency (TEA), this group of education-minded professionals updated teaching standards and built a rubric that could be tied to the updated standards during the 2013-2014 schoolyear (TEA, 2015). Very shortly after its development, T-TESS was piloted during the 2014-2015 schoolyear and began to be required of all Texas schools in the 2016-2017 schoolyear (Paufler et al., 2020;

TEA, 2022c; TEA, 2022d). This system consists of three components, broken down into weighted sections: (1) establishment of a goal-setting and professional development plan, which amounts to 10% of one's evaluation score; (2) the observation cycle, which counts for 70%; and (3) analysis of student growth, which comprises the remaining 20% of a teacher's evaluation (Krimbill et al., 2019; TEA, 2015). The observation and goal-setting components are frequently thought of together as the rubric portion of the evaluation, causing many to consider the evaluation to be 80% rubric and 20% student growth (TEA, 2015). During evaluations, evaluators focus on a teacher's ability to thoroughly address Texas state standards of learning (known as Texas Essential Knowledge and Skills, or TEKS), which are considered the pillars of for each subject within the state's required curriculum (TEA, 2022a). Consideration for and the addressing of standards must be evident not only in one's teaching (as assessed during observation), but also when demonstrating quantifiable measures of student growth and in establishing one's own professional development goals (TEA, 2015).

Consideration of the Tool's Development from an ID Perspective. It should be noted that, based on available information, formal instructional designers were not included in the design and development of T-TESS. Still, it is important to consider how an instructional designer might view the design of the evaluation tool such as T-TESS. Although instructional designers often have free reign in selecting or creating their own ID model to follow, designers typically utilize their knowledge of accepted design pathways to their advantage throughout the design process. For instance, a designer might have chosen to utilize Allen's Successive Approximation Model (SAM) in the design of the evaluation tool; though there are numerous ID models that might prove appropriate in designing an evaluation tool, this model will be referenced throughout this section because SAM highlights many of the iterative steps used in

other ID models. Following the general steps of SAM, the designer would divide their design process into three easily recognizable phases: preparation, iterative design, and iterative development. Knowledge of these overarching stages would allow the designer to understand precisely where they were in the design process, which would help inform one's planning and timeline for implementation. Continuing with the example of SAM, a designer would have gathered background information (including client needs and previous efforts to improve teacher evaluations, including fine arts evaluations) during the preparation stage. The designer would have then progressed on to the iterative design phase, which would involve creating a prototype of the evaluation tool to; this prototype would be assessed by stakeholders, with opportunities for feedback to be given and adjustments to be made before the finalizing the product. From here, during the iterative development stage, the evaluation tool would be continuously implemented, evaluated, and further developed until a "gold" (or finalized) version of the evaluation tool was agreed upon by all parties (Allen Interactions, 2022).

While the creators of T-TESS may have moved through the preparation and iterative design stages of SAM, there is no evidence that the tool's design was carried out with the same eye to detail that an instructional designer would have brought to the process. By analyzing perceptions on the effectiveness and impact of T-TESS and by compiling suggestions for the improved long-term use of the tool in evaluating fine arts educators, this study aimed shed light on potential areas of growth concerning the design of future iterations and implementation of the tool.

T-TESS Rubric. The T-TESS rubric is divided into four domains: (1) Planning, (2) Instruction, (3) Learning Environment, and (4) Professional Practices and Responsibilities. Each of these domains is further broken into dimensions or subsections (TEA, 2022b); for example,

Domain 1 (Planning) includes such components as Standards and Alignment (Dimension 1.1), Data and Assessment (Dimension 1.2), Knowledge of Students (Dimension 1.3), and Activities (Dimension 1.4). Each dimension contains a one sentence summary, along with the state standards (TEKS) that relate to that dimension. The one-sentence summary for Dimension 1.1 states, "The teacher designs clean, well-organized, sequential lessons that reflect best practice, align with standards and are appropriate for diverse learners" (TEA, 2022e). A rubric for each dimension lists descriptors of observable teacher behavior, divided into the following performance ratings, from lowest- to highest-achieving: Improvement Needed, Developing, Proficient, Accomplished, Distinguished (TEA, 2022b; TEA, 2022e). The T-TESS rubric is meant to support the professional development of teachers by facilitating self-reflection to aid in the planning and delivery of instruction, allowing for the ability to track progress towards professional goals, and providing clear designations of which teaching practices fall within each performance rating. When appraising teachers, evaluators are instructed to begin by rating teachers as Proficient, and then use observable evidence to justify upward or downward movement to another performance rating. While the Proficient performance rating falls in the middle of the rubric and a middle rating may not seem desirable to some teachers, TEA stressed that a Proficient teacher is a strong teacher, and that the upper performance ratings should only be assigned to teachers of the highest caliber to promote the use of the rubric as a growth tool (TEA, 2022b).

T-TESS Goal-Setting and Establishment of a Professional Development Plan. T-TESS utilizes a goal-setting and professional development plan to emphasize the importance of long-term growth, which should be viewed over the course of a teacher's entire career. TEA recommended the goal-setting and professional development plan be formulated at the end of a

teacher's first year of teaching and at the end of every subsequent school year so that the goal for the upcoming year may be directly related to the teacher's current standing. The appropriateness of professional goals should be reassessed at the start of each school year to best accommodate any shifts in teaching assignment, student population, or teacher aspirations. Throughout the year, the teacher and their evaluator should meet to discuss progress towards the goal and completion of professional development; at the end of each year, the teacher and evaluator should assess attainment of goals and set new goals for the following year. TEA offers districts a goal setting template and training modules, but the actual structure of the goal setting template is ultimately determined by each school district using T-TESS (TEA, 2022b).

T-TESS Observation Cycle. The observation cycle consists of a pre-conference, observation, and post-conference (TEA, 2015). The pre-conference provides an opportunity for the teacher and evaluator to discuss the teacher's upcoming formal observation, which often takes place on an agreed upon date, allowing the teacher to know when to expect their evaluator in class. Typically, the teacher will describe the lesson the evaluator should expect to see, highlighting such aspects as classroom routines and particulars, unique student needs, challenges faced in the classroom, pacing, differentiation, and student goals. The pre-conference is intended both to allow the teacher to prepare for their upcoming observation by talking through key details in advance and to allow the evaluator to ask necessary questions to better inform their upcoming evaluation. Though the pre-conference is not a required element of T-TESS, TEA recommended including this aspect in the observation cycle (TEA, 2022b).

Once the pre-conference is completed, the evaluator will complete a formal observation in which the evaluator documents evidence and determines ratings for the first three domains of the T-TESS rubric. Though T-TESS only requires that one formal observation be completed per

school year and allows districts to set their own stipulations regarding number and length of observations, it is recommended that multiple informal walkthroughs (and, when appropriate, multiple formal observations) be conducted to allow the evaluator to fully understand of the teacher's capabilities (TEA, 2022b)

An in-person post-conference takes place after shortly after the completion of a formal observation. During this time, evaluators use questioning and other coaching tactics (as established by each district) to discuss the observation, highlighting teacher strengths and areas for future growth. The evaluator also uses this time to discuss the teacher's performance ratings. All ratings should be supported by direct evidence pulled from the formal observation, providing a sense of transparency as to how the evaluator arrived at the assigned performance ratings. TEA suggested that a skillful evaluator should be able to direct the post-observation conference in such a way as to allow the teacher to self-identify areas of refinement, thus making the explanation of performance ratings a non-intimidating process (TEA, 2022b).

Benefits of Teacher Evaluations. Teacher evaluations generally aim to serve two functions: accountability and performance improvement. Evaluations can be used for summative purposes, allowing school leaders to demonstrate accountability through decisions made surrounding teachers' future employment and salary; teacher evaluations can also be formative in nature, leading to identification of areas of improvement for educators (Stronge & Tucker, 2023). When used for summative purposes, teacher evaluations can also provide stakeholders beyond the school building (such as state officials and parents) with a means to ensure teachers are being held to high standards and are performing to expectations (Jewell, 2017). In the formative sense, evaluations are typically intended to promote professionalism and continued growth among teachers, while simultaneously facilitating the increased learning and wellbeing of students

(Fresko & Levy-Feldman, 2023). Indeed, efforts like RTTT were implemented with the thought that teachers were the single most important school-based factor in predicting student success, highlighting the need to ensure the continued professional growth of teachers (Gitomer & Marshall, 2023). Teacher evaluations may also be leveraged to allow educators to feel a greater sense of self-efficacy; as Garet et al. (2020) indicated, schools in which principals encouraged teachers to play a more active role in their own evaluations were shown to have more success with the overall implementation of an evaluation process.

Issues with Teacher Evaluations. Of course, teacher evaluations are not without their issues. The use of VAMs in teacher evaluations has been highly disputed due to the plethora of factors beyond the teacher's capabilities that could influence a student's growth on such objective measures as standardized tests (Close & Amrein-Beardsley, 2018). Though some argue VAMs can be accurate in predicting student success (Tutt, 2018), numerous studies have illustrated that VAMs are often biased towards classrooms with high levels of achievement in prior classes and that VAMs in one content area can easily influence the teaching quality in another content area (Gitomer & Marshall, 2023). To add to the problematic nature of VAMs, these measures have been shown to be incorrect approximately 25% of the time, meaning that a quarter of teachers evaluated using VAMs would be scored much higher or lower than they might be otherwise. Moreover, VAMs can only be used to estimate the effectiveness of 30-40% of all teachers (primarily those teaching core content areas with standardized tests (Geiger et al., 2020); it is not uncommon for the other 60-70% of teachers to be assigned the scores of another teacher in a tested subject (Gates et al., 2015; Potter, 2021b). Despite these issues, VAMs continue to be utilized to vary degrees by many states (Gitonmer & Marshall, 2023; Paufler,

2018). In fact, as of 2020, 12 states continued to encourage public schools to make firing decisions based primarily on VAM data (Geiger et al., 2020).

Of course, identifying the pitfalls of the continued usage of VAMs is not enough; there is a need to consider alternative concepts that could be used in the design of an improved performance evaluation. This is where Eisner's concepts of connoisseurship and criticism come in to play (Eisner, 2003). Looking at an educator's contributions and capabilities from the perspective of one who has acquired an understanding and appreciation of a given content area (in other words, from the perspective of a connoisseur), rather than assigning scores based on VAMs that may not effectively reflect a teacher's capabilities, should allow fine arts educator performance assessments to be executed in a manner that allows transparent and intentional appraisals of fine arts educators' capabilities to come to light. It therefore follows that T-TESS, an evaluation tool that does not require the evaluator to be a connoisseur of the fine arts, could be improved to provide a more reliable and valid take on a fine arts teacher's performance. This study aimed to gather data regarding educator and evaluator perspectives on T-TESS as it currently is utilized. In this way, this study aimed to address the literature gap surrounding how to make fine arts educator performance assessments more meaningful; it was anticipated this study would hold promising implications for the design of future fine arts educator performance evaluation tools.

Teacher evaluations also suffer from inconsistencies in scoring as well as biases among evaluators and gamification of the evaluation system (Close & Amrein-Beardsley, 2018). Gitomer and Marshall (2023) found Black teachers are more likely to receive lower evaluation scores than their White colleagues, and lower observation scores are typically given to students who struggle more academically. Likewise, evaluators are known to inflate the evaluation ratings

of teachers they get along with or want to protect, particularly in an effort to counterbalance low test scores, boost teacher morale or avoid having difficult conversations with under-performing teachers (Close & Amrein-Beardsley, 2018; Jones et al., 2022). In some cases, these pitfalls may be attributed to the design of an evaluation tool. For example, designing an evaluation tool to include specific language and concrete examples of what to look for during the observation process has been shown to reduce bias and increase evaluator confidence in conducting teacher performance evaluations (Jones et al., 2022). However, it stands to reason that if the specific examples within the evaluation tool do not directly relate to the content area being observed—as is the case with T-TESS, which lacks direct references to the fine arts—evaluators may not experience the same increased levels of confidence when assessing those subjects.

Lack of transparency plagues teacher evaluations. Considering that much of the evaluation process is left up to each state (Paufler, 2018), and considering that cities like Houston pay $680,000 per year for their measures of student growth (Close & Amrein-Beardsley, 2018), it is no surprise that states are reluctant to publicize details surrounding evaluation processes. However, this tendency towards secrecy negatively impacts teachers. Teachers in Houston, North Carolina, and Tennessee have all asserted that there is very little (if any) support provided in helping teachers understand their scoring or what may be done to improve instructional practices in the future (Geiger et al., 2020). From a design perspective, this lack of transparency can be reduced by designing performance evaluation tools that include content-specific examples to guide evaluators throughout the observation cycle, and by including explicit requirements that the educator and evaluator participate in a collaborative and detailed review of the teacher's performance (Jones et al., 2022, TEA, 2022d). The third research question of this study aimed to examine the needs identified by evaluators and fine arts educators

for improving the long-term use and effectiveness of T-TESS—needs which may be addressed through such design features, among others.

Finally, teacher evaluations often suffer due to aspects related to evaluators. Principals tend to be split as to why they evaluate teachers. Some feel they are merely completing a job requirement, while others attempt to use evaluation as a pedagogical tool to aid their staff; additionally, due to such factors as limited training on evaluation procedures, evaluators may struggle with giving meaningful feedback (Fresko & Levy-Feldman, 2023). Despite training on evaluation systems, Reid (2018) reported that training only minimally influences how evaluators think about and evaluate teachers; many evaluators rely on their own previous experiences to determine what quality teaching looks like to inform their evaluations. One possible outcome of this study was the identification of evidence suggesting that elevating evaluators' connoisseurship and ability to provide meaningful, content-specific criticism may enhance the overall effectiveness of the evaluation process. Such improvements will undoubtedly be critical if performance evaluations (particularly in the fine arts) are to ever earn a less contentious reputation.

Effects of the Evaluation Process. Given that the evaluation process is utilized to determine a teacher's effectiveness and identify areas for growth, it is important to consider the unintentional consequences teachers may face as a result of implications associated with the evaluation process. As mentioned above, in some states, evaluations are used to dictate or inform contractual decisions; for some teachers, a poor evaluation can result in termination (Geiger et al., 2020). Due to their history of being used to help justify terminations, evaluations are still seen as something to be feared or disliked by many teachers (Willey, 2019). Though likely unintentional, teacher evaluations are frequently associated with decreased teacher and principal

morale, as well as with diminished sense of professional value (Paufler, 2018). Additional unintended consequences associated with evaluations include teachers being more likely to shy away from collaboration and maintain privacy in their work, teachers avoiding scenarios that require them to work with high-needs students who may hinder their evaluation results, and teachers leaving schools that tend to consistently produce lower evaluation scores. In some cases, this fear of the repercussions of poor evaluation results may cause teachers to leave the profession altogether (Geiger et al., 2020).

While the aforementioned effects can often be attributed to the way in which current teacher performance assessments are implemented, it is important to remember that implementation is a key part of the ID process, especially given that most ID models follow the general sequential and iterative process suggested by the ADDIE framework: Analysis, Design, Develop, Implement, Evaluate (Cram & Marabelli, 2018; Tessmer & Wedman, 1990). Therefore, the implementation of teacher performance assessments is related to and affected by the design of these tools. In other words, issues with implementation (such as negative perceptions surrounding evaluations due to the way in which teacher assessments are executed) may be tied back to the design of the tool. Designing future evaluation tools with an eye towards ensuring that meaningful criticism results from the tool's implementation may help allay these more negative implications of the evaluation process. As such, this study's investigation of fine arts educator and evaluator perceptions highlighted the critical role that thoughtful design plays in shaping not only the implementation and effectiveness of teacher evaluations, but also in cultivating evaluators' sense of connoisseurship and ability to provide meaningful criticism, in line with Eisner's theories.

Exacerbated Issues with Fine Arts Teacher Evaluations

Fine arts teachers often face additional issues with the evaluation process beyond those faced by teachers of more traditional content areas. These issues include concerns with evaluator preparedness, the impact that increased emphasis on standardized testing and core subjects has on fine arts teaching and evaluation practices, and the lack of input fine arts educators have traditionally had in the evaluation process.

Issues with Evaluator Preparedness and Job Scope. One of the primary issues plaguing the field of fine arts evaluations is that evaluators lacking backgrounds in the fine arts are typically trusted to measure the effectiveness of fine arts teachers, even though these evaluators are often unaware of what constitutes best practices in a given fine arts classroom (Tutt, 2018). Many fine arts content areas are already subjective in nature or are characterized by their temporal nature. Due to the inability of many fine arts subjects to be assessed in terms of concrete numbers, it is understandably difficult for those lacking a detailed knowledge of the arts to evaluate teacher effectiveness with any sense of reliability (Lin, 2013; Tutt, 2018; Yang, 2023). Though instructional designers are not necessarily subject matter experts themselves on every project on which they work, it is important for instructional designers to consult with subject matter experts (such as seasoned fine arts educators) when designing evaluation tools to ensure that expected performance outcomes align with the design criteria of the tool.

Temporal constraints also impact the ability of evaluators to conduct quality fine arts teacher evaluations. Evaluators admit there is often not enough time to conduct teacher evaluations as carefully as they should, leading evaluators to deem the evaluation process unhelpful (Jones et al., 2022). This imbalance between time and evaluative duties may also help explain why evaluators remain uninformed about fine arts best practices (Lin, 2013).

Furthermore, evaluators have been shown to struggle when it comes to consistently assessing and supporting their ratings of teacher effectiveness; one study showed principals initially judged 19% of teachers as being below proficient, but quickly dropped this statistic to a mere 6% when asked to provide more detailed ratings of personnel (Jones et al., 2022). In short, a combination of time limitations and lack of appropriate content knowledge puts evaluators at a disadvantage when it comes to conducting quality fine arts teacher performance evaluations. This study not only focused on understanding if participants perceived these issues as impacting the effectiveness of T-TESS, but also explored suggestions for addressing these challenges to promote meaningful evaluations of fine arts teachers.

Influence of the Emphasis on Core Subjects. Given that American school districts are frequently pressured to improve standardized test scores, it is no wonder that many schools have placed increasing emphasis on instruction related to those subjects that tend to be featured on standardized tests, such as mathematics and English (Gates et al., 2015; Kimelberg et al., 2019). High-stakes evaluations and standardized testing (in both the fine arts and other content areas) frequently force teachers to narrow the scope of their curriculum in an effort to counterbalance decreased teaching time and allow more time for test preparation; this results in decreased feelings of autonomy and an inability to demonstrate one's full teaching capabilities during observation (Kimelberg et al., 2019; Potter, 2021b). In some cases, school districts and evaluators may look for examples of writing and mathematics in all classes, including fine arts classes. This may lead fine arts teachers to be forced to devote classroom time to writing or math assignments that break the flow of instruction in order to adhere to evaluator expectations, even when there is inadequate space for students to write (such as using the floor or a music stand instead of a desk (Gates et al., 2015)). Even when not directly instructed to do so by their

evaluators, some arts teachers report feeling the need to resort to “fake teaching” in order to present evaluators with what they expect to see, rather than an authentic fine arts education (Potter, 2021b). Similarly, knowing the precarious position that the fine arts hold in many districts and understanding the emphasis placed on “core” content areas, fine arts teachers often feel the need to justify art as a way to help teach other course content, which in turn encourages evaluators and other school officials to take the fine arts less seriously (Kimelberg et al., 2019).

As mentioned previously, many schools incorporate students’ standardized test scores in subjects such as mathematics and English into the evaluations of all teachers within a school, including fine arts teachers (Gates et al., 2015; Potter, 2021b). Although the arts were identified as a core subject under the 2001 No Child Left Behind Act, legislation only necessitated testing for math and reading, a requirement that still influences testing decisions today (Shaw, 2016). It is not uncommon for this practice to diminish the individual performance evaluation ratings for fine arts teachers, many of whom are rated highly when it comes to their individual classrooms and teaching capabilities, but whom see marked declines in their overall evaluative rankings when school scores pertaining to other subject areas are factored into the teachers’ rankings (Kimelberg et al., 2019). In this way, the intense focus on achievement in the core subjects negatively affects the performance evaluation outcomes for fine arts teachers across the nation, making it difficult for fine arts teachers to receive an accurate and meaningful appraisal of their teaching abilities.

In order to address this issue, an instructional designer working on a new evaluation tool geared towards fine arts teachers must consider how to best highlight the fine arts as their own content area, worthy of evaluation separate from ties to such core subjects as mathematics and English. While efforts – described below—have been made to stress the importance of

evaluating the fine arts differently than other content areas, these efforts have not been without their pitfalls (Gates et al, 2015; Potter, 2021b; Yang, 2023). One anticipated outcome of this study was to contribute to these efforts by identifying participant needs for the improving the long-term use and effectiveness of T-TESS in assessing fine arts educators.

Lack of Input. Yet another issue with fine arts performance evaluations stems from the limited role fine arts educators are expected to play in designing the evaluation process. This lack of input in teacher assessment practices is exemplified by such measures as Pearson's Teacher Performance Assessment (edTPA), which is used in over forty states as a means of high-stakes evaluation by teacher preparatory programs, with several states requiring prospective teachers to pass edTPA in order to become certified. The standards for edTPA were set in 2013 by a group of twenty-six policymakers and twenty-one practitioners; none of the policymakers were fine arts educators, and only one of the practitioners had experience in the fine arts. Not only were fine arts educators largely left out of the discussion surrounding teacher evaluations, not a single arts-based organization was included among the plethora of subject matter organizations that helped inform the edTPA performing arts handbook. The fact that edTPA evaluation standards for arts teachers were developed by non-arts teachers is undoubtedly cause for concern and logically explains why so many fine arts teachers report feeling their teaching abilities are hampered by edTPA standards (Potter, 2021a).

Similarly, the hyperfocus on standardized testing mentioned above leaves little room for fine arts educators to voice their concerns regarding the evaluation process. Insistence on using formalized testing procedures to measure student growth often prevents fine arts teachers from utilizing meaningful projects that would better demonstrate both student learning outcomes and the effectiveness of the teacher (Kimelberg et al., 2019). As Shaw (2016) pointed out, teacher

evaluations typically rely on strict categories that assign fine arts teachers a single effectiveness level (such as ineffective, effective, or highly effective) without providing adequate space for the fine arts educator to suggest other means of evaluation that may more accurately demonstrate their aptitude. Number-based effectiveness categories and VAMs were created with a small group of teachers in mind (those whose subjects can be easily tested and broken into numbers), and with no regard for or inclusion of fine arts educators in the creation of performance evaluation policies (Gates et al., 2015; Kimelberg et al., 2019). This study provided support for the notion that using an evaluation tool devised in line with the input of fine arts educators might allow for more accurate performance evaluations, thus inspiring meaningful and lasting changes in the field of fine arts teacher performance evaluations.

Efforts to Improve Fine Arts Teacher Evaluations. In recent years, numerous efforts have been made (with varying levels of success) to try to improve the way in which fine arts teachers are evaluated. For example, the Colorado Department of Education includes a statement in their *Practical Ideas for Evaluating Teachers of the Arts* instructing fine arts educators to only incorporate core content areas into their classes in authentic ways, such as writing a response to a performance or naturally incorporating math into elements of a visual arts class. However, this intent for the fine arts to be taught in their most genuine manners has been somewhat undermined by the state's simultaneous focus on the development of additional evaluation and assessment resources (Gates et al., 2015).

The Tennessee Fine Arts Portfolio Model (TNFAPM) provides another example of an effort to improve the evaluation process for fine arts teachers. The TNFAPM, designed by music teachers, provided an alternative to typical music teacher evaluation practices, which historically incorporated students' reading and math scores. Though the portfolio represented a positive step

forward by providing music teachers with an alternative to test scores in their evaluations, non-music educators have asserted their influence on the nature of the TNFAPM over time, leading to increases in automation—thus undermining the efforts that originally led to the creation of this portfolio. Though the inception of the TNFAPM represented a massive step forward in giving fine arts educators a voice in the evaluation process, fine arts educators find themselves continuing to fight for their voices to be heard in the evaluation process (Potter, 2021b).

A third instance of efforts to influence the evaluation of fine arts teachers may be found in Yang (2023), who recently published a multi-dimensional evaluation index system (EIS) intended to better assess teachers' role in visual arts courses. While the validity of the model was verified and the EIS showed promise, Yang's (2023) study represents one instance of implementation, and such a model has yet to be adapted for wide-spread use. In summary, though efforts have been made to improve evaluation practices for fine arts teachers, very little lasting progress has been made in this field.

Analyzing these past efforts to improve fine arts educator performance evaluations points to a commonality that may have contributed to the pitfalls identified in each example. Each of the three efforts relied on the input of a single group of subject matter experts (fine arts educators) to inform the design and implementation of new evaluation measures. No known research has, as of yet, or explored perspectives from multiple stakeholder groups, as is the case with this study. An anticipated outcome of this study was a demonstrated need for instructional designers to consult multiple types of stakeholders when designing an evaluation tool like T-TESS, especially when the tool caters to a niche audience, such as those involved in fine arts educator performance evaluations. By compiling the perspectives of multiple stakeholder groups,

this study also explored suggestions for replicable ways to improve the accuracy of quality in fine arts teacher evaluations.

Calls for Additional Research. While previous efforts to improve the evaluation process for fine arts teachers do not go unrecognized, the need to continue these efforts and make lasting changes to the face of fine arts teacher evaluations has been voiced by a number of researchers. For example, Willey (2019) and Garet et al. (2020) insisted future educational reforms and evaluation tools must consider the emotions of educators, and treat educators as masters of their content areas, if they are to be successful. Researchers further asserted that fine arts educators continue to collaborate with education leaders and make their voices and relevance known to those at the state and national education levels; this would help shed light on the need to reassess the way in which fine arts educator evaluations are designed (Potter, 2021b). Tutt (2018) specifically called for fine arts educators to advocate for ways in which success may be better measured in the fine arts so as to lead to more meaningful teacher evaluations. Though there is still great room for growth when it comes to improving the way in which fine arts teachers are evaluated, there is a clear desire to advance practices in this field, indicating the necessity of studies such as this one, which focused on understanding the perceptions of fine arts educators and evaluators regarding a specific evaluation tool and ways in which this tool could be improved to ensure optimal effectiveness for fine arts educators.

Ethical Assurances

This study received approval from National University's Institutional Review Board (IRB) prior to data collection. During data collection and throughout the completion of the study, intentional efforts were taken to address such ethical considerations as informed consent, anonymity, and confidentiality, as outlined in the Belmont Report (National Commission, 1978).

Participation in this study was voluntary, and participants maintained the right to withdraw from the study at any time. All participants were provided a clear explanation of the research procedure, purpose of the study, and anticipated benefits and potential risks. Participants were also given the opportunity to ask questions (via email or in person) before committing to this study; an informed consent was also provided prior to data collection. All participant experiences were regarded equally, and participants were debriefed at the conclusion of each interview. Participant anonymity was ensured by assigning each participant a number and letter identifier (E1, E2, and so forth for evaluators, and T1, T2, and so forth for teachers), and these identifiers were used to refer to participants in all notes and documentation related to the study. All documents containing confidential or identifying information have been stored in a secured location: electronic documents were stored on my password-protected computer, while paper documents will be kept in a locked cabinet for three years in accordance with IRB requirements.

As a researcher, I aimed to conduct this study in an unbiased manner, though I would be remiss if I did not acknowledge my background and personal interest in this study. During my time a public-school dance teacher, I have had many moments where I felt my evaluators did not fully understand my position or the methodology and rationale behind my teaching decisions. In some cases, my evaluators would admit that my content area fell outside their scope of expertise and that they were not sure if their feedback was going to be beneficial in aiding my professional development; in other cases, evaluators would simply take off points during my evaluations for items that were not applicable to a fine arts classroom. Many of my colleagues have shared similar anecdotes. This led to my interest in better understanding the perceptions of both fine arts educators and their evaluators, as well as in exploring ideas for efforts to change the face of fine arts teacher evaluations.

Though my experience as a fine arts educator and current employment as a dance teacher with a Texas school district afforded me an insider perspective that allowed me to more easily access the phenomenon being studied, ask more detailed questions, and develop a stronger sense of trust with participants (Holmes & Gary, 2020), I was aware that my positionality on this topic could be viewed as a limitation to this study. Positionality refers to the way in which social position and power may shape one's identity and access, as well as one's world view, which may inform their position in research (Bloomberg, 2023; Holmes & Gary, 2020). I recognized the importance of practicing reflexivity, or the continual examination of my positionality as a researcher and the acknowledgment of the intentional and unintentional effects I may have had on the research process (Berger, 2015; Bloomberg, 2023). To this end, I engaged in journaling throughout the research process and maintained a "paper trail" of field notes (Nowell et al., 2017) to reflect on my personal beliefs and analyze their potential impact on this study. I actively acknowledged differences among participants and made every effort to ensure personal perspectives did not influence research outcomes. I also avoided unethical research practices, maintained transparency throughout my work, and verified my understanding of all participant perspectives through peer debriefing (Bloomberg, 2022) in order to present my findings in the most accurate manner possible. Still, my personal opinions and positionality have been listed as a potential limitation of this study.

Summary

While concerns about the teacher evaluation process in America plague all content areas (Close & Amrein-Beardsley, 2018; Gitomer & Marshall, 2023; Paufler, 2018), fine arts educators have long faced additional disadvantages when it comes to navigating the evaluation process (Gates et al., 2015; Tutt, 2018). Unlike other subjects, demonstrating clear and objective

student growth proves challenging for fine arts educators due to the inherently subjective nature of their content areas. (Lin, 2013; Yang, 2023). Additionally, these educators are often observed by evaluators who lack the appropriate content knowledge to provide quality evaluation feedback (Tutt, 2018). Many fine arts educators are pressured to inauthentically incorporate such "core" subjects as mathematics and English into their lessons in order to meet the expectations of evaluators who expect to see these subjects explicitly taught in all classes (Gates et al., 2015). These challenges can cause fine arts teachers to fear the evaluation process, especially considering that in some cases, a poor evaluation may result in pay cuts or termination (Geiger et al., 2020; Willey, 2019). Stressors surrounding evaluations also have been shown to lead to declines in morale, "fake" teaching, and a diminished sense of professional value (Paufler, 2018; Potter, 2021b). Needless to say, the current landscape surrounding fine arts teacher evaluations profoundly impacts the personal and professional well-being of fine arts educators, highlighting the need for a more effective and supportive approach to evaluation within the fine arts.

The purpose of this action research was to explore perceptions held by evaluators and fine arts teachers regarding the effectiveness of T-TESS in terms of how well the tool achieves validity and alignment with expected performance goals in the evaluation of fine arts educators. Validity was operationally defined as the extent to which performance evaluation rankings were perceived as fair and were supported by underlying reasoning. Throughout the study, focus was placed on ascertaining participant perceptions regarding how effectively the aforementioned tool gauges the quality of fine arts teachers, as well as on how the design of the tool aligned with expectations regarding the way in which teacher quality is assessed and defined. The sample for this study involved ten fine arts educators and three evaluators with at least one year of prior experience in their current role with their school district. Data was

collected through questionnaires (administered through Qualtrics) and one-on-one interviews (held and recorded through Zoom). Participants were asked to share their perspectives on the usage and impact of T-TESS, with questions focusing on such topics as accuracy within evaluations, influence of the tool on educator morale, the ability of the evaluation tool to facilitate fine arts educator's reflection on their teaching capabilities, and influence of the tool on educators' daily teaching practices. Data was thematically analyzed with the aid of NVivo. The following section expands on the established challenges and research gap surrounding the evaluation of fine arts educators by outlining the methodological approach and design employed in this study to explore fine arts educator and evaluator perceptions surrounding T-TESS.

Section 2: Methodology and Design

The face of teacher evaluations in the United States is contentious (Close & Amrein-Beardsley, 2018; Gates et al., 2015), particularly when it comes to evaluating fine arts teachers (Gates et al., 2015; Lin, 2013; Potter, 2021b; Tutt, 2018). It is common for fine arts educators to receive observation scores that fail to take into account the unique nature of their content areas; this is often due to a combination of evaluators' lack of fine arts content knowledge and temporal constraints affecting consistency in performance evaluations (Jones et al., 2022; Lin, 2013; Tutt, 2018). Likewise, research indicates that standardized testing and the tendency of many school districts to focus on the core content areas (such as mathematics and English) frequently force fine arts teachers to narrow the scope of their curriculum to allow more time for test preparation in core subjects, which results in an inability for fine arts teachers to demonstrate their full teaching capabilities during evaluation-mandated observations (Kimelberg et al., 2019). This is an important issue in states like Texas, where observations count for 70% percent of one's evaluation score (TEA, 2015). The problem addressed by this study was the ineffectiveness of current fine arts teacher evaluation practices (Gates et al., 2015, Potter, 2021b, Tutt, 2018). The purpose of this action research was to explore perceptions held by evaluators and fine arts teachers regarding the effectiveness of T-TESS in terms of how well the tool achieves validity and alignment with expected performance goals in the evaluation of fine arts educators.

This section opens with a description of the methodology and design that best fits the research problem and purpose of this study, though the merit of alternative designs is also explored to better justify the researcher's choice in design. An explanation of the data analysis process and researcher's positionality is provided, followed by an outline of population and

sampling method. This section explains the instruments used throughout the data collection, along with a description of data collection and analysis. The section concludes with a discussion of assumptions, limitations, and delimitations associated with the study, culminating in a summary of the section.

Design and Method

This study embraced a qualitative methodology, which served the nature of the study and aligned with the study's problem, purpose, and research questions. To reiterate, the purpose of the study was to explore perceptions held by evaluators and fine arts teachers regarding the effectiveness of T-TESS in terms of how well the tool achieves validity and alignment with expected performance goals in the evaluation of fine arts educators. Focusing on perceptions strongly aligned with qualitative methodology, which emphasizes crafting a holistic picture of the phenomenon at hand and embracing multiple perspectives to accurately represent participants (Creswell & Poth, 2018). Similarly, the research questions for this study centered on the perceptions of fine arts educators and evaluators. A study that made sense of participant perceptions, without focusing on numerical data or proving trends between variables, aligned with qualitative methodology (Bloomberg, 2023).

This qualitative study used an action research design. Action research typically focuses on investigating, planning, and implementing new practices in order to support change to current practices within a field (Given, 2008). A key component of action research that drives forward this sense of change is the evaluation of results, which allows for revisions in one's future approach to a problem (Caro-Bruce, 2000). The study recorded participant perspectives regarding the effectiveness of an evaluation tool (in essence, providing an evaluation of the perceived effectiveness of the tool). An anticipated outcome of this study was that stakeholders

could familiarize themselves with perceptions regarding T-TESS' use in fine arts education, as well as suggestions for efforts to improve the fine arts educator evaluation process in Texas, which would in turn better inform future approaches to the evaluation cycle. The trajectory of this study also aligned with Mertler's (2019) sentiment that action research is an appropriate methodology for connecting participants and stakeholders while tracking efforts to improve aspects of the school community that affect these parties.

Action research tends to derive such features as its problem and research questions from the perceptions of practitioners within a specific setting; this focus on local context allows action research to further the development of an organization (Herr & Anderson, 2005). To this end, this study focused on perspectives surrounding an evaluation tool created and used specifically by school districts in Texas, and this study's findings directly benefit those school districts by informing the future development and implementation of T-TESS. Action research is typically considered appropriate when investigating fundamental changes to the way in which processes are executed in a systems-oriented setting, such as within an educational setting (Mertler, 2019). A school district's educator evaluation procedures are highly systematic, further accentuating the appropriateness of action research when it comes to exploring perceptions surrounding possible changes to traditionally-accepted evaluation systems.

The structure of this particular study aligned Greenwood and Levin's (2007) description of action research. As Greenwood and Levin (2007) indicated, action research must entail three elements: action (or a goal of promoting change), research (or generating knowledgeable findings), and participation (or the involvement of stakeholders who are directly involved in and affected by the phenomenon of interest). In line with the participation component, this study required direct engagement with stakeholders, who were asked to share their perceptions on the

effectiveness of T-TESS. Stakeholder participation allowed the researcher to analyze perspectives in order to present detailed, data-supported findings, in line with the research component of action research. The study's findings were anticipated to inspire meaningful and lasting improvement to T-TESS as it is used in the evaluation of fine arts teachers (in line with the action component of action research).

Another key tenet of action research is the improvement of practice that results from abandoning commonly accepted habits or customs in favor of informed practice (Kemmis et al., 2014; Whitehead, 2008). With this said, one of the objectives of this study was to facilitate movement towards an evaluation process and tool that are informed by the feedback of and have the support of various stakeholder groups (namely fine arts educators and their evaluators). Finally, action research was a suitable design for this topic, given the need for empirical evidence to support the effectiveness of T-TESS. This need for evidence aligns with Kurt Lewin's belief that action should not exist without research, nor research without action (Adelman, 1993). In other words, it is not enough to implement an evaluation tool and assume its continued effectiveness from year to year, or among all stakeholder groups. One must conduct research to support and further encourage data-driven change in the field of fine arts educator performance evaluations—change that may ultimately lead in a different direction than anticipated, based on the outcomes of one's research. Action research served as the most suitable methodology by which to carry out this sort of research.

Other designs, such as a single case study design and phenomenological design, were considered for this study. Given that a single case study focuses on understanding an individual case or phenomenon rather than coming to a discovery that applies to multiple cases (Stake, 1993), I initially felt a single case study might be appropriate for this study. Single case

studies allow for the study of real situations from the perspective of a single group, with emphasis placed on including the individual accounts and interpretations of participants in the resulting research report (Bloomberg, 2018). Yet, single case studies tend to focus primarily on observing and analyzing an ongoing phenomenon (Bloomberg, 2018); observation and analysis in a case study do not necessitate there is a problem to solve, as is typically the case with action research. For example, a single case study could have been formulated to simply observe how T-TESS was implemented over the course of several years. However, since this study focused on the need to solve a given problem faced by a specific group (the need to improve fine arts educator performance evaluation practices), utilization of a single case study would not have been ideal for this specific study.

A phenomenological design was also considered for this study. Giorgi indicated that a phenomenological study aims to describe a given phenomenon as accurately as possible (Groenewald, 2004), while researchers like Welman and Kruger (1999) stressed the need for a phenomenological design to center on the perspectives of those involved in the phenomenon. I was led to consider the utilization of a phenomenological design due to these overarching aspects of a phenomenology, which could also be applied to this study. However, a phenomenological research design aims to capture the essence of a particular phenomenon by articulating a common meaning experienced by several individuals in relation to the given experience (Creswell & Poth, 2018; Van Manen, 2018). While a study could potentially be formulated to search for a universal thread regarding perceptions towards T-TESS, this study focused instead on showcasing trends among participant perspectives. Rather than attempting to find the commonalities among multiple participant perspectives with the goal of arriving at a single common meaning, this study ensured differing perspectives were explained in detail.

Additionally, a phenomenological study does not place the emphasis on improving practice, as is characteristic of action research (Ungvarsky, 2020). Through the dissemination of perspectives surrounding the effectiveness of the evaluation tool, this study aimed to inspire change to practices in the field of fine arts educator performance evaluations, indicating the purpose of this study aligned with an action research design, rather than with a phenomenological design.

This study made use of questionnaires and semi-structured interviews. The use of the questionnaires allowed the researcher to identify common themes among responses that needed to be explored further due to participant confusion or other uncertainties; the questionnaires also heightened transparency, trustworthiness, and ongoing self-reflection (Keefer, 2009). Furthermore, the questionnaire allowed stakeholders to make meaningful adjustments to future practice by providing insight into participant struggles and perspectives regarding current practices (Jacobs, 2015; Samuel & Conceição, 2022). The choice to utilize questionnaires to collect data made sense when considering that this action research aimed to provide an initial set of perspectives that might be used to inform trends or necessary adjustments to future evaluation practices. The goal of this study was to better understand perceptions regarding the effectiveness of T-TESS in terms of how well the tool achieves validity and alignment with expected performance goals perceptions. Questionnaires provided a starting point to understand these perspectives, with individual interviews allowing for deeper exploration of participant sentiments. Additionally, each participant was interviewed in a one-on-one setting over Zoom within two weeks of their completion of the questionnaire. Semi-structured interviews allowed me to engage with participants in a conversational manner that was intended to encourage participants to share personal feelings and ideas without the sense of restraint that might be felt in more clinical interview settings (Elhami & Khoshnevisan, 2022). Holding interviews within a

two-week timeframe was intended to provide me adequate time to familiarize myself with each participant's responses to the questionnaire and allow me to determine any additional areas of focus that might need to be incorporated in the individual interviews. During data collection, focus was placed on understanding individual perceptions regarding the effectiveness of T-TESS, as well as on how the tool's design aligns with expectations regarding perceived fairness surrounding the way in which teacher quality is assessed and defined.

Other data collection methods were considered and ultimately deemed inappropriate for this study. For example, the use of observation, through which the researcher enters the natural setting of the study and takes field notes based on their personal interpretations (rather than relying on someone else's perspective (Bloomberg, 2023)), was briefly considered. Observations would have allowed participants to be monitored in real time as they interacted with the various aspects of T-TESS. However, since this study focused on understanding the perspectives of fine arts educators and their evaluators, using a method of data collection that would have pulled away from the notion of participant perspectives in favor of the researcher's firsthand interpretations would have be counterintuitive to the aims of this study.

The inclusion of focus groups in this study was also strongly considered. Focus groups allow a small gathering of participants to discuss and provide insight into a series of questions posed by a moderator (Dziak, 2022). While organizing a focus group of fine arts educator participants and a focus group of evaluator participants could have had its advantages, the number of participants (ten fine arts educators and three evaluators) would likely have limited the advantages associated with using focus groups, as it is recommended that focus groups contain between five and ten participants (Bloomberg, 2023), and the evaluator group would only consist of three participants. Additionally, Lazar et al. (2017) suggested it is not wise to

rely on a single focus group session, as a single group of fine arts educators or evaluators could provide inaccurate insight into a given phenomenon due to such factors as power dynamics or mixes of particular personalities within the given focus group. Therefore, the use of focus groups was ultimately considered to be inappropriate for this study.

This study employed thematic analysis to explore and make sense of participant perceptions. The term thematic analysis may be used to describe different techniques used to systematically identify, organize, and provide insight into themes within a set of data (Braun & Clark, 2006). After collecting data and using NViVo to transcribe all interviews, I repeatedly read over the data and took general notes to familiarize myself with the data set. From here, the data was be coded using NVivo. When engaging in thematic analysis, it is common practice to generate codes and consider themes while actively engaging in the coding process (Naeem et al., 2023). Therefore, I approached coding inductively, allowing the data to "speak" for itself, rather than coming up with a listing of themes that had to be found in the data (Demystifying Research Methodology, n.d.). From here, I discerned themes from the coded data. Throughout this process, I often returned to early steps of the coding process to ensure I was representing the data as accurately as possible. I engaged in what Naeem et al. (2023) referred to as conceptualization, refining and confirming themes, while ensuring all findings were accurate and could be logically derived from the data. I also utilized a visual map to help connect themes. Once I was confident I had derived the most accurate meaning possible from the dataset, I began to construct a narrative and reported my findings, all the while keeping in mind my role as the researcher in actively constructing themes (Braun & Clarke, 2006; Braun & Clark, 2022). The choice to use thematic analysis allowed me to fully understand overarching themes across the numerous interviews and questionnaires that comprised this action research study.

Positionality may be thought of as the researcher's role, identity, and social location in relation to the setting and framework of the research (Bloomberg, 2019). As Rowe (2014) indicated, positionality affects all stages of the research process, especially since a researcher's position may not be static throughout a given study (Berger, 2015). My positionality within this case study is that of a fine arts (Dance) educator employed by a Texas school district. This study was informed by my experience as a fine arts educator, so it is possible that my positionality influenced the type and depth of information participants chose to share with me, as well as my analysis of their perceptions. In an effort to practice reflexivity and address my positionality, I remained aware of my personal values and potential biases. I also included a log of my thoughts and reflections in the Appendix section of this paper to transparently illustrate my thoughts throughout the research process.

Population and Sample

While this study holds implications relevant to all fine arts educators and evaluators across the country, the target population for this study comprised all fine arts educators and evaluators working for Texas school districts utilizing T-TESS. There are 1,207 school districts in the state of Texas, containing 9,000 school campuses (TEA, 2024a). Though there is no publicized, comprehensive count concerning exactly how many districts use T-TESS, roughly 5% (or 60) of these Texas school districts piloted T-TESS during the 2014-2015 schoolyear, and 250 schools implemented a refined version of T-TESS in the 2015-2016 schoolyear; T-TESS remains the statewide recommended evaluation system at the time of this study (American Institutes for Research, 2016; TEA, 2024b). Members of the target population have worked for a Texas school district for upwards of one year; fine arts educators in the target population have taught and currently teach a variety of disciplines: band, choir, dance, orchestra and strings,

piano, theatre arts, and visual arts. This target population included all fine arts educators and evaluators who interact directly with, and thus, are impacted by, T-TESS throughout the teacher evaluation and observation process.

As of the time of writing, there was no comprehensive or publicly available count of all fine arts teachers employed in the state of Texas, nor a centralized record denoting how many of these teachers are evaluated through T-TESS. This absence of data posed a challenge in establishing an accurate estimate of the target population for this study. TEA publishes an annual teacher full-time equivalent count, which quantifies the instructional time allocated to each course. A manual aggregation of FTEs related to theatre-, art-, dance-, and music courses for the 2024-2025 schoolyear indicated an FTE of 26,736.58 (TEA, 2025). Therefore, there is the equivalent of nearly 27,000 full-time fine arts teaching positions statewide. However, FTE may obscure the actual number of educators; for example, two part-time fine arts teachers teaching half a courseload each would be categorized as one FTE teacher. Furthermore, no publicly available data delineates which of these teachers are employed by districts using T-TESS. Still, it was anticipated that the target population comprised thousands of fine arts educators and the administrators who conduct their evaluations.

From this population of fine arts educators and evaluators, purposive sampling was used to select at least ten fine arts educators and two evaluators for inclusion in this study (ultimately, ten fine arts educators and three evaluators were included in this study). More specifically, participants were recruited via email, phone calls, and word of mouth, with inclusion and exclusion criteria being used to select desired participants from those who showed interest until the target sample size was achieved and resources for participant recruitment were exhausted (Daniel, 2012). Purposive sampling was appropriate for this study: not only does

purposive sampling align well with qualitative methodology, but this choice also allowed individuals who were most likely to be rich sources of information to be selected and included in the study (Bedford, 2020). This study focused on understanding the perceptions of fine arts educators and their evaluators regarding the effectiveness of an evaluation tool; in order for this study to be as effective as possible, I anticipated it would be beneficial for me to work with participants representing a variety of schools and districts, fine arts content areas, and grade levels across the district. This is because one school or district may be utilizing T-TESS in a slightly different way than another school or district, and these unpredictable variances in implementation and utilization of the tool (or even differences in individual school cultures) could have influenced participant perspectives on the tool. It was therefore important that I used purposive sampling to ensure the chosen participants represented as many schools across the district as possible. Likewise, I anticipated it might have been possible that Dance teachers would have a different perspective than Choir teachers on the effectiveness of the tool. In other words, I imagined teachers of one fine arts content area might perceive T-TESS differently than educators of another fine arts content area. Purposive sampling allowed me to be selective in ensuring the participants represented as many content areas as possible. Furthermore, the identified sample size was appropriate for this study in that the sample size was manageable for the scope of the study while allowing me to be thorough in dissecting and assimilating information from participant perspectives so as to reach data saturation (Bloomberg, 2023). I anticipated that it would be feasible for me to recruit ten fine arts educators that represented a variety of fine arts content areas and grade levels, as well as two fine arts evaluators.

Participant eligibility for inclusion in the sample was based on a series of inclusion criteria. Inclusion criteria for the sampling frame included that the fine arts educator or evaluator must possess at least one year's worth of experience working in their current role for the district, be fluent in English (so as to be able to easily complete the questionnaire and one-on-one interview), have good attendance and professional standing in the workplace, and possess previous experience with the T-TESS evaluation system.

Adherence to these criteria allowed this study to focus specifically on the fine arts educators and evaluators that were at the center of this study's problem, purpose, and research questions. The problem addressed by this study revolved around the ineffectiveness of current fine arts teacher evaluation practices; targeting fine arts educators and evaluators was necessary in order to better understand the firsthand perspectives of those directly involved in and affected by the fine arts educator evaluation process. The purpose of this action research was to explore how fine arts educators and their evaluators perceive the effectiveness T-TESS as it is used in the evaluation of fine arts educators. The inclusion criteria ensured all participants had the necessary background to be able to provide perspectives on all elements of the T-TESS evaluation process. In particular, the requirement that participants possessed a year's worth of experience with their district guaranteed participants moved through the full evaluation process at least once before, which allowed participants to provide an informed perspective on the effectiveness of the evaluation tool. The research questions focused on insights regarding the usage and impact of T-TESS, as well as perceived needs related to the effective and prolonged use of the tool in the future. It was anticipated that the inclusion criteria would allow me to recruit fine arts educators and evaluators who were able to share perspectives that were relevant to the research questions. Measures were taken to protect participant rights throughout the

recruitment process and over the duration of this study. IRB approval was obtained prior to recruitment for this study, and the voluntary nature of this study was stressed, with all participants completing an informed consent form (Appendix D). Participants also maintained the right to ask questions prior to consenting to participate and were able to withdraw from the study at any time. Abiding by these inclusion criteria and measures to protect participant rights allowed me to assemble a sample of participants that was appropriate for the study, with all participants possessing an appropriate background in fine arts education (or, in the case of evaluators, experience evaluating fine arts educators), as well as a firsthand understanding of traditional evaluation practices. This allowed participants to provide their perceptions of T-TESS in alignment with the other elements of this study.

This population was also appropriate from an ID perspective; a designer would normally begin designing an evaluation tool by gathering background information (such as gauging client needs) and using feedback from stakeholders to refine the evaluation tool throughout its development; feedback would be used to ensure the final published version of the evaluation tool met all stakeholder needs while performing in line with intended metrics. In order to gain the most helpful feedback on the tool, a designer would want to assess the perspectives of those who actually interact with the tool—those who comprised this study's population. In a study such as this one, feedback obtained from participants might be used to determine such matters as if the evaluation tool is seen as equally effective by both fine arts educators and evaluators, how the tool impacts the observation and overall evaluation process, and what needs participants indicated must be taken into account in order for the tool to be used effectively over a prolonged period. Given that an anticipated outcome of this study was the establishment of an initial set of perspectives regarding the effectiveness of T-TESS tool, an instructional designer

could use feedback gathered from participants to ensure that future versions of this tool contained intentional modifications meant to address the concerns and needs of the tool's users and other stakeholders. Of course, meaningful development concerning the design and implementation of the evaluation tool would not be able to be achieved without taking the needs of the tool's users – particularly fine arts educators and evaluators—into account, thus further justifying the choice in population for this study.

Materials/Instrumentation

This study made use of both questionnaires and semi-structured, open-ended interview guides (Appendices A and B, respectively). The two participant groups (evaluators and fine arts educators) received nearly identical copies of each tool, with only slight modifications made to the wording of questions to allow the instruments to be geared towards the appropriate participant group. The questionnaires, consisting of five items, were used to capture initial reactions to the participants' interactions with the evaluation tool.

The questionnaire responses helped inform my interview with each participant. While the semi-structured interview guide provided a starting point for interviews, the questionnaire responses allowed me to assess if additional items needed to be included in the interviews. I anticipated the data I collected from the questionnaires would allow me to approach the individual interviews from a more informed standpoint, allowing me to gather richer data than if I had not utilized a questionnaire. The questionnaires used in this study assessed perspectives surrounding interactions with T-TESS, specifically positive and negative interactions with the tool, helpful and confusing aspects of the tool, and any surprising elements regarding the usage and impact of the evaluation tool. Collecting this sort of initial information allowed me to gain a general sense of how fine arts educators and their evaluators perceive the usage and impact of the

tool (relating back to RQ1 and RQ2), and these general perceptions were further dissected in the individual interviews to shed further light on the research questions.

The semi-structured interview guides were developed based on Castillo-Montoya's (2016) Interview Protocol Refinement Framework (IPR), which facilitates the ability to obtain quality data while strengthening reliability in interview protocols used in qualitative research. In following Castillo-Montoya's (2016) recommendations, I ensured that the interview questions aligned with the research questions; I also constructed the interview guide such that the interview questions maintained a conversational tone while deeply probing into a participant's perceptions and feelings. A matrix was used to map out interview questions, ensure alignment between research questions and interview questions, and allow for manual note-taking during interviews (Castillo-Montoya, 2016). The interview guides, which each contained 15 questions, and the note-taking matrix are included (see Appendix C). As I constructed my data collection tools, I was sure to seek and receive feedback on the data collection protocols prior to conducting this study. I had one associate principal who served as a fine arts evaluator and two fine arts educators (one of whom taught Visual Arts at the high school level, and one of whom taught middle school Dance) review and provide feedback on both the questionnaire and interview protocol. These individuals possessed five to fifteen years of experience each working in their respective fields for the Texas school district in question and had a plethora of experience navigating T-TESS over the course of their careers, allowing them to view and critique the data collection tools from an informed background. The associate principal also possessed experience in constructing and administering qualitative surveys among fine arts staff throughout the district, while both teachers had basic experience in collecting and analyzing qualitative data from student surveys. These non-participants possessed the niche background necessary to

provide valuable feedback on and suggestions for changes to the data collection tools; in this way, these non-participants helped me conduct a field test of the data collection protocol to ensure the appropriateness of the questionnaire and semi-structured interview guide prior to the implementation of the tools in the study.

The questions contained in the semi-structured interview protocol may be divided into three types: demographic questions (such as, “What is your current role in the district?”), key questions (such as, “What role, if any, does the evaluation tool play in your choice of teaching strategies, both when being observed and in your daily teaching (when not being observed)?”), and closing questions (such as “Do you have any other comments you would like to share about the fine arts educator performance evaluation tool?”). However, I remained flexible in using probes and follow-up questions as deemed appropriate in the moment, as suggested by Roberts (2020). For example, I cued participants to continue speaking by nodding or using a verbal cue (such as, “go on”). I also asked participants to provide further detail, walk me through an experience, or explain perceptions on a given topic. These question types allowed the interviews to maintain a conversational tone (as suggested by Castillo-Montoya, 2016), while also allowing participants to provide detailed perspectives. IRB approval was obtained prior to the start of the data collection process.

Data Collection and Analysis

The following sections discuss the pathways for engaging in data collection and analysis. Such components as data collection tools, timeline, and steps for appropriate data analysis are outlined as they pertained to this study.

Data Collection

Approval from the National University IRB was obtained prior to the start of this study. The recruitment process began with email solicitation; I reached out to fine arts educators and evaluators of fine arts educators working for various Texas school districts known to utilize T-TESS using the emails publicly listed on each school district's website. I also recruited participants on Facebook. Individuals who were interested in participating responded via email, and I verified inclusion criteria (taking participant's written word that they met all criteria) prior to sharing the informed consent (Appendix D) with each participant. Participation in this study was voluntary. Participants reserved the right to ask questions prior to consenting to participate and maintained the right to withdraw from the study at any time. The recruitment process took approximately two weeks.

Each participant completed a five-item questionnaire (see Appendix A) via Qualtrics. The questionnaire consisted of five questions to be completed after concluding the T-TESS cycle for the school year. A link to access and complete the questionnaire was emailed to participants, and the researcher followed up via email as needed to ensure participants completed the questionnaire. Once a participant's questionnaire was been received, the participant was scheduled for a one-on-one Zoom interview (see Appendix B for the semi-structured interview guide). Interviews lasted between 30 and 60 minutes and took place within two weeks of the participant's completion of the questionnaire. I sent an email to remind participants of their interview time and date. The semi-structured interviews were recorded through Zoom to allow for transcription via NVivo.

To heighten credibility, I engaged in member checking (Creswell & Poth, 2018) by sending each participant's transcribed interview back to them within three days of their

interview. Participants were asked to confirm the transcript accurately captured their perspectives within three days of receiving the transcript. Additionally, I maintained a personal journal (see Appendix E) to encourage reflexivity throughout the data collection process. All electronic files, including interview recordings, transcripts, and questionnaires, were saved to my password-protected computer; any resulting paper documents will be kept in a locked cabinet for three years in accordance with IRB requirements.

Data Analysis

Thematic analysis of the data was conducted once all interviews were transcribed with NVivo and affirmed by participants. As I conducted my analysis, I moved through Clarke and Braun's (2013) six steps for thematic analysis. After familiarizing myself with the data (the first stage of analysis) and taking general notes, I moved on to Clarke and Braun's second defined phase of thematic analysis: coding. I used NVivo to code the data inductively; inductive coding allowed the data analysis process to unfold organically, rather than forcing a collection of themes onto the dataset (Willgens et al., 2016). As I used NVivo, I utilized description-focused coding, extracting relevant statements from the data, taking the time to understand each chunk of data, and labelling each piece of data so that I could easily relocate the data chunks as needed. From here, I attempted to discern potential themes (the third phase of analysis), returning to the coding process as needed to best represent the data. More specifically, I organized the coded data into clusters of related information and discerned what overarching themes could be found among each cluster (Clarke & Braun, 2013; Demystifying Research Methodologies, n.d.). It was anticipated that this approach would prove helpful as I worked to construct the most accurate representation of the data set. A visual map was also utilized in conjunction with coding to connect and solidify themes (see Appendix H). Only once I reviewed the themes to ensure the

themes provided the most complete and truthful meaning from the dataset (phase four—reviewing themes) did I begin to construct a narrative that analyzed each theme (phase five—defining themes) and reported my findings (phase six—writing the narrative). As I moved through the analysis process, I revisited the various phases more than once to build the most compelling and appropriate set of findings possible. I was sure to acknowledge that, as the researcher, I played an active role in determining and reporting the themes associated with the study's findings (Braun & Clarke, 2022). Throughout the data collection and analysis processes, each participant was assigned a number and letter identifier, with the letter "E" connected to evaluators and "T" connected to teachers, such that E1 referred to an evaluator, while T1 indicated a teacher. These identifiers were used in all notes and documentation related to the study to ensure participant anonymity.

Assumptions

This study relied on a series of assumptions. For example, it was assumed that participants answered all questions truthfully and participated in the study with integrity, rather than trying to align their questionnaire and interview responses with what they believed to be "desirable" replies. In other words, I assumed participants would trust that their anonymity would be protected and would not feel pressured to reply to the questionnaire or interview questions in a way that they feel would be preferred by any other stakeholders. Depending on the individual, a participant might have felt that a "desirable" reply was more positive or negative than their genuine feelings; any deviation from an honest response would damage the accuracy of the findings. With this said, the overall value of this study was dependent on the assumption that participants will practice integrity. This assumption stemmed from my belief that participants would not or agree to participate in this study if they were not genuinely willing to

contribute honestly, as requested in the informed consent, in the instructions for the questionnaire, and at the start of each interview.

A second assumption was that the inclusion criteria would allow the sample to be appropriately representative of the population. As previously discussed, the inclusion criteria were meant to ensure participants possessed the necessary background, school standing, and language skills to provide meaningful insight into the research questions. Minimal inclusion criteria were outlined for this study to allow for the largest number of participants possible; this, combined with the use of purposive sampling, was anticipated to allow the participant population to accurately represent the population. A third assumption was that participants would genuinely be interested in participating in the study and that this study would yield enough participants to achieve data saturation. This assumption was linked to the previous assumption: if inclusion criteria were strict enough to attract quality participants while still allowing for the largest number of participants possible to be recruited, it stood to reason that enough participants would be recruited to achieve data saturation. The assumption that participants would be genuinely interested in participating is rooted in the idea that, in general, people do not commit to studies in which they have no interest. Furthermore, it is assumed that all participants would possess the knowledge and ability to access both Zoom and Qualtrics to complete the interview and questionnaire. Working for a school district requires one to have a basic understanding of technology; many school-related meetings are held via Zoom, so it made sense to assume participants would be able to navigate a Zoom interview and survey hosted on a user-friendly platform such as Qualtrics. Of course, it was also assumed that the results of this study would prove useful in answering the outlined research questions and in establishing an initial set of perspectives regarding the usage and impact of the evaluation tool. This final assumption was

based in my personal aims when it comes to producing meaningful research, as well as my faith in the National University dissertation process. In other words, I was confident that my coursework leading up to and carrying me through my dissertation adequately prepared me to execute research in such a way that I would be able to produce quality study results that connected to my research questions and problem and purpose statements.

Limitations

There were several limitations associated with this study. This study was limited to fine arts educators and evaluators of fine arts teachers working for the school districts that use T-TESS. The study was limited because it cannot provide evidence relating to the perceptions of fine arts educators or evaluators of fine arts teachers working for districts that do not implement T-TESS, as these parties fall outside the participant pool. The inclusion criteria could have affected the ability to achieve data saturation and, by extension, the transferability of the study, if an insufficient number of participants were recruited. However, multiple emails were used to reach out to potential participants, with word of mouth and and social media used to maximize the participant pool. Additionally, given the nature of action research, certain defining features of this study may have prevented findings from being transferable to settings that do not closely mirror that of this study. Since action research focuses on a specific phenomenon occurring in a singular setting, this study is limited in that it is unable to present findings that are universally applicable to other settings. To mitigate this limitation and increase the trustworthiness of the findings, I provided detailed accounts of all procedures, background, and data pertaining to this study so as to proactively indicate when and how the findings may be transferrable to other studies and situations. While it was assumed that participants would be honest in completing the interviews and questionnaires, participant honesty and openness could have posed a limitation in

this study if participants felt pressured to provide responses that they believed were correct. This limitation could have arisen if participants felt pressured to show overwhelming support for their school district and their district's implementation of T-TESS (leading to overly positive perspectives regarding the evaluation tool), or if they assumed their participation in the study was meant to help locate flaws in the evaluation system (leading to overly negative perceptions regarding the evaluation tool). While all questionnaire and interview questions were worded in a non-biased and non-leading manner, and while participants were reminded that their honest feedback was valued, the sort of pressure that may lead to dishonest or exaggerated participant input may have come from other sources. These sources may have included individual personality traits and tendencies, fear that one's anonymity would be compromised in the reporting of results, or relationships with one's faculty and staff (such as a principal that dictates all staff must put on a face of overwhelming support for their district). In an effort to encourage honesty and openness among participants, I took advantage of the semi-structured nature of the interviews to ask clarifying questions and confirm the accuracy of my understanding of participant experiences; member checking was also utilized to mitigate this limitation. Finally, since qualitative analysis ultimately relies on the interpretation of the researcher, my own positionality may have posed a limitation to this study, though I engaged in journaling and reflexivity throughout the course of the study in the hopes of addressing any biases or pre-conceived notions.

Delimitations

One possible delimitation associated with this study involved the requirement that all participants must have worked for at least one year in their current role within their district. This decision regarding inclusion criteria was necessary to ensure that all participants had gone

through the T-TESS cycle of evaluation at least once, allowing participants to provide detailed perceptions on how the evaluation process, in line with the research questions. Another delimitation was that neither of the chosen methods for data collection (questionnaire administered electronically through Qualtrics and Zoom interview) involved in-person interaction. This choice to utilize digital means of data collection was intended to allow the participants to complete the questionnaire on their own schedule and to participate in the interview in whatever setting they find most comfortable; it was anticipated that these choices would allow participants to be relaxed during the data collection process, which in turn would yield richer data than might otherwise be obtained in a more clinical setting. Yet another delimitation was posed by the conceptual framework of this study. It was anticipated that this study would provide insight into how Eisner's concepts of connoisseurship and criticism could inform the design and implementation of evaluative tools in the field of fine arts education. While a study focusing on the implementation of T-TESS could have been crafted based on a different framework, the choice to situate this study's framework in Eisner's ideas regarding connoisseurship and criticism aligned well with my studies in ID, as well as the problem and purpose of this study.

Summary

The purpose of this action research was to explore perceptions held by evaluators and fine arts teachers regarding the effectiveness of T-TESS in terms of how well the tool achieved validity and alignment with expected performance goals in the evaluation of fine arts educators. Participants were fine arts educators and evaluators of fine arts educators who have been working in their current position for at least one year; the researcher recruited thirteen participants (ten educators and three evaluators) through use of emails, social media, and

word of mouth. Data was collected digitally through five-item questionnaires administered through Qualtrics and one-on-one semi-structured Zoom interviews. Interviews were transcribed and data was coded with the aid of NVivo; the researcher engaged in thematic analysis to make sense of the data. Assumptions, limitations, and delimitations were considered in order to enhance trustworthiness and outline necessary mitigation procedures.

Section 3: Findings, Implications, and Recommendations

The problem addressed by this study was the ineffectiveness of current fine arts teacher evaluation practices (Gates et al., 2015, Potter, 2021b, Tutt, 2018). Fine arts teachers often face a variety of struggles related to the evaluation process, such as administrators who do not possess content knowledge of the classes they are observing, heightened focus on standardized testing that competes with the subjective nature of fine arts classes, and issues with consistency in performance evaluations (Jones et al., 2022; Kimelberg et al., 2019; Lin, 2013; Tutt, 2018). Under the Texas Teacher Evaluation and Support System, or T-TESS, observations count for 70% percent of one's evaluation score (TEA, 2015), highlighting the need for research efforts focused on better understanding perceptions of T-TESS, particularly as it pertains to fine arts evaluations. The purpose of this action research was to explore perceptions held by evaluators and fine arts teachers regarding the effectiveness of T-TESS in terms of how well the tool achieves validity and alignment with expected performance goals in the evaluation of fine arts educators. Validity was operationally defined as the extent to which performance evaluation rankings were perceived as fair and were supported by underlying reasoning.

This section provides findings of the perceptions of fine arts educators and evaluators of fine arts educators regarding the effectiveness and impact of T-TESS. This section also details participant suggestions for additional supports and practices to improve the long-term use and effectiveness of T-TESS in assessing fine arts educators. The data presented in this section was collected through participant completion of five-item questionnaires (via Qualtrics) and one-on-one Zoom interviews; data was thematically analyzed with the aid of

the tool NVivo. This section also presents an evaluation of outcomes, implications and recommendations for practice, and recommendations for future research.

It is important to note that several limitations may have influenced the interpretation of findings. One notable limitation concerns the participant pool for this study. Ten fine arts educators and three evaluators participated; although efforts were made to ensure participant representation across grade levels and content areas, the perspectives displayed in this chapter may not encapsulate the views of all fine arts educators and evaluators who interact with T-TESS.

The timing of this study and availability of participants may have posed additional limitations. Data collection occurred near the end of the school year. This period was chosen intentionally, as I anticipated participants would be better able to reflect on T-TESS shortly after completing the annual evaluation cycle. However, it is possible that participant responses may have differed if data collection had occurred at a different point in the evaluation cycle.

Participant availability may also have affected the nature and depth of responses. Participants may have been managing competing demands or personal stressors, which may have affected the information they were willing to share. Member checking was utilized in an effort to mitigate this limitation and support the trustworthiness of the findings.

Findings

This section details findings obtained through the analysis of five-item questionnaires and semi-structured interviews completed by all participants. The five-item questionnaires were completed on Qualtrics, and the one-on-one, semi-structured interviews were held via Zoom. The informed consent letter, which proceeded the five-item questionnaire, was also reviewed at

the start of each interview and participants were given the chance to ask questions or stop the interview at any time. All interview transcripts were saved to my password-protected computer, with alphanumeric participant identifiers used to label all documents. Transcripts were inductively coded with the aid of NVivo, and I followed Clarke and Braun's (2013) six steps for thematic analysis to make sense of the data: (1) familiarizing myself with the data, (2) coding, (3) discerning potential themes, (4) reviewing themes, (5) defining themes, and (6) writing the narrative.

More specifically, I began coding by looking at the questionnaire responses to gain a foundational understanding of potential themes regarding T-TESS' effectiveness in terms of validity and alignment with expected performance goals, the impact of T-TESS, and recommendations for improving the long-term use and effectiveness of T-TESS in the evaluation of fine arts educators. The questionnaire responses indicated that participants appreciated that T-TESS provides a clear rubric and means of obtaining feedback for fine arts educators, but participants noted that fine arts teachers should not be evaluated in an identical manner to core content teachers; all questionnaire responses were succinct and did not provide examples or further description to support these potential themes. I then analyzed the interview transcripts to assess whether the interview responses aligned with those obtained from the questionnaire. The interview responses offered deeper insights, expanding on the brief answers given in the questionnaires, as the semi-structured format of the interviews was designed to elicit more detailed information from the participants. I followed an inductive coding process on my first review of the interview transcripts to prevent any potential biases from affecting my identification and definition of themes (Demystifying Research Methodology, n.d.). I then revisited my three research questions and reorganized my coding schema (adding and combining

codes) to ensure all participant views were represented and connected back to the three research questions. Each of the three research questions was translated into a broad category within the coding schema—I came up with three categories around which to base my coding schema: (1) evaluator perceptions, (2) teacher perceptions, and (3) recommendations for future practice. Following the advice of Bloomberg and Volpe (2019), I devised a series of descriptors for each category, making sure to include all themes and codes generated during my first round of coding in these descriptors. For example, the initial code "Teacher Impact on Role" was separated into five separate codes, as listed in Table 1. Likewise, the initial codes "Confusing," "Work Backwards," and "Redundancy," were all combined under the code "(2.2/FR) – Source of Frustration" in the ultimate coding schema. After several rounds of coding, during which I continued to condense, modify and add to my coding schema, I then created a visual map, which may be found in Appendix H, to connect and solidify the emerging themes. The coding schema for this study, along with associated frequency counts for each code, may be found in Table 1.

Table 1

Coding Schema and Associated Frequencies

Code	**Number of Mentions per Participant**		
(1) Evaluator Perceptions	E1	E2	E3
1.1 Effectiveness of T-TESS			
(1.1/ECR) Effective – Demonstrate Clear Reasoning for Scoring	3	2	2
(1.1/ECE) Effective – Clear Expectations		3	
(1.1/ERH) Effective if Used in the Right Hands	5	2	3
(1.1/IL) Ineffective – Language of Rubric	3	1	1
(1.1/IRS) Ineffective – Performance Rating Scale	1	2	2
(1.1/LEN) Ineffective – Too Lengthy	2		
1.2 Impact of T-TESS			
(1.2/FIN) Financial	1	2	
(1.2/EA) Evaluator Accountability	3	1	2
(1.2/TA) Teacher Accountability	2	4	2
(1.2/TG) Teacher Growth	5	2	2
(1.2/NI) Negligible Impact on Employment	2	2	1

Code	Number of Mentions per Participant									
(2) Teacher Perceptions	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10
2.1 Effectiveness of T-TESS										
(2.1/ESC) Effective – Self-Check on Best Teaching Practices	1	2	2	1	1		1		1	2
(2.1/DEB) Effective – Dependent on Evaluator Background		2	1	2	1	2	1	1		
(2.1/IIN) Ineffective – Inaccurate Means of Assessing Fine Arts	3	3	2	4	3	3	2	2	2	1
(2.1/IF) Ineffective – Feedback		1		3	1	2	2	2		1
(2.1/IC) Ineffective – Confusing	1			3			3			2
2.2 Impact of T-TESS										
(2.2/AP) – Affirm Practices		1	3		3		1			
(2.2/LTP) – Little Impact on Teaching Practices	1	1	1	3	2	1	2	1	1	2
(2.2/FR) – Source of Frustration	2	1	2	1	2	1	4	1	2	2
(2.2/TG) – Teacher Growth		1		1	2		3		1	1
(2.2/TA) – Teacher Accountability		2	1	1	2	1	2		1	2

Code	Number of Mentions per Participant												
(3) Recommendations for Future Practice	E1	E2	E3	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10
(3.1/FAE) – Fine Arts Exposure	3	2	1	1	2	1	1	1	1	2	1	2	1
(3.2/CA) – Communication/Advocacy	2	5	4	1				5	1	2	1		2
(3.2/FI) – Financial Incentives	1	1		2				1					
(3.2/FAR) – Create Supplementary Fine Arts Rubrics	2	1	1		2	1					1	1	
(3.2/CCE) – Consult Content Experts		1	5							1	6	1	

Demographics

This study examined the perspectives of 13 participants—three evaluators and ten fine arts educators. The participants, six of whom were female and seven of whom were male, ranged in age from 31 to 72 and had been teaching in the public school system for 4 to 51 years. The fine arts educators taught a variety of subjects at the elementary, junior high, and high school levels, including Piano/Music Theory, Theatre, Instrument Repair, Choir, Orchestra, and Dance.

Participant identifiers consisting of a number and a letter ("E" for evaluator and "T" for teacher) were used to ensure confidentiality and protect identities. Further details concerning participant demographics may be found in Table 2.

Table 2

Participant Demographic Data

Participant	Gender	Age	Position Held	Grade Level	Years in Public Education
Evaluator 1 (E1)	Female	50	Associate Principal	High	15
Evaluator 2 (E2)	Male	47	Assistant Principal	Elementary	11
Evaluator 3 (E3)	Female	48	Associate Principal	High	18
Teacher 1 (T1)	Male	52	Visual Arts Teacher	High	26
Teacher 2 (T2)	Female	72	Piano/Music Theory Teacher	High	51
Teacher 3 (T3)	Male	47	Piano/Music Theory Teacher	High	5
Teacher 4 (T4)	Male	52	Instrument Repair Teacher	High	10
Teacher 5 (T5)	Male	36	Theatre Teacher	High	10
Teacher 6 (T6)	Male	48	Choir Teacher	Junior high	11
Teacher 7 (T7)	Female	49	Dance Teacher	Elementary	22
Teacher 8 (T8)	Female	47	Dance Teacher	High	8
Teacher 9 (T9)	Female	32	Orchestra Teacher	Elementary	10
Teacher 10 (T10)	Male	31	Dance Teacher	High	4

Trustworthiness of the Data

Throughout the course of this study, intentional efforts were made to uphold all four trustworthiness criteria: credibility, dependability, transferability, and confirmability. The following sections provide a brief overview of each criterion, along with an explanation of steps taken to address each.

Credibility. Credibility refers to the believability of a qualitative study, in the sense that findings are not only believable to readers, but also receive the approval of the participants involved in the study (Bloomberg, 2017). Credibility may also be thought of as a clear link between the respondents' contributions to the study and the researcher's representation of these contributions (Nowell et al., 2017). I actively worked to build trust with participants throughout the data collection process by encouraging a sense of honesty and using iterative questioning (Shenton, 2004). Although participants did not request any changes to their transcripts, I shared interview transcripts with each participant and offered the opportunity for participants to request transcript contents be amended; member checking allowed me to ensure accurate documentation of all participant perspectives (Creswell & Poth, 2018). The use of both questionnaires and semi-structured interviews as data collection methods allowed me to further bolster the credibility of my findings through triangulation. Finally, I maintained a reflective journal throughout the data collection process to maintain an awareness of biases, expectations, or assumptions that could have influenced the credibility of the study (Bloomberg, 2019).

Dependability. The phrase dependability refers to the consistency of the inquiry processes as they are used over time (Bloomberg, 2017). Dependability indicates that one's research process is traceable, logical, and well-documented, allowing readers to judge for themselves if research was conducted in a reliable manner (Nowell et al., 2017). I addressed

dependability by clearly outlining my research process throughout this study (Creswell & Poth, 2018). Similarly, I maintained a "paper trail" of field notes and journaled throughout the data collection process to ensure my research process was documented accurately and in sufficient detail.

Transferability. Transferability may be thought of as the applicability of inquiry or findings from one context to another on a case-to-case basis (Bloomberg, 2017; Nowell et al., 2017). I strove to utilize clear descriptions in my explanation of the research process and in my detailing of findings to allow readers to judge transferability for their own individualized contexts (Nowell et al., 2017). I also proactively indicated ways in which my findings may be applicable to other contexts or populations towards the end of this chapter (Bloomberg, 2019; Bloomberg & Volpe, 2019).

Confirmability. The final trustworthiness criterion, confirmability, is linked with the other criteria mentioned above in the sense that confirmability is established once the other three criteria are achieved (Nowell et al., 2017). Confirmability indicates the degree to which the findings of a study are determined to accurately reflect the data, rather than the biases or interests of the researcher (Frey, 2018). Confirmability has been addressed below through a detailing of how my findings align with the existing body of research into fine arts teacher evaluations; I anticipated this would illustrate that my interpretations are not based in personal biases. Referencing relevant sources in my literature review, engaging in member checking, and taking detailed field notes contributed to the confirmability of my research (Bloomberg, 2019; Nowell et al., 2017).

Research Question 1

How do public-school evaluators of fine arts educators perceive the effectiveness and impact of T-TESS in evaluating fine arts educators? This study aimed to understand perceptions of evaluators concerning the perceived effectiveness of T-TESS in terms of validity (defined here in terms of perceived fairness and reasoning underlying performance evaluations) and alignment with expected performance goals in the evaluation of fine arts educators, as well as the perceived impact of T-TESS. The perceptions of the three evaluators who participated in this study centered around two themes connected to the effectiveness of T-TESS, and three themes related to the impact of T-TESS. Table 3 details a summary of these themes; a discussion of each theme follows.

Table 3

Themes Pertaining to Research Question 1

Research Question	Organizing Principle	Themes
RQ1. How do public-school evaluators of fine arts educators perceive the effectiveness and impact of T-TESS in evaluating fine arts educators?	Effectiveness of T-TESS	1. T-TESS is effective for use in fine arts educator evaluations in that it allows evaluators to demonstrate clear reasoning underlying performance evaluations. 2. T-TESS' language and performance rating scale make the tool ineffective in evaluating fine arts educators.
	Impact of T-TESS	3. T-TESS has negligible impact on employment. 4. T-TESS may have a financial impact on fine arts educators. 5. When evaluators adapt an arts-minded lens, T-TESS contributes to the growth and accountability of fine arts educators.

Theme 1: T-TESS is effective for use in fine arts educator evaluations in that it allows evaluators to demonstrate clear reasoning underlying performance evaluations. All

the evaluators who participated in this study remarked that T-TESS was beneficial in evaluating fine arts educators in the sense that the evaluation system provides a clear, transparent means for assigning performance ratings. Evaluator 3 remarked that T-TESS provides a clearer rubric than its predecessor, PDAS: “I like T-TESS better than PDAS. The rubric is much clearer as an evaluator—to be able such as give concrete examples and explanations.” Evaluator 1 discussed their use of the T-TESS rubric during post-conferences to explain performance ratings:

> I always have it out. And there have been times I have had a teacher who wanted to learn to be an administrator. I had a couple that wanted to really learn or a couple that were not really amazing and they were just wanting to go on and level up or they were good and they were trying to. And so, we look at the actual rubric and I'd point out, ‘here's the exact reason why I had to score you here, instead of here.’ I've definitely used it during the post-conference because I want [teachers] to understand that I'm not just pulling this rating out of thin air. I have to follow the rubric.

Evaluator 2 agreed with this sentiment, stating

> What I like about a system that uses a rubric… I feel like as a teacher, you have the opportunity to like take the bull by the horns and be like, ‘okay, these are very concrete ways that I can improve my practice.’

Evaluator 1 elaborated that T-TESS’ recommendation to conduct multiple observations of a teacher heightens the validity of the tool in terms of perceived fairness and reasoning underlying performance evaluation rankings: “I can't just show up one time. You also can't just put on a dog and pony show the one time I come in because I'm in there multiple times giving feedback.”

Evaluator 2 was most supportive of the language used by the evaluation tool and stated

> I feel like the language is broad enough to encompass everything it needs. There's nothing on there that I would say like, 'oh, well, a fine arts teacher doesn't do that or couldn't do that.' And so, yeah, I do think it's effective.

Theme 2: T-TESS' language and performance rating scale make the tool ineffective in evaluating fine arts educators. Though evaluators appreciated that T-TESS allows for performance ratings to be substantiated by a clear rubric, all participating evaluators were quick to point out concerns with the language and rating scale used by the evaluation tool. Evaluator 1 stated that T-TESS is simultaneously too bulky and too vague to make it truly helpful in evaluating fine arts educators:

> I don't look at the T-TESS rubric when I'm in the classroom. It's too cumbersome. If I'm looking at the rubric, I'm not paying attention to the class… However, it still leaves a lot of room for discretion. And if I'm a principal who doesn't really understand what I'm seeing, it still can be hard for me to make those things match.

Evaluator 2 reflected on their time as a Theatre teacher, before becoming an administrator:

> Administrators didn't know how to evaluate me as a theater teacher using the rubric because the way it looked in a traditional core classroom was very different. The things that they were looking for as far as lesson structure and DOLs and exit tickets and evaluation, all of those things looked very different in a fine arts classroom than it did in the traditional classroom. So they would come in and either rate me really poorly because they didn't see it the same way, or they would just rate me in the middle because they would say, 'I don't really know,' which was kind of frustrating as a teacher.

All three evaluators noted the structure of the performance rating scale makes T-TESS less than effective. When asked about their thoughts on the fairness of performance ratings assigned

through T-TESS, Evaluator 3 noted, “you can make the data look however you want.” Evaluator 1 took issue with the guidance surrounding the assignment of performance ratings:

> I think that the fact that proficient is rock solid teacher, which they've just beat into our heads, still doesn't feel good even all these years later to tell someone, ‘oh, you're proficient, congratulations,’ because that feels like the bare minimum. That's not their intention, but that's what it feels like. And we as administrators are not encouraged to give higher than proficient. That's not really talked about, but we're not. And I don't agree with that because if I have a stellar teacher, I'm going to give them what I feel like they deserve. So I don't love that part of it because why can't you be accomplished or whatever? Why can't you be?

Evaluator 2 described the T-TESS performance rating scale as

> a real disadvantage for the arts teachers because they were doing fantastic teaching that really everybody should model themselves after. But it wasn't rating that way because people didn't know how to [assess fine arts educators]. You know, again, we didn't have that shared language.

Theme 3: T-TESS has negligible impact on employment. Two of the three evaluators indicated that while elements of T-TESS may be less than ideal, it does not generally impact a fine arts educator’s employment. According to Evaluator 1, “the reality is that as long as you're making proficient, it's fine. The only reason T-TESS ever is going to be an issue is if they are looking at you as someone they may not want to ask back.” Evaluator 2 expressed, “Before when I was a teacher, I got one evaluation a semester. It didn't matter anyway. At the end of the day, it doesn’t really impact you.”

Theme 4: T-TESS may have a financial impact on fine arts educators. While the evaluators who participated in this study acknowledged T-TESS does not have bearing on a fine arts educator's employment, two of the evaluators noted that in districts participating in the Teacher Incentive Allotment (TIA), T-TESS ratings can impact one's pay. Evaluator 2, who works for a district that utilizes TIA, stated

> Your evaluation directly impacts your pay. And I tell my teachers, 'I want you to get paid the most amount of money. I want you to be happy. I want you to be amazing. And I want you to get paid for it.'

Evaluator 1 also touched upon the financial impact of a fine arts educator's T-TESS rating:

> I hate the idea that how much you get or don't get is determinant on me and my evaluation. And the reality is that most districts are adding point values to each of the T-TESS indicators. So if you get proficient, you might get three points. But if you get accomplished, it's four points. And sometimes one or two points is the difference between you getting $5,000 and $10,000.

Theme 5: When evaluators adapt an arts-minded lens, T-TESS contributes to the growth and accountability of fine arts educators. Despite the evaluators' mixed feelings on the effectiveness of T-TESS, all participants agreed T-TESS positively impacts fine arts educators when evaluators adapt an arts-minded lens. All three evaluators commented on the benefit afforded by their background in the arts. Evaluator 1 stated, "I think that, you know, with fine arts you've got to take a look at the overall picture. You just can't look at scores… fine arts [evaluators] recognize the whole picture." Similarly, Evaluator 2 noted

> As an administrator, I face a little bit of an uphill battle because people say, 'oh, well, you didn't teach math,' or 'you didn't teach reading, and you come from the fine arts,' but I actually think that it's a benefit because I do go in looking for things in a different way.

Of T-TESS' impact and purpose, Evaluator 1 said "when it's done right, it's designed to be a growth model." Evaluator 3 uses the pre-conference of the observation cycle to help fine arts educators think about ways in which they would like to grow by asking, "What is your ultimate student goal—for every single kid, not just for your high achievers?" Evaluator 2 leverages T-TESS as a tool to get fine arts educators to develop their teaching abilities in a way that is tied to the evaluation tool's rubric:

> In art, for example, one of the things I pushed our art teachers to do was come up with a way for students to evaluate one another's work and to self-evaluate their work. Because one of the domains on the rubric was cognitive demand, which was really about student voice, student analysis, self-reflection and group reflection, giving each other feedback… I tried to kind of shift the way that they looked at their practice and say like this is where this will shoot you up on the rubric. This is the practice that you could do that would really make sure that you're hitting all these marks.

The three evaluators each mentioned T-TESS' ability to promote accountability in fine arts educators. Evaluator 2 stated T-TESS encourages "a little bit of self-efficacy. You're able to control some of your own destiny there. It's not just the whim of somebody." Evaluator 1 remarked that T-TESS is "designed to be a partnership. It forces the teacher to do a lot more reflection on their teaching." Evaluator 3 stated they have fine arts educators preemptively reflect on their teaching before entering their post-conference so that the teachers are better prepared to discuss and take accountability for their work:

> I offer [the post-conference] as 'what pieces am I missing from my documentation, from my observation'… and we walk-through each of the bullet points of what is between 'proficient' and 'accomplished,' and a lot of times it's not so much that I've missed something; it's that they didn't meet the criteria.

Research Question 2

How do fine arts educators perceive the effectiveness and impact of T-TESS in evaluating their performance? Much like the first research question, this question focused on both effectiveness and impact of T-TESS. Analysis and coding of the data revealed three themes related to the effectiveness of T-TESS, and three themes related to the impact of T-TESS. These themes are outlined in Table 4 and are explored in detail in the following sections.

Table 4

Themes Pertaining to Research Question 2

Research Question	Organizing Principle	Themes
RQ2. How do fine arts educators perceive the effectiveness and impact of T-TESS in evaluating their performance?	Effectiveness of T-TESS	1. T-TESS is effective in helping teachers ensure they are utilizing best practices in their teaching. 2. T-TESS does not provide an accurate or fair assessment of a fine art educator's teaching. 3. T-TESS does not provide adequate feedback for fine arts educators.
	Impact of T-TESS	4. T-TESS has little impact on the teaching practices one employs. 5. T-TESS is a source of frustration for some fine arts educators. 6. T-TESS can help facilitate growth and accountability.

Theme 1: T-TESS is effective in helping teachers ensure they are utilizing best practices in their teaching. Seven of the ten teachers who participated in this study noted the clear design of T-TESS allows them to ensure they are incorporating teaching strategies needed to meet desired performance goals. Teacher 7 noted, "I did often refer back to the rubric to make sure that I was demonstrating effective teaching strategies according to T-TESS standards." Similarly, Teacher 10 stated,

> There is a side of it where I can see how I am almost evaluating myself and it's forcing me to think about my teaching… The ultimate goal is to make sure… that I am feeling confident in what I'm doing.

Teacher 3 indicated T-TESS can help newer teachers gain a sense of clarity and confidence. in their teaching. They said, "I actually had [the rubric] posted on my wall, and I did have it near my desk and I would refer to it every once in a while." Of T-TESS, Teacher 5 remarked, "It's a good check-in point to make sure that I'm still incorporating everything that I'm supposed to do that was set by the requirements of the beginning of the school year." Several teachers explained that, in addition to using T-TESS to hold themselves personally accountable, the T-TESS rubrics allow them to ensure they are also complying with evaluator expectations. Teacher 9 noted, "Sometimes I will look at the rubric to align SLO goals, DOLs, success criteria, and learning objectives, but this is only for personal use so I can ensure I am 'checking all the boxes' that administration would want to see." Teacher 5 stated, "T-TESS lets me prepare and helps me kind of get in the mindset of an [evaluator]. What are they looking for? What do I not want to get in trouble for?"

Theme 2: T-TESS does not provide an accurate or fair assessment of a fine art educator's teaching. All ten fine arts educators felt T-TESS is not an effective tool when it

comes to providing a fair performance evaluation, largely due to either T-TESS not being crafted with the fine arts in mind, or due to evaluators not having a strong grasp on the fine arts. Teacher 2 said,

> People throughout the district have been getting dinged on their observations because the administrator doesn't know what to expect... It's just not an accurate judge of your teaching abilities. I know it's not. The teacher knows when they're coming in for the most part. They're on their best behavior and… the kids know that a teacher is being evaluated also and they act differently as well.

Teacher 1 reiterated, "The kids know that a teacher is being evaluated also and they act differently as well. Whether they like you or not." Teacher 4 stated, "I think they're trying to en masse just say 'this is what the teacher needs to be,' but doesn't always take into consideration what the subject matter is." Teacher 8 noted the specificity of words used in the T-TESS rubric: "There were times when I can't do 'all,' but I can do 'many'… I don't know that arts people are getting full credit on the T-TESS in general... because the arts are so subjective." Teacher 5 said,

> Theater only has one page of state standards compared to some of the other subjects that have 25 to 30 [pages]. So for an AP or principal to hold my state standards to the same as another teacher—it's just unreasonable and not really going to help.

Teacher 6 shared that evaluators "don't know what it looks like—formative assessment. If all the kids sing 'Happy birthday,' you're formally assessing every single kid. You notice if the notes they're singing are wrong, whereas [evaluators] don't recognize if that's what you're doing." Teacher 9 commented

> I think [T-TESS] impacts fine arts teachers in a negative way because we are seen through the same exact scope as core content teachers. There are a lot of differences in

> how fine arts teachers plan and teach compared to core teachers… Most fine arts teachers have to basically develop their own curriculum and their own lesson plans… We are teaching students physical skills, mental skills, SEL skills all at once. That doesn't happen in say, a math class. We have to adjust, scaffold, reteach, rehearse and practice those skills not only with individual students but sometimes (in the case of all fine arts except for visual arts) with entire groups. Our students' learning just happens in different way and at a different pace than in a math classroom.

Theme 3: T-TESS does not provide adequate feedback for fine arts educators. Six of the teachers who participated in this study disclosed T-TESS does not give meaningful feedback for fine arts educators. Teacher 4 noted, “T-TESS at [my previous school] was a joke. I did not get any kind of feedback or guidance or anything like that, so those were really tough years.” Teacher 7 stated

> I would have somebody walk in my classroom and spend like five minutes and walk out, and follow-up interviews with the administrators were five or ten minutes long, and sometimes they didn't even talk about what they saw in my classroom.

Teacher 8 commented, “nobody's teaching you how to teach dance,” when asked if they received helpful feedback from their yearly evaluations. Teacher 5 remarked, “I’ve had a few evaluators in different places that’ve only come in for a few minutes and then were able to give me a full evaluation. And I didn't think that was very helpful.” Teacher 6 recalled an evaluator said, “I'm starting you out low [on the rubric] so that we can finish high. The purpose of this is to show growth,” in lieu of giving evaluative feedback. Teacher 10 said,

> As far as them doing feedback back to you—I don't think it helps me because sometimes you're going to see something written down that you're going to be like, 'okay, that had nothing to do with what you should've been looking at.'

Theme 4: T-TESS has little impact on the teaching practices one employs. All ten of the teacher participants felt the T-TESS rubric and associated documents have little to no impact on their teaching practices or lesson plans. Teacher 8 noted that fine arts educators "refer to T-TESS because we know it, but I think that our work goes far beyond the like the scale of what T-TESS is." When asked if T-TESS informed their lesson planning, Teacher 10 stated, "I think the appropriate answer is yeah you should read that [but] it's never been my teaching philosophy to be honest." Teacher 1 noted trying to utilize T-TESS to inform their instructional practices was "like I was working backwards." Teacher 4 also noted, "I don't use the T-TESS to guide my lesson plans or anything like that," while Teacher 9 said, "I would not say it influences my lesson planning or how I teach, as I use the years of content experience I have to create and teach my lessons." Teacher 7 remarked, "Never at the high school level did I ever use any of the information that was given to me to prepare a lesson to be frank with you."

Theme 5: T-TESS is a source of frustration for some fine arts educators. Each of the fine arts educators whom participated in this study relayed that T-TESS has caused them personal frustration. As Teacher 2 recounted,

> I can remember early on having been evaluated during a rehearsal and thinking 'that was stellar. That was perfect.' And then on the observations, then the administrator saw it in a totally different way. And I was thinking, talk about perfect start to finish. I had an opener. I had a closer. I did that perfectly and it wasn't what they saw. They were looking for things that you would not in the arts.

Teacher 3 expressed their frustration with an evaluator who deducted points from an observation because the evaluator was expecting to see practices typically associated with core classes in Teacher 3's Piano class:

> I had this idiot for an administrator who, really, I don't think she ever taught. She's not there anymore, but she dinged me on some things I did not do—think, pair share, that type of thing. And that pissed me off, especially since she came up with no background of my class.

Teacher 6 stated T-TESS can feel like "a dog and pony show," while Teacher 4 noted, "I find it to be tiresome because some of it is set up where it's not logical. It doesn't pertain to all the different genres and the different arts classes that that are offered." Teacher 7 was bothered by the redundancy within the goal-setting aspect of T-TESS: "I could've spent so much more time focusing on my teaching goals if I wasn't having to answer the same questions in three different ways." Two teachers also expressed frustration with evaluators who seemed to make up scores without going through all the steps of T-TESS. Teacher 4 stated, "I was pretty much told specifically 'don't worry about it, you know, just don't worry about it. Put something in, I'll approve it, and we'll just go on with life.'"

Theme 6: T-TESS can help facilitate growth and accountability. Though participating teachers expressed frustration with T-TESS, they also noted that T-TESS does have the potential to facilitate growth and accountability among fine arts educators. Teacher 2 said, "T-TESS could provide possibly a more fair evaluation if it was used in the right hands." Teacher 5 stated, "I think if used effectively and correctly…I think it could be really helpful for the teacher." Teacher 7 noted, "I feel like the evaluation process helps us become lifelong learners, which is what, to my understanding and my belief, is the desire of every administrator." Teacher 2 discussed how

one particular evaluator who possessed a background in the fine arts affected their perception of T-TESS, saying, “I have grown tremendously as a teacher and I think it's through a lot of [my evaluator’s] guidance and her ability to help me with that.” Teacher 10 noted T-TESS provides an opportunity for a fine arts teacher to advocate on behalf of their content area and reflect on their teaching practices:

> It gives you an opportunity to just continue to spread that awareness and help [evaluators] understand what you're doing… when you're able to put it into words… it helps you think ‘oh, you know I said this, but I didn't even talk about that.’ So it makes me think about my teaching.

Research Question 3

Reflecting on their experiences with T-TESS, what additional supports and practices, if any, do fine arts educators and evaluators identify as beneficial for improving the long-term use and effectiveness of T-TESS in assessing fine arts educators? The semi-structured interviews with the 13 participants provided the data for this research question. Analysis of the data revealed three emergent themes: (1) Regular exposure to the fine arts is critical for evaluators to better understand the teachers they are evaluating, (2) Advocacy and communication are helpful in ensuring T-TESS is used to facilitate the growth of fine arts educators, and (3) Fine arts “experts” should assist in the evaluation of fine arts educators. These themes are outlined in Table 5, and a discussion of each theme follows.

Table 5

Themes Pertaining to Research Question 3

Research Question	Themes
RQ3. Reflecting on their experiences with T-TESS, what additional supports and practices, if any, do fine arts educators and evaluators identify as beneficial for improving the long-term use and effectiveness of T-TESS in assessing fine arts educators?	1. Regular exposure to the fine arts is critical for evaluators to better understand the teachers they are evaluating. 2. Communication and advocacy are helpful in ensuring T-TESS is used to facilitate the growth of fine arts educators. 3. The evaluation of fine arts educators should be informed by the work of content "experts."

Theme 1: Regular exposure to the fine arts is critical for evaluators to better understand the teachers they are evaluating. All three evaluators and ten teachers who participated in this study felt having a background in the fine arts, or having regular exposure to the fine arts, would aid evaluators in more effectively using T-TESS to evaluate fine arts educators. Evaluator 1 noted their background in the arts helped them evaluate fine arts teachers, stating, "I was a dance teacher, so I know it's different for me." Evaluator 2, who previously worked as a Theatre teacher, indicated "I've typically been the person at every campus I've been at that evaluates those things. In my opinion, for the best when you have that background." Teacher 3 commented, "principals need to spend a day or part of a day in a fine arts class observing." Teacher 9 relayed, "I think administration could possibly take training on how to evaluate a fine arts classroom and teacher properly. Most administrators have no idea the difference between band and orchestra, drill team and dance, musical theater and choir." Teacher 4 noted the fine arts staff at their school is "fortunate to have a principal that was a fine arts teacher; I think she recognized that she wasn't always assessed fairly based on these kinds of rubrics either." Teacher 7 commented, "I felt like the [evaluation] process was way more

effective with having someone that had experience [in the arts]." When asked what steps could be taken to make T-TESS more effective when it comes to evaluating fine arts educators, Teacher 8 said,

> I think, honestly, the first and foremost would be the admin attending concerts or watching the performances, because that's where you're going to see how that educator is progressing. Because you see it through the kids and you see it through their excitement through their performance onstage and what they look like and how they're performing… The investment is so low for the admin in the ultimate goal, and the ultimate goal for us is to get those kids on stage and performing and preparing them for that—that's a huge component of what we do in the arts.

Theme 2: Communication and advocacy are helpful in ensuring T-TESS may be used to facilitate the growth of fine arts educators. Evaluators and fine arts educators noted both communication between evaluator and teacher and self-advocacy by the fine arts educator are critical for ensuring T-TESS may be used more effectively to evaluate fine arts educators. Evaluator 1 commented,

> I think it depends on the teachers being able to advocate for themselves and know how to speak to the administrator, and having an administrator that is willing to realize that you can't be exactly black and white when you have to listen to the teacher; you have to look at the overall picture.

Teacher 8 said it is important for fine arts educators to "go in with the mindset that there are things your evaluators need to know." Teacher 10 noted fine arts educators need to be

> their own advocate and have a conversation with admin because… the evaluator needs to be educated by the teacher or at least give them the opportunity to identify [the fine arts

> educator] and what the subject matter is… The best practice is scheduling a meeting with your administrator so that they can, you know, get to know you and get to know what you do.

Teacher 5 found T-TESS to be more beneficial when they proactively organized

> a one-on-one meeting with [my] AP… because a lot of the APS do not have an artistic background, so kind of helping them understand what your goals are as an educator in the classroom, and what things they should be looking for within your classroom."

Evaluator 2 stated,

> I think the biggest hurdle with fine arts teachers… is developing relationships with administrators. There is a long history of being ignored, of being pushed to the side, of having your programs cut. There's a lot of trauma that school districts have done to arts teachers. On the flip, there's a lot of ego in a lot of arts teachers as a result and also as a natural extension of being artists... I think the action step is really just creating those conversations—making sure that the arts teachers feel heard from their administrators… And for administrators to hear them and respond in kind.

Four of the participants noted that advocacy efforts should also focus on fairly compensating the work that fine arts teachers do, as T-TESS scores can are often tied to incentive initiatives. Teacher 1 said, "In a regular business, people get bonuses… They come to us every year and they add on all these things they want us to. We don't get extra money for that." Teacher 5 supplied, "In order to have teachers stay, teachers need to be fairly compensated. The time we put in as fine arts teachers is not equivalent to what we get out." Evaluator 2 noted incentives tied to evaluation scores provide "concrete ways [to earn incentives] that you can achieve by

ticking off these boxes or making sure that you're doing this. And honestly, I think that it's good practice."

Theme 3: The evaluation of fine arts educators should be informed by the work of content "experts." Evaluators and fine arts teachers alike suggested that fine arts "experts" or the work of such "experts" should be utilized to make T-TESS more effective, with several participants sharing personal examples. Evaluator 1's district offers a "partner document" written by "extremely quality teachers" to highlight best practices that should be observed in fine arts classrooms during the observation cycle. Evaluator 2 said,

> When I worked with my fine arts team, it was really about, well, what's a way to formalize that so that you could show an administrator if they walked in. So I really leaned into the idea of using rubrics and even just creating a general rubric for your year.

Teacher 9 suggested, "Having seasoned and successful fine arts directors provide districts with rubrics, training for admin, would really help." Teacher 2 said T-TESS should "use the UIL evaluator" to assess fine arts educators, referring to the University Interscholastic League that oversees the vast majority of performative public-school contests. Teacher 8, who currently works as a third-party contractor and fine arts instructional coach for several school districts, shared that their company supplements the work of campus evaluators by offering informed feedback. Teacher 8 remarked, "I go in once a month and I observe each teacher for their class day and then I give feedback… your principal [or] your observer should know these are the things they should be looking for." Evaluator 3 brings additional parties on walk-throughs to ensure they are able to provide accurate feedback:

> I take my district coordinators with me. I take my instructional coaches with me and we can walk through all of the information… They're the expert… I don't know that I always

> know the curriculum, if [the teacher is] going too deep into the curriculum or if they're being too shallow, and so I do all of my walk-throughs with a district representative that comes from curriculum instruction.

Evaluation of the Outcomes

The emergent themes identified in the findings are closely aligned the study's problem and purpose, as well as to the conceptual framework and the body of literature outlined in Chapter 1.

Evaluation of Research Question 1: How do public-school evaluators of fine arts educators perceive the effectiveness and impact of T-TESS in evaluating fine arts educators?

All three evaluators noted that although T-TESS offers a widely accepted rubric that allows evaluators to articulate clear reasoning underlying performance ratings, the language of T-TESS and the rating scale are ineffective for evaluating fine arts educators. All three evaluators came from fine arts backgrounds and used phrases like "overall picture" and "guide map" when describing how they used an arts-minded lens to provide what they viewed as more meaningful and accurate evaluative ratings and feedback to fine arts educators. The three evaluators agreed that, although T-TESS may not have been structured with fine arts educators in mind, in the correct hands, the tool can be used to facilitate growth and accountability among fine arts educators. Two evaluators also noted that when tied to an incentive such as TIA, T-TESS may have a financial impact on fine arts educators, though the results of one's performance evaluation would rarely affect one's employment.

Eisner's concepts of connoisseurship and criticism formed the framework for this study; The findings from this study demonstrated that T-TESS may be able to be used in a more meaningful manner in the evaluation of fine arts teachers if users of this evaluation system apply

an arts-minded lens to their use of the tool. These findings supported Eisner's concepts of connoisseurship and criticism: all evaluators who participated in this study indicated their belief in the importance of being knowledgeable about the content area one is evaluating so as to provide feedback that is valid in terms of perceived fairness and ability to be supported by clear reasoning. These findings are in alignment with the problem and purpose of this study, as evaluators acknowledged the limitations of the T-TESS tool and the way in which the presence of subject-specific expertise ensures that evaluations of fine arts educators are valid, meaningful, and conducive to the educators' professional growth.

These findings support the body of literature that formed the foundation for this study, while expanding on previous studies to provide new information regarding evaluator perceptions of a specific evaluation system (T-TESS) as it is used with a particular group of teachers (fine arts educators). Reid (2018) found that administrator training marginally influences how evaluators assess teachers, with evaluators instead drawing confidence from their own experiences to make appropriate judgments during the evaluation process. Indeed, the evaluators in this study indicated they leveraged their personal backgrounds and previous experiences as fine arts educators to inform their application of the evaluation tool. In a 2023 study, Fresko and Levy-Feldman found most principals (evaluators) believed the evaluation process to be characterized by transparency, though they also noted the evaluation process is not always comprehensive or appropriately detailed— this study yielded similar perspectives.

Evaluation of Research Question 2: **How do fine arts educators perceive the effectiveness and impact of T-TESS in evaluating their performance?**

After acknowledging the necessity of having a teacher evaluation system in place (as Teacher 4 stated, "If it's not T-TESS, it's going to be something else), eight fine arts

educators relayed that the design of T-TESS allowed them to easily and effectively ensure they were utilizing best practices in their teaching, which in turn guaranteed they were meeting administrative expectations. The ten teachers universally agreed that T-TESS is not effective when it comes to providing fair or accurate assessments of fine arts educators; seven teachers highlighted the evaluation tool's inability to provide adequate feedback for fine arts educators. When asked how T-TESS impacts the work of fine arts educators, the participants noted that T-TESS has little impact on the manner in which they teach. While all ten evaluators relayed the evaluation process as a whole can be a source of frustration, eight teachers noted T-TESS can be a positive tool in aiding growth and accountability. Seven fine arts teachers noted that the specific evaluator with whom they were completing the T-TESS cycle played a part in determining whether the evaluation process would truly yield actionable and meaningful feedback.

Much as was the case for RQ1, the teachers' experiences and perceptions supported Eisner's theories of educational connoisseurship and criticism. Participants said the evaluation process was more beneficial when they had evaluators who were connoisseurs of the arts, as this background allowed evaluators to view the educators' work from an informed perspective and provide meaningful feedback (Nordin and Wahlström, 2019).

The outcomes of RQ2 support Tutt's (2018) recommendation that different avenues must be explored to ensure fine arts educator's skills are measured more accurately and that the evaluation of fine arts educators is able to take place in a more productive and meaningful manner. This study also echoes the notion that evaluators of fine arts educators may inflate evaluation ratings of teachers they want to maintain positive relationships with (Close & Amrein-Beardsley, 2018; Jones et al., 2022). For example, Teacher 4 noted one of his evaluators

told him, "'don't even concern yourself with [T-TESS]; you know I've got this taken care of,' and so there was never that formal observation and assessment," and Teacher 6 said, "my first evaluation I ever got was by the principal. I'd already been to her house five times, you know, because of [my wife] so I got a glowing one."

Evaluation of Research Question 3: Reflecting on their experiences with T-TESS, what additional supports and practices, if any, do fine arts educators and evaluators identify as beneficial for improving the long-term use and effectiveness of T-TESS in assessing fine arts educators?

All thirteen participants communicated that regular exposure to the fine arts is critical for improving the long-term use and effectiveness of T-TESS in assessing fine arts educators, as increased knowledge and awareness of the fine arts would allow evaluators to develop a heightened understanding of the teachers they are evaluating. Nine participants stated that maintaining honest and open lines of communication between evaluator and educator is important for ensuring genuine teacher growth occurs as a result of the evaluation cycle. Another emergent theme was the importance of incorporating fine arts-specific expertise into the evaluation of fine arts educators, whether through the use of a supplemental rubric tailored to the arts or through consultation with district-level or third-party content experts.

These findings support this study's conceptual framework, which at its core asserts that connoisseurship of a given content area is a necessary characteristic for teacher evaluators to possess in order to critique teacher performances accurately. Furthermore, these findings support and contribute to the existing literature. Willey (2019) and Garet et al. (2020) implored future teacher evaluation tools to treat educators as masters of their content areas. The findings of this study echo this sentiment and further the conversation surrounding fine arts evaluations by

suggesting that third-party expertise may enhance the accuracy and value of feedback given to fine arts educators. Potter (2021b) focused on the need for fine arts educators and education leaders to work together to bring attention to the need for more effective fine arts educator evaluation systems. While the outcomes of this study also heavily emphasized the need for communication and collaboration between fine arts educators and school leaders (more specifically, evaluators), this study's findings focused on smaller-scale, one-on-one advocacy efforts. Though Potter's (2021b) call for state and national level change remains important, this study contributes complementary perspectives to the conversation by focusing on proposing feasible ways in which fine arts educators may improve the effectiveness of their own evaluations within the currently existing T-TESS evaluation system.

Implications and Recommendations for Practice

While fine arts educators and evaluators recognize the necessity of teacher evaluation systems like T-TESS in promoting teacher growth and accountability, both evaluators and fine arts educators reported that T-TESS ultimately has a minimal impact on employment decisions or instructional practices. All ten fine arts educators noted that T-TESS does not offer educators an accurate assessment of their performance or provide meaningful, content-specific feedback; however, seven teachers and all three evaluators pointed out that the exception to this would be when evaluators possess a background in the fine arts and approach the evaluation process with an arts-minded lens. This study's findings corroborate Lin's (2013) suggestion that evaluators of fine arts educators should bolster their knowledge of the fine arts. The emergent data from this study also support the need for fine arts educators to continue advocating for more relevant and effective fine arts teacher evaluations, aligning with Tutt's (2018) recommendations. RQ3, in particular, focused on potential recommendations for improving the long-term use and

effectiveness of T-TESS as it is used in the evaluation of fine arts educators; the emergent themes from this research question formed the basis for the recommendations described below.

Recommendation for Practice: Fine Arts Exposure for Evaluators

The first recommendation is to provide evaluators with experiences to deepen their knowledge of the fine arts. The participants in this study shared anecdotal evidence suggesting evaluators who had heightened background knowledge of the fine arts were able to offer more meaningful feedback and accurate evaluation scores. Yet, none of the participants described existing programs designed to help evaluators develop a better understanding of the fine arts; it is difficult for those lacking background knowledge in the fine arts to evaluate fine arts teachers reliably (Yang, 2023). Therefore, evaluators would benefit from attending professional performances and demonstrations, including as art galleries, instrumental and choir concerts, art museums, dance concerts, musicals, and plays. It would also benefit evaluators to attend trainings and conferences focused on the fine arts and teacher evaluations within these areas, such as those held by the Texas Dance Educators Association (TDEA), Texas Art Education Association (TAEA), Texas Music Educators Association (TMEA), Texas Fine Arts Administrators (TFAA), and Texas Commission on the Arts (TCA). However, given that evaluators are inundated with a large number of responsibilities that often interfere with evaluative duties (Jonest et al., 2022; Lin, 2013), any such professional development must be funded by school districts and scheduled during working hours. This would allow evaluators to dedicate their full attention to improving their knowledge of and appreciation for the fine arts without having to balance additional work responsibilities, ultimately leading to more effective and meaningful evaluations of fine arts educators.

Recommendation for Practice: Enhanced Communication

The second recommendation is to create district- and state-level training modules aimed at establishing open lines of communication and mutual understanding between fine arts educators and evaluators. Teachers emphasized the importance of proactively engaging with evaluators to explain the distinctive aspects of fine arts instruction, which often differ significantly from core subjects. However, this practice of self-advocacy was generally characterized as a learned strategy that fine arts educators adopted after gaining experience with the T-TESS evaluation system, rather than as an instinctive approach employed during their first year of teaching. Though TEA recommended all educators and evaluators engage in a pre-conference to talk through a teacher's upcoming observation (TEA, 2022b), the pre-conference is not mandatory and does not provide structured guidance tailored to the needs of fine arts educators. Potter (2021b) advised fine arts educators to collaborate with school leaders, while Garet et al. (2020) recommended future evaluative reforms strive to consider the emotions of educators. With these recommendations in mind, it would be beneficial for both fine arts educators and evaluators to access district- and state-level training modules that address the unique teaching and evaluative considerations for fine arts classrooms. These modules should provide questions and collaborative activities to promote open dialogue between evaluator and fine arts educator. Unlike the current recommended pre-conference, these training modules should be compulsory, with the fine arts educator and evaluator completing the training concurrently to ensure mutual learning and alignment throughout the evaluation process.

Recommendation for Practice: Utilize Content Expertise in Evaluations

The final recommendation is for the expertise of content specialists to be leveraged to inform and strengthen the evaluation outcomes of fine arts educators. Participants outlined

various methods used by their schools to heighten the accuracy and fairness of fine arts educators' evaluations. For instance, Evaluator 1 referenced a fine arts-specific rubric used by their district to provide evaluators with items to look for in each fine arts classroom during observations. Similarly, Evaluator 3 and Participant 8 described collaborations with district-level and third-party content experts who helped ensure teacher evaluations were carried out through an arts-minded lens. In general, participants noted evaluators who possessed backgrounds in the fine arts were better equipped to provide relevant and applicable evaluation feedback. It logically follows that a tool intended to help all evaluators gain greater insight into the arts and approach T-TESS through an arts-minded lens would likely be beneficial in improving the long-term effectiveness of the evaluation process. This notion also aligns with the prior research of Gates et a. (2015), and Kimelberg e al. (2019), both of whom called for greater consideration of fine arts educators in the development of teacher evaluation systems. As such, it would be advantageous for fine arts evaluations to utilize either of the aforementioned tactics identified by participants: a supplementary fine arts rubric or consultation with content experts. Though this study did not specifically focus on the implementation or feasibility of either approach, it is anticipated that the design and implementation of a fine arts supplementary rubric would be more practical for most districts. This option does not require that a content expert be available for consultation throughout the evaluation process, and a district- or state-created rubric could be used across multiple schools and districts, thus improving cost-effectiveness. The recommendation for the design of a fine arts rubric to use in conjunction with T-TESS is logical from an ID perspective, as most ID models follow an iterative cycle involving design, implementation, and evaluation. Applying an ID model to the design of a supplementary fine arts-specific rubric would ensure T-TESS remains effective for its intended users—particularly fine arts educators.

Recommendations for Future Research

Four recommendations for future research were proposed in line with the results and implications of this study.

Recommendation 1

Future research should focus on understanding evaluator and fine arts educator perceptions on the effectiveness of other evaluation systems beyond T-TESS. Previous research efforts have focused on understanding administrator perspectives on T-TESS (Krimbill et al., 2019; Paufler et al., 2020) or on understanding how evaluations affect fine arts educators (Fresko & Levy-Feldman, 2023; Lin, 2013; Tutt, 2018). This study was unique in its efforts to encapsulate both the perspectives of fine arts educators and evaluators; yet, this study only focused on one evaluation system. While participants referenced specific aspects of T-TESS in their one-on-one interviews, many of the themes that arose from this research – such as the notion that regular exposure to the fine arts would allow evaluators to provide more informed and applicable evaluative feedback—may apply to other evaluation systems beyond T-TESS. Consequently, it would be valuable to replicate this study, modifying the inclusion criteria to identify participants with at least one year of experience with any given teacher evaluation system, rather than T-TESS specifically. This would allow researchers to determine whether the themes identified in this study were specific to users of T-TESS, or whether the recommendations from this study might be applicable to a wider population of fine arts educators and evaluators.

Recommendation 2

Future research efforts should focus on comparing perceptions regarding T-TESS between two groups of evaluators: those with a teaching background in the fine arts, and those

without such a background. Though it was not a requirement for this participation in this study, all three evaluators had previous experience teaching a fine arts course in the public-school sector. These evaluators felt their prior experience in the fine arts allowed them to provide more accurate and meaningful feedback throughout the evaluation process, and the teachers who participated in this study echoed this perspective. However, it would be valuable to understand how evaluators who do not possess content knowledge of the fine arts perceive the effectiveness and impact of T-TESS. Conducting a comparable study with evaluators lacking experience in the fine arts could help substantiate the current study's emergent findings, while supporting the work of researchers like Tutt (2018), who indicated fine arts teacher evaluations often suffer due to evaluators' limited understanding of what best teaching practices look like in the fine arts. The conversation surrounding T-TESS' effectiveness for usage in evaluating fine arts educators would be propelled forward even if divergent themes arose from this proposed study; if this were to be the case, there would be a need for additional research concerning how one's educational background informs their perception of the evaluation tool.

Recommendation 3

Future studies should focus on currently existing practices that utilize the expertise of fine arts content experts to further the effectiveness of T-TESS, such as the use of best practices rubric and the tactic of consulting with fine arts specialists, with the goal of better understanding how these specific practices affect the evaluation of fine arts educators in terms of perceived evaluation accuracy and quality of evaluative feedback. Though participants recommended these practices as potential pathways to improve the effectiveness of the evaluation process, it would be prudent to further explore the perceived impact of such measures before engaging in implementation on a district level, especially as individual school districts may have unique

needs when it comes to the execution of T-TESS. Potter (2021b) indicated, it is important for fine arts educators' voices to be heard in the evaluation process, and the aforementioned efforts would allow for the voices of content experts to better inform the evaluation of fine arts educators. Yet, it is important to continuously evaluate and develop any new evaluation tool before establishing a finalized version for widespread use (Allen Interactions, 2022), and this third recommendation allows stakeholders to collect the data and feedback necessary to make the most educated decision regarding utilizing the knowledge of content experts to inform fine arts educator evaluations in a given school district.

Recommendation 4

The final recommendation for future research would be to study how the ability to provide meaningful and applicable criticism changes with increased levels of connoisseurship. The framework for this study was based in Eisner's theories of connoisseurship and criticism (Eisner, 2003), and this study's findings support Eisner's premise in that participants indicated increased knowledge of the fine arts (increased connoisseurship) allows evaluators to give more helpful evaluative feedback (improved criticism). Eisner's theories, and by extension, the current study's findings, could be bolstered by additional studies specifically focusing on how developing a greater awareness of and appreciation for the fine arts affects an evaluator's ability to provide feedback that is perceived as meaningful and helpful by fine arts educators. Such a study might involve analyzing the perspectives of fine arts educators and evaluators on the quality of evaluative feedback given to fine arts educators both before and after the evaluators participated in a training focused on fine arts appreciation.

Conclusions

The problem addressed by this study was the ineffectiveness of current fine arts teacher evaluation practices (Gates et al., 2015, Potter, 2021b, Tutt, 2018). The purpose of this action research was to explore perceptions held by evaluators and fine arts teachers regarding the effectiveness of T-TESS in terms of how well the tool achieves validity and alignment with expected performance goals in the evaluation of fine arts educators. This study found that when evaluators lack fine arts content knowledge and when evaluative tools are not designed with the subjective, more creative nature of fine arts classes in mind, the evaluation process is perceived as less meaningful or accurate for fine arts educators; this also hampers opportunities for educator growth. This study's findings may help positively influence the face of fine arts educator evaluations in the future by demonstrating a need to improve how T-TESS is used to evaluate fine arts educators. This study also aimed to continue the work of such researchers as Potter (2021b), who suggested fine arts educators should collaborate with education leaders and reassess how fine arts educator evaluations are designed.

The target population for this study comprised the thousands of fine arts educators and evaluators working for Texas school districts utilizing T-TESS. Inclusion criteria for this study stipulated that participants be fluent in English, have at least one year of experience working as a fine arts educator or evaluator of fine arts educator, have completed the T-TESS evaluation cycle at least once as a fine arts educator or evaluator, and maintained good attendance and professional standing in their position with their school district. A sample of three evaluators and ten fine arts educators were recruited through email and social media.

All participants completed a five-item questionnaire via Qualtrics; the questionnaires asked participants to briefly describe positive and negative interactions with T-TESS. These

general perceptions were explored in detail through individual semi-structured interviews, held on Zoom. After transcribing all interviews with the aid of NVivo, I followed Clarke and Braun's (2013) six steps for thematic analysis to make sense of the data.

Participants in this study stated aspects of T-TESS can be effective for evaluating fine arts educators. While evaluators appreciated that T-TESS allows evaluators to easily provide clear reasoning underlying performance ratings, fine arts educators noted that the T-TESS rubrics can be used as a resource to allow teachers to ensure they are employing best teaching practices. Yet, both evaluators and fine arts educators found T-TESS to be ineffective in the holistic evaluation of fine arts educators. Evaluators specified that T-TESS' language and rating scales are ineffective for use in evaluating fine arts educators. Meanwhile, the fine arts teachers themselves stated T-TESS does not provide an accurate or fair assessment of a fine arts educator, nor does the tool enable evaluators to provide adequate feedback to fine arts educators.

Evaluators and fine arts educators also spoke on the perceived impact of T-TESS. Both participant groups indicated T-TESS has minimal impact: evaluators stated T-TESS has negligible impact on fine arts teacher employment, while fine arts educators noted T-TESS does not influence their choice in teaching strategies. Though T-TESS may have little impact on fine arts educators' teaching practices, this participant group noted that T-TESS can be a source of frustration for fine arts educators. Evaluators conveyed that T-TESS can sometimes have financial ramifications for fine arts teachers, particularly when the evaluation system is tied to a district-wide incentive system. Both evaluators and fine arts educators indicated that when implemented by an evaluator with experience in the fine arts, T-TESS can contribute to fine arts educator growth and accountability.

When asked about additional supports and practices that could be implemented to improve the long-term use and effectiveness of T-TESS, both fine arts educators and their evaluators said that regular fine arts exposure is necessary in allowing evaluators to better understand the work of fine arts educators. Participants also stated proactive communication and advocacy efforts may allow fine arts educators and their evaluators to improve communication and understanding between one another, thus ensuring T-TESS may be used to facilitate growth in fine arts educators. Finally, participants described potential methods of incorporating the work of fine arts "experts" into the T-TESS cycle to allow fine arts educators to receive more meaningful and applicable evaluative feedback. Some suggested examples of using fine arts expertise to enhance T-TESS included utilizing district-level or third-party content experts to assist in evaluations, as well as using supplementary fine arts rubrics to determine T-TESS performance ratings. These findings highlight a unified desire for the addition of arts-specific guidance to the evaluation process, as well as a need to increase connoisseurship in the evaluation process. This notion supports the Eisner's theories of educational connoisseurship and criticism by suggesting that the quality of evaluative feedback given to fine arts educators could be improved if evaluators were to enhance their understanding of and appreciation for the fine arts.

This study's findings indicate a need for evaluators to engage in additional training related to the content areas taught by the teachers they are evaluating, particularly the fine arts. While evaluation tools like T-TESS should be redesigned to encompass the unique characteristics of fine arts instruction, smaller-level practices such as developing and implementing supplementary fine arts-specific rubrics and integrating fine arts expertise into the development of evaluation tools and evaluator training would enhance the ability of the

evaluation process to facilitate the growth of fine arts educators. Failure to explore the aforementioned practice-based implications could potentially prolong inequities in the evaluation of fine arts educators, which would ultimately hamper the growth potential of these educators.

Summary

This section synthesized the study's findings as they pertained to the three research questions. While participants indicated several aspects of T-TESS were effective, such as the evaluation tool's ability to provide clear reasoning underlying performance ratings and to serve as a means to ensure teachers are employing best practices, participants perceived other aspects of T-TESS to be ineffective, such as the ability of the tool to accurately assess fine arts educators. Participants discussed the perceived impact of T-TESS, as well as recommendations for increasing the effectiveness of T-TESS as it is used in the evaluation of fine arts educators. This section also explored implications, recommendations for practice, and recommendations for future research. This study's findings suggest the need for evaluators to strengthen their understanding of and appreciation for the content areas taught by the teachers they are evaluating, particularly the fine arts. The findings support Eisner's theories of educational connoisseurship and criticism, which formed the framework for this study, by highlighting the importance of acquiring content-specific knowledge to improve the quality of criticism one is able to provide. Through the analysis of the perceptions of both fine arts educators and their evaluators, this study contributed to the call for equity in fine arts teacher evaluations and presented action steps to improve the implementation of T-TESS. This study aimed to further the body of research focused on ensuring evaluative systems support fine arts educator growth, while simultaneously providing a foundation upon which future research efforts may draw.

References

Allen Interactions. (2022). *SAM: The successive approximations model*. Allen Interactions. https://www.alleninteractions.com/services/custom-learning/sam/elearning-development

Adelman, C. (1993). Kurt Lewin and the origins of action research. *Educational Action Research, 1*(1), 7–24. https://doi.org/10.1080/0965079930010102

American Institutes for Research. (2016). *Texas teacher evaluation and support system FAQ.* The Texas Comprehensive Center. https://sedl.org/txcc/our_work/tx_educator_evaluation/teachers/teacher_faq.php?

Anderson, C. (2023, August 3). *Fine arts update*. [Live presentation]. Arlington Independent School District.

Bedford, L. (2020). *Generic qualitative research: An often overlooked and misunderstood qualitative approach* [Webinar]. Northcentral University/Center for Teaching and Learning.

Berger, R. (2015). Now I see it, now I don't: Researcher's position and reflexivity in qualitative research. *Qualitative Research, 15*(2), 219–234.

Bloomberg, L. (2017). *Addressing standards of trustworthiness in qualitative research* [Webinar]. Northcentral University/Center for Teaching and Learning.

Bloomberg, L. (Ed.). (2018). "Case study method." *The SAGE encyclopedia of educational research, measurement, and evaluation,* vols. 1–4, SAGE Publications, https://doi.org/10.4135/9781506326139

Bloomberg, L. (2019). *Role & responsibilities of the qualitative researcher* [Webinar]. Northcentral University/Center for Teaching and Learning.

Bloomberg, L. D. (2023). *Completing your qualitative dissertation: A road map from beginning to end* (5th ed.). Sage.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Braun, V., & Clarke, V. (2022). Thematic analysis: A practical guide. *Revia Za Sociologiju.*

Caro-Bruce, C. (2000). *Action research facilitator's handbook*. National Staff Development Council.

Castillo-Montoya, M. (2016). Preparing for interview research: The interview protocol refinement framework. *The Qualitative Report, 21*(5), 811–831. https://doi.org/10.46743/2160-3715/2016.2337

Clarke, V., & Braun, V. (2013). Teaching thematic analysis: Overcoming challenges and developing strategies for effective learning. *The Psychologist, 26*(2), 120–133.

Close, K., & Amrein-Beardsley, A. (2018). Learning from what doesn't work in teacher evaluation. *The Phi Delta Kappan, 100*(1), 15–19.

Cram, W. A., & Marabelli, M. (2018). Have your cake and eat it too? Simultaneously pursuing the knowledge-sharing benefits of agile and traditional development approaches. *Information and Management, 55*(3), 322–339.

Creswell, J. W., & Poth, C. N. (2018). Qualitative inquiry and research design: Choosing among five approaches (4th ed.). Sage.

Daniel, J. (2012). *Choosing the type of nonprobability sampling*. SAGE Publications, https://doi.org/10.4135/9781452272047

Demystifying Research Methodology. (n.d.). *Videos* [YouTube channel]. YouTube. Retrieved June 16, 2023, from https://www.youtube.com/@drphilipadu/videos

Derrington, M. L., & Campbell, J. W. (2018). High-stakes teacher evaluation policy: US principals' perspectives and variations in practice. *Teachers and Teaching Theory and Practice, 24*(3), 246–262. https://doi.org/10.1080/13540602.2017.1421164

Donmoyer, R. (2014). Elliot Eisner's lost legacy in memoriam. *American Journal of Evaluation, 35*(3), 442–452. https://doi.org/10.1177/1098214014537398

Dziak, M. (2022). Focus group. *Salem press encyclopedia*.

Eisner, E. (2003). Educational connoisseurship and educational criticism: An arts-based approach to educational evaluation. In T. Kellaghan, D. Stufflebeam, & L. Wingate (Eds.), *Springer international handbooks of education: International handbook of educational evaluation.* Springer Science Business Media.

Eisner, E. W. (2004). Multiple intelligences: Its tensions and possibilities. *Teachers College Record: The Voice of Scholarship in Education, 106*(1), 31–39.

Eisner, E. W. (2005). *Reimagining schools: The selected works of Elliot W. Eisner*. Routledge.

Eisner, E. W., & Day, M. D. (2004). *Handbook of research and policy in art education*. Routledge.

Elhami, A., & Khoshnevisan, B. (2022). Conducting an interview in qualitative research: The modus operandi. *MEXTESOL Journal, 46*(1).

Fresko, B., & Levy-Feldman, I. (2023). Principals' implementation of teacher evaluation and its relationship to intended purpose, perceived benefits, training, and background variables. *Assessment in Education: Principles, Policy & Practice, 30*(1), 18–32. https://doi.org/10.1080/0969594X.2023.2166461

Garet, M. S., Robyn, A., Stecher, B. M., Baird, M. D., Choi, L., Todd, I., Engberg, J., Holtzman, D. J., Steiner, E. D., Gutierrez, I. A., & Huguet, A. (2020). *Variation in improvement among schools in the intensive partnerships for effective teaching*.

Gates, K., Hansen, D., & Tuttle, L. (2015). Teacher evaluation in the arts disciplines: Three state perspectives. *Arts Education Policy Review, 116*(4), 162–175. https://doi.org/10.1080/10632913.2015.1068657

Gaunt, H., & Treacy, D. S. (2020). Ensemble practices in the arts: A reflective matrix to enhance team work and collaborative learning in higher education. *Arts and Humanities in Higher Education: An International Journal of Theory, Research and Practice, 19*(4), 419–444. https://doi.org/10.1177/1474022219885791

Geiger, T. J., Amrein-Beardsley, A., & Holloway, J. (2020). Using test scores to evaluate and hold school teachers accountable in New Mexico. *Educational Assessment, Evaluation and Accountability: International Journal of Policy, Practice and Research, 32*(2), 187–235. https://doi.org/10.1007/s11092-020-09324-w

Gitomer, D. H., & Marshal, B. L. (2023). The sizzle and fizzle of teacher evaluation in the United States and the selective use of research evidence. *Frontiers in Education, 8*. https://doi.org/10.3389/feduc.2023.1221569

Given, L. M. (Ed.). (2008). *Action research*. SAGE Publications, https://dx.doi.org/10.4135/9781412963909

Greenwood, D. J., & Levin, M. (2007). Introduction: Action research, diversity, and democracy. In Introduction to Action Research (2nd ed., pp. 3–12). SAGE Publications, https://doi.org/10.4135/9781412984614

Groenewald, T. (2004). A phenomenological research design illustrated. *International Journal of Qualitative Methods*, *3*(1), 42–55. https://doi.org/10.1177/160940690400300104

Herr, K., & Anderson, G. L. (2005). Introduction: What is an action research dissertation? In The Action Research Dissertation: A Guide for Students and Faculty (pp. 1–7). SAGE Publications, https://doi.org/10.4135/9781452226644

Holmes, A., & Gary, D. (2020). Researcher positionality—A consideration of its influence and place in qualitative research—A new researcher guide. *Shanlax International Journal of Education, 8*(4), 1–10.

Jacobs, M. A. (2015). By their pupils they'll be taught: Using critical incident questionnaire as feedback. *Journal of Invitational Theory and Practice, 21*, 9–22.

Jewell, J. W. (2017). From inspection, supervision, and observation to value-added evaluation: A brief history of U.S. teacher performance evaluations. *Drake Law Review, 65*(2), 363–420.

Jones, E., Bergin, C., & Murphy, B. (2022). Principals may inflate teacher evaluation scores to achieve important goals. *Educational Assessment, Evaluation and Accountability: International Journal of Policy, Practice, and Research, 34*(1), 57–88. https://doi.org/10.1007/s11092-021-09366-8

Keefer, J. M. (2009). The critical incident questionnaire (CIQ): From research to practice and back again. *Adult Education Research Conference*.

Kemmis, S., McTaggart, R., & Nixon, R. (2014). *The action research planner: Doing critical participatory action research*. Springer.

Kimelberg, S. M., Adelman, R. M., Thompkins, J., & Rabbi, W. (2019). 'It's just art': Experiences of K-12 visual arts teachers in the era of neoliberalism, assessment, and accountability. *Sociation Today, 18*(1), 28–38.

Krimbill, E. M., Goess, D. E., & Escobedo, P. V. (2019). Organizational citizenship and teacher evaluation: Using the T-TESS to promote OCB and improve student outcomes. *School Leadership Review, 14*(1), 12–25.

Lazar, J., Feng, J. H., & Hochhesiser, H. (2017). *Research methods in human computer interaction* (2nd ed.). Morgan Kaufmann.

Li, J., Zhao, J., & Xue, G. (2011). Design of the index system of the college teachers' performance evaluation based on AHP approach. *2011 International Conference on Machine Learning and Cybernetics (ICMLC), 3*, 995–1000. https://doi.org/10.1109/ICMLC.2011.6016905

Lin, M. C. (2013). The development of a performance assessment with performing arts teachers in Taiwan: From national policy to classroom practice. *Ride- The Journal of Applied Theatre and Performance, 18*(3), 296–312. https://doi.org/10.1080/13569783.2013.810926

Matthews, W. K., & Koner, K. (2022). Beyond the headlines: Perspectives of charter school music teachers. *Arts Education Policy Review, 123*(2), 39–49. https://doi.org/10.1080/10632913.2020.1766615

Mertler, C. (2019). *The Wiley handbook of action research in education*. John Wiley & Sons.

Naeem, M., Ozuem, W., Howell, K., & Ranfagni, S. (2023). A step-by-step process of thematic analysis to develop a conceptual model in qualitative research. *International Journal of Qualitative Methods, 22*. https://doi.org/10.1177/16094069231205789

National Council on Teacher Quality. (2020). *State of the states 2019: Teacher & principal evaluation policy*. National Council on Teacher Quality. https://www.nctq.org/pages/State-of-the-States-2019:-Teacher-and-Principal-Evaluation-Policy#:~:text=As%20of%202019%2C%20more%20than,adhere%20to%20a%20binary%20system.

Nordin, A., & Wahlström, N. (2019). *Transnational policy discourses on 'teacher quality': An educational connoisseurship and criticism approach.*

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). https://doi.org/10.1177/1609406917733847

Paufler, N. A. (2018). Declining morale, diminishing autonomy, and decreasing value: Principal reflections on a high-stakes teacher evaluation system. *International Journal of Education Policy & Leadership, 13*(8), 1–15. https://doi.org/10.22230/ijepl.2018v13n8a813

Paufler, N. A., King, K. M., & Zhu, P. (2020). Delivering on the promise of support for growth? Evaluator perceptions of a new state teacher evaluation system. *Journal of Educational Supervision, 3*(2), 32–50. https://doi.org/10.31045/jes.3.2.3

Potter, D. (2021a). (In)visible power: A critical policy analysis of edTPA. *Arts Education Policy Review, 122*(2), 101–114. https://doi.org/10.1080/10632913.2020.1744053

Potter, D. (2021b). Music teacher perceptions of evaluation: A case study of the Tennessee fine arts portfolio model. *Arts Education Policy Review*, 1–13. https://doi.org/10.1080/10632913.2021.1992324

Reid, D. B. (2018). How schools train principals to use new teacher evaluation systems: A comparative case study of charter school and traditional public school principals. *Journal of School Choice, 12*(2), 237–253. https://doi.org/10.1080/15582159.2018.1447737

Roberts, R. E. (2020). Qualitative interview questions: Guidance for novice researchers. *The Qualitative Report, 25*(9), 3185–3203. https://doi.org/10.46743/2160-3715/2020.4640

Rowe, W. E. (2014). Positionality. In D. Coghlan & M. Brydon-Miller (Eds.), *The SAGE encyclopedia of action research*. SAGE.

Samuel, A., & Conceição, S. C. O. (2022). Using the critical incident questionnaire as a formative evaluation tool to inform online course design: A qualitative study. *International Review of Research in Open and Distributed Learning, 23*(2), 151–169. https://doi.org/10.19173/irrodl.v23i2.5959

Shaw, R. D. (2016). Arts teacher evaluation: How did we get here? *Arts Education Policy Review, 117*(1), 1–12. https://doi.org/10.1080/10632913.2014.992083

Shenton, A. K. (2004). Strategies for ensuring trustworthiness in qualitative research projects. *Education for Information, 22*(2), 63–75. https://doi.org/10.3233/EFI-2004-22201

Stake, R. E. (1993). *The art of case study research*. Sage.

Steinberg, M. P., & Kraft, M. A. (2017). The sensitivity of teacher performance ratings to the design of teacher evaluation systems. *Educational Researcher, 46*(7), 378–396.

Stronge, J. H., & Tucker, P. D. (2023). *Handbook on teacher evaluation with CD-ROM: Assessing and improving performance*. Eye on Education.

TEA. (2015). *Texas teacher evaluation and support system FAQ.* Texas Education Agency. https://www.nctq.org/dmsView/Texas_Teacher_Evaluation_and_Support_System_FAQ_June_5_74815

TEA. (2022a). *Fine Arts Texas Essential Knowledge and Skills*. Texas Education Agency. https://tea.texas.gov/academics/curriculum-standards/teks-review/fine-arts-texas-essential-knowledge-and-skills

TEA. (2022b). *Implementation guidebook.* Texas Education Agency. https://teachfortexas.org/resource_files/Guides/T-TESS_Implementation_Guidebook.pdf

TEA. (2022c). *Texas evaluation and support system*. Texas Education Agency. https://tea.texas.gov/texas-educators/educator-evaluation-and-support-system

TEA. (2022d). *Texas teacher evaluation and support system*. Texas Education Agency. https://tea.texas.gov/texas-educators/educator-evaluation-and-support-system/texas-teacher-evaluation-and-support-system

TEA. (2022e). *T-TESS rubric*. Texas Education Agency. https://teachfortexas.org/Resource_Files/Guides/T-TESS_Rubric.pdf

TEA. (2024a). *Finding a school*. Texas Education Agency. https://tea.texas.gov/texas-schools/general-information/finding-a-school-for-your-child/finding-a-school#:~:text=Finding%20a%20school%20among%20Texas,9%2C000%20campuses%2C%20can%20be%20daunting.

TEA. (2024b). *TESS pilot districts*. Texas Education Agency. https://tea.texas.gov/texas-educators/educator-evaluation-and-support-system/tess-pilot-districts

Tessmer, M., & Wedman, J. (1990). A layers-of-necessity instructional development model. *Educational Technology Research & Development, 38*(2), 77–85. https://doi.org/10.1007/BF02298271

Tutt, K. (2018). Issues in performing arts teacher preparation program evaluation policy. *Arts Education Policy Review, 119*(4), 185–193.

Uhrmacher, P. B., Moroye, C. M., & Flinders, D. J. (2017). *Using educational criticism and connoisseurship for qualitative research*. Routledge.

Van Manen, M. (2016). *Phenomenology of practice: Meaning-giving methods in phenomenological research and writing*. Routledge.

Welman, J. C., & Kruger, S. J. (2001). *Research methodology for the business and administrative sciences* (2nd ed.). Oxford University Press.

Whitehead, J. (2008). Using a living theory methodology in improving practice and generating educational knowledge in living theories. *Educational Journal of Living Theories, 1*(1), 103–126.

Willey, G. P. (2019). The perceptions of teacher evaluation by teachers and campus administrators in a suburban Texas district. *School Leadership Review, 14*(1). 4–11.

Willgens, A. M., Cooper, R., Jadotte, D., Lilyea, B., Langtiw, C. L., & Obenchain-Leeson, S. (2016). How to enhance qualitative research appraisal: Development of the methodological congruence instrument. *The Qualitative Report, 21*(12), 2380.

Yang, S. (2023). Multi-dimensional evaluation of teachers' leading role in art-training courses. *International Journal of Emerging Technologies in Learning, 18*(15), 209–224. https://doi.org/10.3991/ijet.v18i15.42381

Zerull, D. S. (1990). Evaluation in arts education: Building and using an effective assessment strategy. *Design for Arts in Education, 92*(1), 19.

Appendix A

Questionnaire

Please take about five minutes to respond to the following questions about your experience with T-TESS. If nothing comes to mind for any of the below questions, simply type N/A. You will be assigned a unique participant identifier so as to keep your responses anonymous. What you write below will help inform our discussion during your Zoom interview, so please be sure to remain honest and as detailed as possible in your responses.

1. What aspect of T-TESS do you find most positive/helpful in navigating your role during the observation process?

2. What aspect T-TESS do you find most negative/unhelpful in navigating your role during the observation process?

3. What about the T-TESS itself or its usage in the evaluation of fine arts educators do you find most affirming?

4. What about T-TESS itself or its usage in the evaluation of fine arts educators do you find most confusing?

5. What supports or practices, if any, would be beneficial in improving the way T-TESS is used in the evaluation of fine arts educators? This could be about your reactions to the evaluation process, ideas to help make T-TESS more applicable to fine arts educators, or anything else that comes to mind.

Appendix B

Semi-Structured Interview Guide

Introduction: Hello, and thank you for agreeing to participate in this interview today. My name is Brianna Miller, and I am a doctoral candidate at National University conducting my dissertation research. This interview is expected to last 45-60 minutes. I will be recording our discussion and taking notes to make sure I have complete information. Your responses will be confidential and will not be shared with anyone. Additionally, I will send you a copy of the interview transcript within the next three days so that you may have the chance to review and/or clarify your responses.

Consent: I would like to review the consent letter with you before we begin the interview. *Do you agree to participate in the study?*

Lead into Interview: I'm asking you to take part in an interview about how fine arts educators and evaluators perceive the effectiveness and impact of T-TESS on the evaluation of fine arts educators. Your perspectives and experiences are important in understanding the impact and usage of the tool in its current form.

Demographic Questions

1. Gender: Male, Female, Prefer Not to Answer
2. Age
3. Which school district do you work for?
4. What is your current role, or which role have you previously held? Is this the only role you have held?
5. How long have you held this role?

Study Questions

Fine Arts Educators

1. Describe your interaction, if any, with T-TESS, including any T-TESS-related documents or rubrics, prior to your performance observation. This may include during your pre-observation conference.
2. Describe your interaction, if any, with T-TESS, including any T-TESS-related documents or rubrics, during your observation.

3. Describe your interaction, if any, with T-TESS, including any T-TESS-related documents or rubrics, during your post-observation conference.
4. Please explain any additional specific experiences that have shaped your perception of T-TESS and your role as a fine arts educator in the T-TESS observation cycle.
5. What do you understand to be the purpose of T-TESS?
6. Explain any differences, if you feel any exist, in the way that T-TESS impacts fine arts educators opposed to educators of other content areas.
7. What role, if any, does T-TESS play in your choice of teaching strategies, both when being observed and in your daily teaching (when not being observed)?
8. In what ways, if any, does the T-TESS system provide a true and fair evaluation of fine arts educators?
9. Is there anything about T-TESS' design or the way it is used that would need to be changed to help make the T-TESS system more effective when it comes to the evaluation of fine arts educators?
10. Do you have any other comments you would like to share about T-TESS and the evaluation of fine arts educators?

Additional questions, such as those listed below, may be added based on responses to the questionnaire:

- Am I correct in understanding that you feel/stated...?
- Could you explain that a bit more?
- Why do you feel this way?
- How did that make you feel?
- How did you come to that conclusion?
- Did anything else stand out to you?
- Give me additional background on that.
- Is there anything else you think I should know about that?

Evaluators of Fine Arts Educators

1. Describe your interaction, if any, with T-TESS, including any T-TESS-related documents or rubrics, before engaging in the observation process. This may include during any trainings you may have participated in or during your pre-observation conference with a teacher.
2. Describe your interaction, if any, with T-TESS, including any T-TESS-related documents or rubrics, during your observation.
3. Describe your interaction, if any, with T-TESS, including any T-TESS-related documents or rubrics, during your post-observation conference.
4. Please explain any additional specific experiences that have shaped your perception of T-TESS and your role as an evaluator in the T-TESS observation cycle.
5. What do you understand to be the purpose of T-TESS?
6. Explain any differences, if you feel any exist, in the way that you navigate the T-TESS system when it comes to fine arts educators, as opposed to educators of other content areas.
7. What role, if any, does T-TESS play in your choice of evaluation strategies and your approach to the evaluation process, specifically when it comes to fine arts educators?
8. In what ways, if any, does the T-TESS system provide a true and fair evaluation of fine arts educators?
9. Is there anything about T-TESS' design or the way it is used that would need to be changed to help make the T-TESS system more effective when it comes to the evaluation of fine arts educators?
10. Do you have any other comments you would like to share about T-TESS and the evaluation of fine arts educators?

Additional questions, such as those listed below, may be added based on responses to the questionnaire:

- Am I correct in understanding that you feel/stated...?
- Could you explain that a bit more?
- Why do you feel this way?
- How did that make you feel?
- How did you come to that conclusion?
- Did anything else stand out to you?
- Give me additional background on that.
- Is there anything else you think I should know about that?

Closing: Once again, thank you for your participation in this study. As a reminder, I will email you a copy of the interview transcript within three days, and you will have three days to confirm that the transcript accurately captures your viewpoints.

Appendix C

Note-Taking Matrix for Semi-Structured Interviews

RQ1: How do public-school evaluators of fine arts educators perceive the usage and impact of a fine arts educator performance evaluation tool?	
Interview Questions (Educators Only)	**Notes**
Describe your interaction, if any, with T-TESS, including any T-TESS-related documents or rubrics, prior to your performance observation. This may include during your pre-observation conference.	
Describe your interaction, if any, with T-TESS, including any T-TESS-related documents or rubrics, during your observation.	
Describe your interaction, if any, with T-TESS, including any T-TESS-related documents or rubrics, during your post-observation conference.	
Please explain any additional specific experiences that have shaped your perception of T-TESS and your role as a fine arts educator in the T-TESS observation cycle.	
What do you understand to be the purpose of T-TESS?	
Explain any differences, if you feel any exist, in the way that T-TESS impacts fine arts educators opposed to educators of other content areas.	
What role, if any, does T-TESS play in your choice of teaching strategies, both when being observed and in your daily teaching (when not being observed)?	
Additional questions prompted by responses to the questionnaire	
RQ2: How do fine arts educators perceive the usage and impact of a fine arts educator performance evaluation tool?	
Interview Questions (Evaluators Only)	**Notes**
Describe your interaction, if any, with T-TESS, including any T-TESS-related documents or rubrics, before engaging in the observation process. This may include during any trainings you may have participated in or	

during your pre-observation conference with a teacher.	
Describe your interaction, if any, with T-TESS, including any T-TESS-related documents or rubrics, during your observation.	
Describe your interaction, if any, with T-TESS, including any T-TESS-related documents or rubrics, during your post-observation conference.	
Please explain any additional specific experiences that have shaped your perception of T-TESS and your role as an evaluator in the T-TESS observation cycle.	
What do you understand to be the purpose of T-TESS?	
Explain any differences, if you feel any exist, in the way that you navigate the T-TESS system when it comes to fine arts educators, as opposed to educators of other content areas.	
What role, if any, does T-TESS play in your choice of evaluation strategies and your approach to the evaluation process, specifically when it comes to fine arts educators	
Additional questions prompted by responses to the questionnaire	
RQ3: Reflecting on their experiences with a fine arts educator performance evaluation tool, what needs do evaluators and fine arts educators identify related to the effective and prolonged use of this tool in the future?	
Interview Questions	**Notes**
In what ways, if any, does the T-TESS system provide a true and fair evaluation of fine arts educators?	
Is there anything about T-TESS' design or the way it is used that would need to be changed to help make the T-TESS system more effective when it comes to the evaluation of fine arts educators?	
Do you have any other comments you would like to share about the fine arts educator performance evaluation tool?	
Possible Follow Up Questions (to be asked throughout interview as deemed appropriate)	

- Am I correct in understanding that you feel/stated...?
- Could you explain that a bit more?
- Why do you feel this way?
- How did that make you feel?
- How did you come to that conclusion?
- Did anything else stand out to you?
- Give me additional background on that.
- Is there anything else you think I should know about that?

Appendix D

Consent Form

My name is Brianna Miller, and I am a doctoral student at National University (NU). I also hold a role as the Modern/Contemporary Dance Director at the Dr. Marcelo Cavazos Center for Visual and Performing Arts with Arlington Independent School District.

I'm asking you to take part in a research study about how fine arts educators and evaluators perceive the effectiveness and impact of T-TESS on the evaluation of fine arts educators. The name of this research is "Evaluating the Effectiveness of the Texas Teacher Evaluation and Support System (T-TESS) for Fine Arts Educators."

You may participate in this research if you meet all of the following criteria:

1. You are age 18 or older.
2. You are fluent in English.
3. You work(ed) for a Texas school district as a fine arts (band, choir, dance, orchestra and strings, piano, theatre arts, or visual arts) educator or an evaluator of fine arts educators.
4. You have completed the T-TESS system at least once as a fine arts educator or evaluator of fine arts educators.
5. You maintain(ed) good attendance and professional standing with your school district.

I would like to recruit 12 volunteer participants for this study.
Please read this form carefully and ask any questions you may have before agreeing to take part in the study.

What you will be asked to do: If you agree to be in this study, you will be asked to do the following activities:

1. Complete an online five-item questionnaire on Qualtrics for 5-15 minutes.
2. Participate in a 1:1 online interview over Zoom for 45-60 minutes.
3. Review your interview transcript via email for 10-15 minutes.

During these activities, you will be asked questions about:

- Your age, gender, and work experience,
- Your experience with T-TESS.
- Your thoughts on the purpose, impact, and effectiveness of T-TESS as a means of evaluating fine arts educators

Risks: There are minimal foreseeable risks or discomforts associated with this research. You can still skip any question you do not wish to answer, skip any activity, or stop participation at any time.
Benefits: If you participate, there are no direct benefits to you. This research may increase the body of knowledge in the subject area of this research.
Recording: I would like to audio/video record your responses with Zoom during the interview. You can disable the video function of the online meeting platform at any time.
Mandated Reporting: My professional role outside of NU requires me to report suspicion of child or elderly abuse, suspicion of possible harm to self or others, and committed crimes to the appropriate authorities.
Confidentiality: I will keep the records of this study private and take reasonable measures to protect the security of all your personal information. In any report I make public, I will not include any information that will make it possible to identify you. A code system will be used to identify all participant responses. No actual names will be used; each participant will be assigned a letter and number, and these participant letter-number codes will be used in all reports. All documents pertaining to this study will be stored on the researcher's password protected computer, while all paper documents will be kept in a locked cabinet in the researcher's home. All documents connected with this study will be destroyed after three years following the completion of the thesis.

Taking part is voluntary: Participation in this study is completely voluntary. You may quit at any time.

If you have questions: Please ask any questions you have now. If you have questions later, you may contact me at b.miller0444@o365.ncu.edu or at 845-863-7457.
If you have any questions or concerns regarding your rights as a subject in this study, you may contact the Institutional Review Board (IRB) via email at irb@nu.edu

Appendix E

Personal Journal Maintained throughout Study

5.5.2025

- E1 indicated frequent attendance in classrooms allows them to have a better idea of best practices for each subject
- Best Practices Tool (used by district): E1 was largely in charge of creation of supplementary rubric; uses BP to double check for subjects less sure on
- Evaluations are not meant to be a "gotcha"
- The way E1 speaks on benefits of T-TESS very much gives the feeling of one who has gone through training as an admin/who has had the merits of evaluation drilled into their brain- curious if all evaluators will share similar sentiments/have faith in T-TESS
- lots of good info on merits of fine arts classes, but does it fit the RQs?
    - Arts cater to whole student; students come to school for arts and the connections formed in the class; content shapes student in a way core classes do not (ex. confidence)
- Technology component of T-TESS not helpful- check for thread among others
- Assume good intent

5.12.2025

- lots of follow-ups on questionnaire to schedule interview, but end of school year is making scheduling interviews difficult- was this a bad time of year to plan for data collection? Hopefully the merits of time (end of eval cycle) make up for difficulty in scheduling

5.27.2025

- some participants discussed skipping T-TESS (same with potential participant who responded to cold email); difficult to provide commentary on T-TESS when you haven't gone through it recently, but still valuable to know T-TESS isn't valuable enough not to skip
- Classroom management much different in visual arts- kids can socialize/listen to music/relax (same in middle school), very different; always good evaluations though; kids know teacher being evaluated
    - Do evaluations need to be a surprise? Is there a need to make a different rubric to evaluate arts?
- different views on lesson between fine arts teachers and evaluators- need for evaluators to use arts-minded lens
- Teachers able to express similar "textbook" response on purpose of T-TESS, but acknowledge T-TESS doesn't necessarily meet those goals
- T2 indicated other teachers already are inclined to "look down" on fine arts, so changes to T-TESS/accommodations could breed jealousy
    - Disconnect between FA and other areas- potential theme for investigation?
- Are fine arts people more self-motivated than core content teachers, or is this just participants being biased/self-inflated?
    - Explore in future interviews

6.3.2025

- T4 echoed many themes/topics of other interviews (disconnect between FA and other content areas; evaluators not understanding arts; purpose of T-TESS not really holding true in reality for arts)

- The evaluator makes a difference- so far, there's been a lot of discussion on looking at the whole picture and praise for evaluators w/ FA backgrounds, but T4 able to provide insight into having a "regular" evaluator and a fine arts evaluator; described a world of difference between experiences with T-TESS

6.4.2025

- Piece seem to be coming together; E2 helps solidify notion that fine arts lens is helpful in providing meaningful evaluative feedback.
- Notion that contention historically exists between fine arts educators and evaluators based in history of arts in public education. This could be true, but could also be the focus of a future study. For now, may this is a recommendation for future research (proactively building relationships between the parties?).
- Last few interviews have not shown support for prior theory about FA teachers being more motivated/passionate (I assumed this theory might arise in multiple interviews, but apparently not)
- Interestingly, E2 noted issues in T-TESS when they were a FA teacher, but implies their approach/arts lens helps make T-TESS less unfair than in the past. Important to note this could be a moment of self-bias… maybe not in the sense that an arts lens matters, but more so in the sense that T-TESS is automatically more equitable if one has an arts background

6.11.2025

- Importance of arts advocacy has been a recurring theme in interviews
- FA teachers need to explain to evaluators how arts are different and class flow so evaluators know what to expect to improve evaluations

- I have SO much data/anecdotes/explanations regarding how the fine arts are different than core classes, but unfortunately, a lot of the great quotes I got on this topic came from participants going on slightly off-topic and the info does not directly relate to my RQs. I was thinking about how I would structure a section on what differentiates the arts from core classes, but I need to recognize that is not related to T-TESS/the purpose of my study.
- I wonder if many of the repeated topics from the various interviews (need for more arts exposure, need for repaired relationships between evaluators and FA educators) would apply to evaluation tools other than T-TESS?
    - The choice to focus on T-TESS was intentional to ensure participants were speaking on a singular evaluation tool, but it would be interesting to see if perspectives differ with different tools

6.14.2025

- T7 had experience at both elementary and high school levels and began interview saying that they had very different experiences with the evaluation process between the grade levels, but as the interview progressed, participant acknowledged the biggest difference was admin support/previous experience in FA (one had great support, the other not so much)
- Redundancy of T-TESS mentioned- other teachers indicated frustration with various elements (not specifically redundant elements)… can these frustrating elements be combined into a single theme?
- Will have to revisit transcripts and see what coding indicates, but seems like there is a link between connoisseurship and criticism in that fine arts teachers prefer to receive

feedback from evaluators with a FA background because the feedback truly helps them grow

  - This is my experience as well, so I must be aware of biases as I navigate data analysis

6.16.2025

- It is interesting how similar themes pop up between interviews, but in different ways (ex. T8 mentioned serving as a third party to help with FA evaluations, while E3 described using district-level experts to confirm evaluation outcomes). These sort of anecdotes/recommendations are similar to the use of the FA rubric mentioned by E1 and others, so this may evolve into a theme related to RQ3.
- Again, the need for advocacy and educating evaluators on FA has come up. Why is this the teacher's responsibility? Wouldn't it be great if this sort of education/knowledge was built into T-TESS or evaluator training?
- Are these problems and desire for more arts-expertise specific to T-TESS, or applicable to all evaluation tools?
- I assumed the participants would have issues with aspects of T-TESS, but I am surprised by how many participants say they trust their own background/knowledge in teaching or evaluating.

Ongoing Thoughts During Data Analysis

- I am nervous I will analyze the data "incorrectly" and my biases will unintentionally affect my interpretation of the data
- As I review the transcripts, I am just creating new codes as I go to try to let the data speak for itself; also writing down quotes that seem important for later use

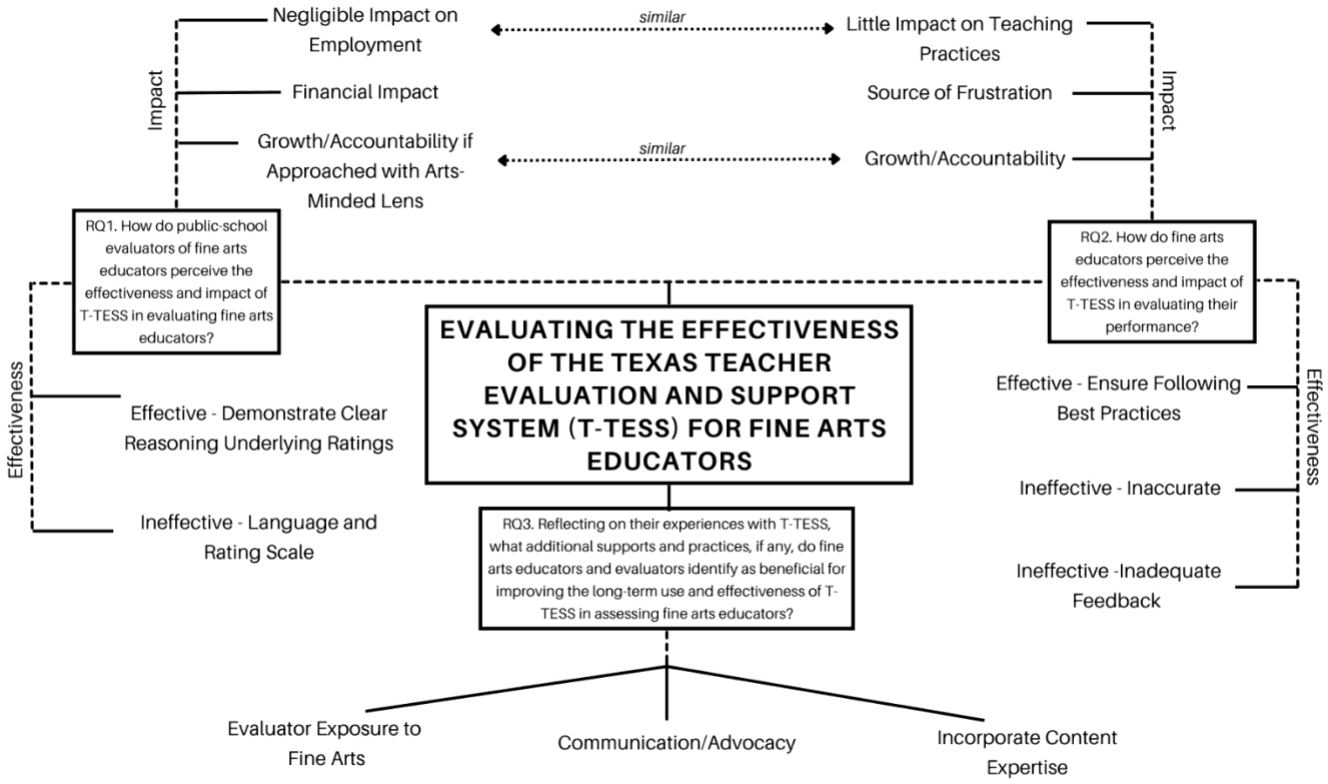
- First round of coding has left me with many codes, some of which will likely need to be combined or otherwise modified as I go
- I have created a completely different set of codes based on the research questions and the knowledge of what my first round of coding seemed to reveal thematically (ex. lots of focus on preference for evaluators with arts lens, as well as frustration that T-TESS is not always applicable to FA teachers)
    - Once this list of descriptors was complete, I went back to my descriptors created through the first round of inductive coding to add in any descriptors I had missed
- I then engaged in another round of coding to put all data into the new, still-evolving code list, now organized by research question. I found myself adding and combining descriptors as I went
- Descriptors under RQ1 and RQ2 are similar in some ways- will want to highlight similarities/connections in my visual map
- After sitting on my coding schema and coming back a few days later, I am satisfied with my schema, but I also recognize someone else might have been able to construct a slightly different schema from the same dataset… BUT I am fairly confident that even if the descriptor names were different, the overall category groupings/themes would be similar
- Determination of themes/narratives is a bit tricky- I can look at the coding schema and construct a basic map of themes, but as I begin to draft my narrative and pull specific participant quotes, I find the wording of my themes needs to change slightly. Hopefully this is a good sign that I am allowing the data to truly "speak for itself" instead of trying to impose my own themes and phrasing on the dataset

  - Do I have too many themes? I don't think so, but only time (and SME feedback) will tell
  - Why does so much of this analysis process feel like I am an imposter? I am able to supply a plethora of quotes to support the themes, but AM I DOING IT RIGHT? Or rather, is there any "correct" answer in research? How do you know your findings are of the highest caliber?

- I feel like I have mentally been constructing a visual map of my thoughts as I engaged in coding, so sketching out a visual map is turning out easier than I thought it would be (although I wonder if my map should be simpler—it makes sense to me though).

Appendix H

Visual Map of Emergent Themes



Note. Image created with the aid of Canva