

<u>`USING ACHIEVEMENT GOAL THEORY TO INVESTICATE PRE-SERVICE</u> <u>MUSIC TEACHERS' ATTITUDES TOWARD GRADING PRACTICES</u>

by

Nathan	A. St. Pierre
A D	ussertation
Suon Gradu	nued to the
Grade	of
George M	01 Jason University
in Partial	I Fulfillment of
The Requiren	nents for the Degree
The requirement	of
Doctor	of Philosophy
Musi	c Education
Committee:	
	Director
	Program Director
	Director of the School of Music
	Dean, College of Visual and
	Performing Arts
Date:	Fall Semester 2017
	George Mason University
	Fairfax, VA

Using Achievement Goal Theory to Investigate Pre-service Music Teachers' Attitudes toward Grading Practices

A Dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at George Mason University

by

Nathan A. St. Pierre Master of Education Loyola University Maryland, 2011 Bachelor of Music The Hartt School, University of Hartford, 2007

Director: Charles Ciorba, Associate Professor Department of Music Education

> Fall Semester 2017 George Mason University Fairfax, VA

COPYRIGHT 2017 NATHAN A. ST. PIERRE ALL RIGHTS RESERVED

Dedication

This dissertation is dedicated to Mixolydian, my favorite mode.

Acknowledgements

I would like to thank my wife, Katie, for entertaining my whims and pretending to be interested in this "boring stuff." Also, my parents, Valerie and Gerald Paquin, for their support. I want to thank Dr. Brian Wuttke, for his leadership and for teaching me to love research. I also thank the members of my cohort, for being my sounding block, devil's advocates, and friends. I would like to extend sincere thanks to Drs. Edward Bolkovac and Stuart Younse. Without whose artistry, I never would have fallen in love with choral music. I also thank Dr. John Feierabend, who was a mentor to me when I was a preservice teacher. He recognized my potential, and gave me my favorite comeback: "show me the research." I cannot give enough thanks to Dr. Lisa Billingham, who (thanks to a very serendipitous meeting) has become a trusted friend and mentor. Without her, I would not be writing this document. I also thank Philip Desrosiers, my high school music director, for setting me on this amazing path. I would like to acknowledge the (past and present) administration, faculty, and staff of Ronald Reagan Middle School, where it is my honor to grow young adults with music. I also thank Mrs. Betty Wade, who always inspires and reminds me that everything music teachers do now is because they stand on the shoulders of those who came before. I thank the brothers of The Lodge of the Nine Muses No.1776 (F.A.A.M), all of whom are smarter than me and always challenge me to be better. I thank Phi Mu Alpha Sinfonia Fraternity of America, for growing my desire to advance in America the best in music. I also thank Susan Dommer, a mentor teacher, who is the only person more excited than me to see this document completed. I need to thank Dr. Angela Miller, who first introduced me to Achievement Goal Theory. I also thank Dr. Frederick Brigham, who taught me to aim for Pluto with a rocket from the Moon. Lastly, I need to thank Dr. Charles Ciorba, for jumping right in and giving me the help I needed when I needed it. To him I am, and always will be, very grateful.

Table of Contents

	Page
List of Tables	viii
List of Figures	ix
List of Abbreviations	X
Abstract	xi
Chapter One: Introduction	1
Background	2
Need for the Study	4
Purpose of the Study	4
Chapter Two: Literature Review	7
Social Cognitive Theories of Motivation	
Achievement Goal Theory	
The dichotomous model	11
The trichotomous model	
A 2 x 2 framework	
Toward a Mastery Orientation	15
The Stability of Achievement Goals	17
Teacher Influence on Achievement Goal Orientations	
Teacher Attitudes Toward Grading	
General education	
Music education	
The Case for Standards-based Grading	
The Effect of Teacher Training	
Summary	
Chapter Three: Methodology	
Research Design	
Variables	

Participants	35 35 36 38
Participants.	35 36 38
	36 38
Sampling	38
Measures	
Achievement goal orientations.	38
Attitude Toward Grading Practices.	39
Anticipated use of SBG.	40
Procedure	40
Summary	42
Chapter Four: Data and Analysis	44
Research Question One: Descriptive Statistics	45
Item Reverse Scoring	45
Descriptive Statistics.	49
Research Question Two: Correlation Analysis	50
Variable Correlations	50
GPAS Factor Structure.	51
AGQ Factor Structure	57
Research Question Three: Cluster Analysis	59
Clustering Procedure.	59
Naming the Clusters.	62
Research Question Four: The Effect of AGO Profile on Other Variables	63
AGO Profile and Attitudes toward Grading Practices	63
AGO Profile and Anticipated Use of SBG.	64
Summary	65
Chapter Five: Discussion and Recommendations	68
Discussion	69
Implications	72
Limitations	74
Recommendations for Future Research	76
Conclusion	77
Appendix A: An Example of a Standards-Based Grading Instrument	79

Appendix B: Achievement Goal Questionnaire	80
Appendix C: Grading Practice Attitude Survey	81
Appendix D: Consent Form	83
Appendix E: Recruitment Letter for Participants	
Appendix F: IRB Decision Letter	85
References	

List of Tables

Table	Page
Table 1	
Table 2	
Table 3	
Table 4	
Table 5	
Table 6	
Table 7	
Table 8	
Table 9	
Table 10	
Table 11	
Table 12	
Table 13	
Table 14	
Table 15	

List of Figures

Figure	Page
Figure 1. Dendrogram using Ward's linkage method	60

List of Abbreviations

Standards-Based Grading	SBG
Achievement Goal Orientation	AGO
Achievement Goal Questionnaire	AGQ
Grading Practice Attitude Survey	GPAS
Cluster Membership	AGO Profile
Performance-Approach AGO	Perf_Ap_Total
Performance-Avoidance AGO	Perf_Av_Total
Mastery-Approach AGO	Mast_Ap_Total
Mastery-Avoidance AGO	
Attitude toward Grading Practices	GPAS_Total
Anticipated Use of SBG	Use_SBG

Abstract

USING ACHIEVEMENT GOAL THEORY TO INVESTIGATE PRE-SERVICE MUSIC TEACHERS' ATTITUDES TOWARD GRADING PRACTICES

Nathan A. St. Pierre, Ph.D.

George Mason University, 2017

Dissertation Director: Dr. Charles Ciorba

The purpose of this study was to examine the interplay between achievement goal orientation (AGO) and attitudes toward grading practices. A secondary purpose was to identify homogeneous groups of participants for further analysis. Pre-service music teachers in their third year of undergraduate study (N = 56) were given measures of their achievement goal orientation (AGQ) and attitudes toward grading practices (GPAS). A cluster analysis homogeneously grouped participants into five AGO Profiles: (a) Non-Mastery-Oriented; (b) Non-Performance-Oriented; (c) Performance-Oriented; (d) Eclectic; and (e) Mastery-Oriented. These AGO Profiles differed significantly with regards to the average GPAS score for each group (GPAS_Total) [$F(4, 51) = 2.94, p = .03, \eta^2 = .19$]. Specifically, the Mastery-Oriented AGO Profile scored significantly ($p \le .03$) higher (M = 162.20, SD = 10.56) than the (a) Non-Mastery-Oriented AGO Profile (M = 145.90, SD = .03, SD = .03, SD = .03, SD = .04, SD = .056) that the (a) Performance-Oriented AGO Profile (M = .045.90, SD = .03, SD = .03, SD = .03, SD = .04, SD = .04, SD = .04, SD = .056) that the (a) Performance-Oriented AGO Profile (M = .045.90, SD = .056) that the (a) Performance-Oriented AGO Profile (M = .045.90, SD = .03, SD = .04, SD = .056) that the (a) Performance-Oriented AGO Profile (M = .045.90, SD = .056) that the (a) Performance-Oriented AGO Profile (M = .045.90, SD = .056) that the performance-Oriented AGO Profile (M = .045.90, SD = .056) that the performance-Oriented AGO Profile (M = .045.90, SD = .056) that the performance-Oriented AGO Profile (M = .045.90, SD = .056) that the performance-Oriented AGO Profile (M = .045.90, SD = .056) that the performance-Oriented AGO Profile (M = .045.90, SD = .056) that the performance-Oriented AGO Profile (M = .045.90, SD = .056) that the performance-Oriented AGO Profile (M = .045.90, SD = .056) the performance-Oriented AGO Profile (M = .056.90, SD = .056) the performance-Oriented AGO Profile (M = .056.90, SD

10.72). These results suggest that neither AGQ scores representing high mastery-based orientations nor AGQ scores representing low performance-based orientations by themselves are enough to affect attitudes toward grading practices (GPAS_Total). Rather, it is the combination of both high mastery-based orientations with low performance-based orientations that leads to increased acceptance of grading practices typical of standards-based grading (SBG). In addition, AGO Profile, as a categorical variable, was not found to be significantly related to anticipated use of standards-based grading (SBG) (p = .51).

Chapter One: Introduction

Student assessment is one of the most important responsibilities that any teacher will undertake. While grades are primarily intended to communicate academic performance to students, parents, and administrators, they are often used for a wide range of purposes (e.g., rewards, punishments, eligibility for extra-curricular activities, class placement, special projects, and college admission) (Brookhart, 1991). Knowing that grades will be used in a variety of ways, it might be even more important for teachers to ensure that student grades are reflective of their actual achievement with regards to the curriculum. Within the field of music education, grades should reflect musical achievement, performance ability, and content knowledge.

It is possible that certain psychological antecedents might exist as obstacles for music teachers when evaluating student achievement based on musical criteria. More specifically, teachers might "teach as they learn" (Malmberg, 2006, p. 59) and therefore teachers' academic goal orientations and attitudes toward grading, could potentially serve as variables influencing their grading policy decisions. This study will examine these psychological antecedents and determine to what extent they affect pre-service music teachers' willingness to accept an achievement-only grading paradigm viz., standardsbased grading (SBG).

Background

Assessment experts recommend that grades should reflect how well a student has mastered what has been taught (Cross & Frary, 1999; Frary, Cross, & Weber, 1993). And yet, for decades music teachers have often based their grades on non-musical criteria (e.g., attendance, effort, etc.) (St. Pierre & Wuttke, 2015), despite the fact that music education experts have recommended student assessment should be based on curriculum standards (Colwell, 1998; Consortium of National Arts Education Associations, 1994; Lehman, 1998). Furthermore, it has been suggested that SBG can be used as a means to bring the recommendations of assessment experts into the classroom (St. Pierre & Wuttke, 2015). Part of the problem lies in the fact that according to recent studies (St. Pierre & Wuttke, 2015; Wuttke & St. Pierre, 2016), 17% to 25% of current and future music teachers were still hesitant to adopt SBG policies. The theoretical basis for the current study implies that psychological antecedents may affect the manner in which music teachers grade their students. Specifically, this study will look at achievement motivation and attitudes toward grading.

There are many social-cognitive theories of motivation within educational psychology, although achievement goal theory holds particular interest for its applications to teacher grading practices. This motivational theory describes the goals individuals set for specific tasks and demonstrates different "ways of approaching, engaging in, and responding to achievement-type activities" (Ames, 1992, p. 261). At the most basic level, Dweck (1986) identified the following achievement goal orientations (AGOs): "(a) learning goals, in which individuals seek to increase their competence, to

understand or master something new, and (b) performance goals, in which individuals seek to gain favorable judgments of their competence or avoid negative judgments of their competence" (p. 1040). There are many other labels for these AGOs in the literature including *task-involvement* and *ego-involvement*, *task* and *ability*, *mastery* and *ability*, or *mastery* and *performance* goals (Ames, 1992; Butler, 2007; Midgley et al., 1998). Among these labels, mastery and performance are the most common (Buluş, 2011). From there, it was determined that an avoidance of the appearance of incompetence was equal to showing competence as an AGO. As a result, a trichotomous model of achievement goal theory was developed, which identified (a) *mastery*, (b) *performance-approach*, and (c) *performance-avoidance* goal orientations (Elliot & Harackiewicz, 1996).

The trichotomous model served as the basis for much of what is known about achievement goal theory, but the model was not as balanced as it could have been. To address this problem, Elliot and McGregor (2001) identified two terms for mastery goals and two terms for performance goals with competence as a focal point. This new model, referred to as the 2 x 2 framework, proposed a fourth term, *mastery-avoidance*, which described individuals who wished to avoid completing a task in an incorrect manner.

Teachers have seemed to prefer students who are mastery-oriented (Schraw & Aplin, 1998). However, the AGOs of teachers can affect their instruction and influence student motivation (Throndsen & Turmo, 2013). Students will sometimes change their AGOs in response to perceived goal preferences in their classroom climates (Ames, 1992; Ames & Archer, 1988; Sandene, 1997). One way in which students perceive teacher goal preference is through teacher assessment (Ames, 1992; Malmberg, 2008;

Pedersen & Williams, 2004). Therefore, teachers might want mastery-oriented students, but their grading policies create performance-oriented classroom environments. The current educational trend of high-stakes testing has also created classroom situations that are heavily performance-oriented (Butler, 2007).

Teacher attitude toward grading is another potential antecedent. St. Pierre and Wuttke (2015) found that 17% of music teachers who participated in their study were not using standards-based grading because they were happy with their current system and saw no reason to change. Even after training, 25% (n = 7) of pre-service music teachers reported they would not use standards-based grading if they were not required to do so (Wuttke & St. Pierre, 2016). Of those, 86% (n = 6) exhibited GPAS scores lower than the mean.

Need for the Study

While it is known that teacher AGOs affect the manner in which they teach, and attitude toward grading affects the way they assess, it is not known to what extent these variables interact to influence teacher grading policy decisions. This gap in the knowledge generally contributed to the need for the current study, which used psychometrics to measure these variables among a sample of pre-service music teachers and compared them to anticipated grading policy.

Purpose of the Study

The primary purpose of this study was to examine the interplay between achievement goal orientation (AGO) and attitudes toward grading practices. A secondary purpose was to identify homogeneous groups of participants for further analysis. This was accomplished through a cluster analysis using scores on the Achievement Goal Questionnaire (AGQ). These groups were then compared to the other variables to determine the effect of AGO on pre-service music teachers' attitudes toward grading practices as well as their anticipated use of SBG. To accomplish this investigation, the following research questions were considered:

- 1. What are the descriptive statistical characteristics of the examined variables?
- 2. To what extent do AGQ scores correlate with the other variables?
- 3. What clusters emerge related to scores on the AGQ subscales?
- 4. What is the effect of cluster membership (AGO Profile) on the other variables?
 - a. On attitude toward grading practices?
 - b. On anticipated use of SBG?

For the final research question, two null hypotheses were considered:

- 1. AGO profile would not be associated with attitude toward grading practices (GPAS) scores with an alternate hypothesis that the GPAS mean scores using AGO Profile as a predictor would differ significantly and be a better fit than the null model: H_0 : $\mu_1 = \mu_2 = ... = \mu_k$
- The probability of anticipating use of SBG would be the same regardless of AGO Profile.

An examination of these research questions is important to the field of music education in that it provides methods course professors an understanding of some of the psychological antecedents that affect whether teachers will use standards-based grading practices. This new-found understanding may help faculty to address these factors and bring the grading practices of music educators more in line with the recommendations of assessment experts.

The target population for this study was third-year pre-service music teachers enrolled in state-approved teacher preparation programs at institutions of higher education in the United States. The participants were obtained using a combination of probability and non-probability sampling techniques. The specific nature of this research methodology will be discussed in greater detail in Chapter Three.

This chapter has explained the research design, participants, and research questions for the current study, all of which will be further examined in Chapter Three. More importantly, this chapter outlined a background of the current problem regarding motivation and music teacher grading policies. This chapter has provided a rationale and theoretical basis for the current study and identified the variables to be considered. Each of the variables examined here will be explained more thoroughly in Chapter Two.

Chapter Two: Literature Review

The primary purpose of this study was to examine the interplay between achievement goal orientation (AGO) and attitudes toward grading practices. A secondary purpose was to identify homogeneous groups of participants for further analysis. This study further investigated the effect of these groups on pre-service music teachers' attitudes toward grading practices as well as their anticipated use of SBG. To accomplish this investigation, the following research questions were considered:

- 1. What are the descriptive statistical characteristics of the examined variables?
- 2. To what extent do AGQ scores correlate with the other variables?
- 3. What clusters emerge related to scores on the AGQ subscales?
- 4. What is the effect of AGO Profile on the other variables?
 - a. On attitude toward grading practices?
 - b. On anticipated use of SBG?

Research in the areas of (a) student motivation for learning and (b) teacher attitudes toward grading practices have been ongoing for many decades. A review of the literature is presented here to provide background on the following major components of the current research: (a) important theories of learning motivation, with an emphasis on achievement goal theory; (b) research regarding teacher—both pre-service and inservice—attitude toward grading practices; and (c) the presence and importance of standards-based grading (SBG) in music education.

Social Cognitive Theories of Motivation

Motivation is a commonly investigated variable in the field of education. Researchers and teachers have long sought to understand the myriad psychological drives behind the classroom behaviors of students and teachers alike. Some of the more common theories of motivation include (a) self-efficacy theory, which deals with an individual's beliefs regarding his or her capabilities in various situations; (b) Expectancy X Value theory, which examines an individual's hope for success as it relates to his or her fear of failure; (c) attribution theory, which describes how an individual will attribute his or her success or failure; (d) intrinsic motivation theory, which describes how individuals satisfy their own needs in addition to social cognitive constructs; and (e) achievement goal theory, which describes the effects social and cognitive goals have on both academic and social behaviors (Maehr, Pintrich, & Linnenbrink, 2002).

Intrinsic motivation theory, attribution theory, and achievement goal theory are of particular interest to music education researchers (West, 2013). Notably, studies that have addressed student attributions for their successes and failures have received a good deal of attention in music education literature (Asmus, 1986; Asmus, 1994; Marlatt, 2004; Sandene, 1997). This is not surprising as 20 percent of achievement can be attributed to motivation (Asmus, 1994). Therefore, it is logical that research in the area of attribution theory can provide teachers with valuable information that can be used to affect student motivation and maximize that portion of achievement.

These thoughts are limited in that attributions are only telling part of the story. Elliott and Dweck (1988) used a pattern recognition task to examine the effect of selfperception on level-of-skill among a group of fifth-grade students (N = 101). Students were split into four experimental groups, and were then asked to complete a specific task. The task instructions provided to participants emphasized either (a) performance or (b) learning. Students were then given feedback (i.e., high or low) based on their ability for the task. When learning was valued (i.e., increased competence), neither the high ability nor low ability students were likely to make attributions for their failure. In the group where performance (i.e., looking competent) was valued (n = 50), students who were told they had low ability (n = 23) made attributions for that failure. Only four percent of the students who were told they had high ability made such attributions. Interestingly, every child in the low ability manipulation made an attribution to an uncontrollable cause, and none of these students attributed their failure to effort (Elliott & Dweck, 1988).

Additional research by Dweck and Leggett (1988) found that children who displayed helplessness often attributed their failure to personal inadequacy, where mastery-oriented children saw failure as a learning opportunity. It was also suggested that attributions, which are internal-stable (e.g., ability or talent), can lead to "helplessness deficits" and "deteriorated performance" (Brunson & Matthews, 1981, p. 917; Elliott & Dweck, 1988). The culmination of this research indicated there are two underlying theories of intelligence. As described in Dweck and Leggett (1988),

Some children favor what we have termed an *incremental theory* of intelligence: They believe that intelligence is a malleable, increasable, controllable quality. Others lean more toward an *entity* theory of intelligence: They believe that intelligence is a fixed or uncontrollable trait (p. 262).

These theories can be extended beyond intelligence to describe personalities. It has been suggested that "viewing [a personal attribute] as a fixed trait will lead to a desire to document the adequacy of that trait, whereas viewing it as a malleable quality will foster a desire to develop that quality" (Dweck & Leggett, 1988, p. 266). This evidence suggests that attributions act as a defensive mechanism for those who value their appearance of competence, as opposed to those who actually value their competence. This is further evidenced by the results reported by Elliot and Dweck (1988), who reported that only 12% (n = 6) of learning goal oriented students made such attributions.

Achievement Goal Theory

Given the situational differences in attributions, it is imperative to isolate a psychological construct that addresses one's learning orientation and subsequent attribution. Elliott and Dweck (1988) proposed the term *goal* as the "construct that organizes these previously distinct cognitive and affective factors and helps us to understand the conditions under which they arise" (p. 6). The earliest conceptions of these *achievement goals* described two goal orientations: (a) *performance goals*, where the objective is to gain positive judgments and/or avoid negative judgments regarding one's ability or competence; and (b) *learning goals*, where the objective is to increase one's competence (Dweck, 1986; Dweck & Leggett, 1988). These goals have been directly related to the incremental and entity theories of intelligence, and these intelligence theories can orient individuals toward different goals. For example, those

who have an (a) incremental theory of intelligence orient themselves toward learning goals, and (b) those with an entity theory of intelligence orient themselves toward performance goals (Dweck, 1986; Dweck & Leggett, 1988). Dweck and Leggett (1988) further explained this trend by stating:

for any personal attribute that the individual values, viewing it as a fixed trait will lead to a desire to document the adequacy of that trait, whereas viewing it as a malleable quality will foster a desire to develop that quality (p. 266).

Elliott and Dweck (1988) tested this theory using a sample of fifth-grade students (N = 101). Results indicated that learning and performance achievement goals can stem from sources of influence on achievement behavior patterns such as task selection ($\chi^2(1, N = 101) = 22.35, p < .001$) and problem-solving strategies (F(1, 94) = 4.08, p < .05).

The dichotomous model. Elliott and Dweck's (1988) research validated a theory in which two major achievement goals can serve as the motivational source for behavior patterns. This dichotomous model of goal orientations became the basis for much of the research in achievement goal theory that was to follow. Their presence in the literature was steadfast, but there was disagreement among researchers regarding how the goal orientations should be labeled. Some used the original labels of *learning* and *performance*; others referred to the goals as (a) *task-involvement* and *ego-involvement*, (b) *task* and *ability*, (c) *mastery* and *ability*, or (d) *mastery* and *performance* goals (Ames, 1992; Butler, 2007; Midgley et al., 1998). Among these labels, mastery and performance are the most common (Bulus, 2011). Armed with a new dichotomous model of achievement goal orientations, motivational researchers began to explore the implications for classroom performance. Ames and Archer (1988) conducted a study using a sample of 8-11 grade students (N =176) to examine perceptions of the goals espoused in their classroom climates. They found that perceived mastery orientation was a significant predictor of learning strategies, and mastery/performance goals provided a useful means of separating student sensitivities to the learning environment.

The trichotomous model. Members of the research community believed the dichotomous model of achievement goal orientation did not fully explain motivation. As such, they sought to incorporate an additional orientation based on earlier motivational research. Specifically, (a) mastery orientation was seen as an *approach* goal, in that the objective was to attain competence/mastery, and (b) performance orientation was an approach goal, in that the objective was to attain competence/mastery, and (b) performance orientation was an approach goal, in that the objective was to attain favorable judgment. It was postulated, however, that performance orientation could also be an *avoidance* goal if the objective was to avoid an unfavorable judgment rather than obtain a judgment that was favorable (Elliot & Harackiewicz, 1996). Elliot & Harackiewicz (1996) tested this theory in a series of experiments using undergraduate psychology students. They concluded that approach-avoidance orientations might affect motivational processes regardless of an individual's perception of competence.

Once the experimental basis for this theory was in place, additional research was conducted to further validate the new trichotomous model of achievement goal orientation. Elliot and Church (1997) determined that splitting performance goal orientation into approach-avoidance partitions accounted for 63.3% of the variance in the questionnaire. This provided strong support for the trichotomous model for achievement goal orientations. These results were similar to those attained by Middleton and Midgley (1997) and Midgley et al., (1998). Perhaps the most well-known measure of the trichotomous model of achievement goal is the Patterns of Adaptive Learning Scales (PALS) (Midgley et al., 2000), which was developed as part of the longitudinal Patterns of Adaptive Learning Study (Midgley, 2002). This trichotomous model has been popular among researchers and has served as the basis for much of what is known regarding achievement goal theory (Elliot & McGregor, 1999; Harackiewicz, Barron, Carter, Lehto, & Elliot, 1997; Midgley et al., 1998).

A 2 x 2 framework. More recently, researchers began to see the trichotomous model of achievement goal theory as slightly unbalanced. To remedy this, it was proposed there should be two terms for approach goals and two terms for avoidance goals using competence as the focal point (Elliot & McGregor, 2001). To test this theory, Elliot and McGregor (2001) conducted a series of three studies using undergraduate psychology students who were administered a number of instruments designed to measure the newly-proposed *mastery-avoidance* goal orientation. In this new construct, mastery-avoidance exists when an individual defines mastery internally but is more concerned about avoiding failure than attaining competence. A factor analysis of the AGQ found four factors and the solution accounted for 81.5% of the total variance (Elliot & McGregor, 2001). When compared to the trichotomous models (Elliot & Church, 1997; Middleton &

13

Midgley, 1997), which accounted for the less than 65% of the total variance, this new 2 x 2 framework does appear to be more balanced.

Other researchers further examined the validity of this new framework in an attempt to examine the extendibility of these goals across various domains. Finney, Pieper, and Barron (2004) administered a modified version of the AGQ to a sample (N = 2,111) of freshman students at a mid-sized east coast university. The instrument was modified to reflect academic achievement in general, rather than focus on a specific class. This large-sample analysis of the AGQ confirmed support for four distinct factors, and also provided evidence for the achievement goal orientations as domain-specific rather than course-specific (Finney et al., 2004).

To further explore the domain-specificity of achievement goal orientations, Baranik, Barron, and Finney (2007) studied a sample (N = 307) of undergraduate psychology students at a midsized southeastern university. Their data collection included items related to both academic and work goals. It was hypothesized that an eight-factor model would provide the best fit for the data, therefore providing evidence for achievement goals as domain-specific orientations rather than global personality traits. Results indicated the eight-factor model was the best fit for the data, and the hypothesis was confirmed.

Researchers in music education have also begun to use the 2 x 2 framework. Miksza (2009) adapted the AGQ and administered it to a sample of high school band students (N = 228). Two two-factor models, two three-factor models, and the 2 x 2 framework were tested using a confirmatory factor analysis to determine which model was the best fit for the data. It was determined the four-factor solution was the best fit for the data. To date, the most balanced framework for achievement goal theory is one that incorporates a total of four achievement goal orientations: (a) mastery-approach, (b) performance-approach, (c) mastery-avoidance, and (d) performance-avoidance. It is also best when these orientations are conceived with situational dependency, as opposed to measuring the aspects of one's personality (Baranik et al., 2007).

Toward a Mastery Orientation

While it has been shown that achievement goal orientations can affect various behavior patterns (Elliott & Dweck, 1988), perhaps it is not too far a leap to imagine there are positive and negative aspects to each goal. The research literature has identified the benefits of adopting a mastery goal orientation including:

- Attributions to effort rather than ability (Ames, 1992; Ames & Archer, 1988; Buluş, 2011; Droe, 2013; Dweck, 1986; Dweck & Leggett, 1988; Elliott & Dweck, 1988)
- Employment of deep learning strategies, which leads to long-term retention of material (Elliot & McGregor, 1999, 2001; Graham & Golan, 1991; Midgley, Kaplan, & Middleton, 2001)
- Increased intrinsic motivation (Elliot & Church, 1997; Elliot & Harackiewicz, 1996; Harackiewicz et al., 1997; Malmberg, 2006)
- The embracement of challenges as learning opportunities (Ames & Archer, 1988; Buluş, 2011; Dweck & Leggett, 1988; Elliot & McGregor, 2001)

- Willingness to seek help (Butler, 2007; Butler & Shibaz, 2008; Middleton & Midgley, 1997)
- Fewer instances of task/test anxiety (Dweck & Leggett, 1988; Eison, 1979; Elliot & McGregor, 1999; Middleton & Midgley, 1997)

It should also be noted that teachers may have a preference for students who are mastery-oriented, at least in the context of teacher training programs, in that they "believe that high-mastery students are harder working, have better in-class social skills, will make better teachers, and are more likely to succeed than low-mastery students are" (Schraw & Aplin, 1998, p. 219).

There are benefits to a performance-approach orientation as well. Performanceapproach goals have been linked to higher standardized test scores (Elliot & McGregor, 2001), and higher classroom grades (Elliot & Church, 1997). Performance-avoidance goals have been shown to lead to lower class grades (Elliot & Harackiewicz, 1996; Elliot & McGregor, 1999; Harackiewicz et al., 1997), decreased intrinsic motivation (Elliot & Church, 1997; Elliot & Harackiewicz, 1996; Harackiewicz et al., 1997; Malmberg, 2006), and low competency expectations (Dweck & Leggett, 1988; Elliot & Church, 1997; Harackiewicz et al., 1997). Essentially, those who believe they have low competence will endorse performance-avoidance goals to avoid the display of incompetence, but they will not exert any effort as they believe it would be seen as evidence of their inability to attain the performance goals (Dweck & Leggett, 1988). This is an important realization as there is a feeling among those who are performance-oriented that effort is indicative of low ability, even if such efforts produce successful outcomes (Dweck & Leggett, 1988). When students perceive their classroom climates support or emphasize achievement goals, the students will alter their learning strategies accordingly (Ames & Archer, 1988).

The Stability of Achievement Goals

While school-age students seem to be able to alter their achievement goals based on perceptions of their classroom environment, it is not entirely clear if they are changing their own goal orientations or simply adapting to their environment. To test the stability of achievement goal orientations, Elliot and McGregor (2001) administered their AGQ to a sample of undergraduates (N = 182) during their exam periods. This allowed the researchers to determine values for subsequent achievement goals. Regression analyses determined that there was "strong evidence for the stability of each learning goal" (Elliot & McGregor, 2001, p. 510). However, it was discovered that mastery-avoidance goals were a positive predictor of subsequent (a) mastery-approach goals and (b) subsequent performance-approach goals (Elliot & McGregor, 2001). These results indicate that, under certain classroom circumstances, achievement goals can be modified to more adaptively suit the student.

It has also been shown that among teachers, goal orientation for student learning is a "relatively stable construct" (Anderson, Anderson, Mehrens, & Prawat, 1990, p. 334). Anderson, Anderson, Mehrens, and Prawat (1990) conducted a longitudinal mixedmethods study using 74 third and fourth-grade teachers. Participants participated in interviews and were asked to complete a Teacher Priorities Questionnaire. It was found that even when teachers did change their goals for student learning, most eventually returned to their original goal orientations (Anderson et al., 1990). It should be noted that while student learning goal orientation is a relatively stable construct, "the fact that some teachers can and do change should serve as an inducement to continue this line of research, and certainly is cause for optimism in the enterprise of improving teacher effectiveness" (Anderson et al., 1990, p. 335).

The stability of achievement goal orientations has also been tested among preservice teachers. Malmberg (2008) conducted a longitudinal study of Finnish students who participated in a teacher-training program. It was discovered that mastery goal orientation was the most stable of the goal orientations, but all three exhibited oscillating patterns. Additionally, goal orientations that were unrelated at the beginning of the study became correlated in the third and fourth years while others, which were relatively strongly related at the beginning, became very strongly related later in the study (Malmberg, 2008). This provides evidence that individuals may be able to change goals in different social or academic contexts (Baranik et al., 2007; Finney et al., 2004). This is an important implication in that pre-service music teachers might be able to change their goals in order to better develop a mastery orientation.

Teacher Influence on Achievement Goal Orientations

Given that certain evidence indicated that teachers preferred students who are mastery-oriented (Schraw & Aplin, 1998), teachers may be able to influence the achievement goal orientations of their students. Some of the early studies of achievement goal theory indicated that students not only adopted the goal orientations valued by their teachers in classroom climates, they modified their learning accordingly (Ames, 1992; Ames & Archer, 1988). This trend has also been seen in the field of music education where teachers who valued ego goals (i.e., performance goals) also had students who valued ego goals (Sandene, 1997). Throndsen and Turmo (2013) provided evidence that "teachers' goal orientations affect their instructional practices, and that teachers are prone to utilize teaching strategies that reflect their personal goal orientation" (p. 317). In addition, Butler (2007) concluded that "teachers' achievement goals might have consequences also for their students" (p. 251). Perhaps, teacher-training programs should determine which goal orientations pre-service teachers possess as these teachers may unknowingly expose their students to their own achievement goals when they enter the profession. According to Butler and Shibaz (2008), this can cause maladaptive learning strategies, that could lead to cheating.

Butler (2007) stated that, "studying teacher goals is important to the extent that these are associated with distinct patterns of meaning and action" and that "analyses showed coherent relations between teachers' endorsement of one or another goal orientation and their help-related perceptions, preferences, and behavior" (p. 249). In addition, when teachers are evaluated based on student test scores they may very well create classroom cultures that value performance-approach goals (Butler, 2007). As such, it is recommended that teachers orient themselves to the achievement goal that not only invites the most adaptive learning strategies but also produces long-term retention of the material being taught (i.e., mastery-approach goals). This could be done through feedback, which values incremental (e.g., effort) rather than entity (e.g., ability) theories of musical performance (Droe, 2013).

19

Another possible way teachers can affect student achievement goal orientations could be through the use of assessment and grading. Ames (1992) argued that "the ways in which students are evaluated is one of the most salient classroom factors that can affect student motivation" (p. 264). When grades are accompanied by improvement opportunities, students might be more likely to self-evaluate based on effort (Ames, 1992). This is important as it has been proposed that students might invest greater effort in graded tasks than ungraded ones (Pedersen & Williams, 2004). While students might change effort for graded tasks, they might also reorient toward different achievement goals, Previous research has found when students are aware they are not graded, there is (a) an increase in mastery goal orientations and (b) a decrease in performance-avoidance goal orientations (Malmberg, 2008). Perhaps, it is not the absence of graded tasks that curbs student focus on performance-oriented goals, but rather teacher-provided opportunities for student improvement that simultaneously focuses on student effort (Ames, 1992). In other words, teachers can foster mastery-oriented goals by celebrating effort and individual improvement.

Teacher Attitudes Toward Grading

General education. There are many different attitudes, understandings, and recommendations regarding the meaning of grades. Measurement experts agree that grades should exclusively represent academic achievement (Brookhart, 1991; Cross & Frary, 1999; Frary et al., 1993). This might be easily accomplished if grades were simply vehicles for interpretation, but they are also used for a variety of factors including rewards, punishments, eligibility for extra-curricular activities, class placement, special projects, and college admission (Brookhart, 1991). Brookhart (1991) identified that "teachers understand, and wish for, a way to provide clear, interpretable feedback to parents about their children's academic learning. But they live in a world where, for both good and ill, grades are used for lots of different things" (p. 36).

Because of this, grades have unintended social consequences, which can contribute to student self-perceptions, even leading to school drop-out (Brookhart, 1991). Teachers have attempted to mitigate these effects by adding non-achievement criteria to grade calculation; creating a "hodgepodge grade of attitude, effort, and achievement" (Brookhart, 1991, p. 36; Cross & Frary, 1999). It has been suggested that teachers are more concerned with how grades make students feel than how grades are used to judge them (Carlton, 1992). Carlton (1992) documented that if teachers were required to grade using achievement-only measures, many would simply ignore the requirement. In fact, it was found that:

If forced to grade in this manner, teachers stated they would fudge grades in the grade book to compensate for not including effort as part of a grade. For example, they stated they would erase a lower test score in the grade book and record a higher one (Carlton, 1992, p. 175).

This shows a deep-seeded and documented trend that teachers grade their students contrary to the recommendations of measurement experts (Brookhart, 1991; Carlton, 1992; Cross & Frary, 1999; Frary et al., 1993). This can partially be explained through a belief in which teachers are of the opinion that effort should be rewarded (Carlton, 1992). Furthermore, many see grades as a classroom management tool to curb undesirable behaviors (Bonner & Chen, 2009; Brookhart, 1994; Frary et al., 1993). There are even some educational contexts in which teachers remain unconvinced that grading is beneficial at all (Richards, 2014). Frary et al. (1993) surveyed 536 secondary teachers in the Virginia public schools, and could not identify a "typical responder" (p. 24). The researchers had to employ a factor analysis of the opinion scores to determine patterns and cluster analysis to identify groups of responders.

This inconsistency toward attitude and grading practice is also present among preservice teachers. Bonner and Chen (2009) studied 222 teacher candidates and determined that "about one-third of candidates generally supported basing grades solely on alternative forms of assessment, and about one-third generally supported relating grades to class or student behavior through the use of pop quizzes or score reductions" (p. 70). There is also conflicting evidence regarding the role assessment training can play on the grading practices of pre-service teachers. Bonner and Chen (2009) found that, after taking formal coursework in assessment strategies, pre-service teachers moved away from reliance on (a) alternative assessments, (b) academic enabling, and (c) the use of grades for classroom management. Even so, participants were unable to abandon academic enabling completely. These results are somewhat in conflict with the findings of Campbell and Evans (2000), who critiqued the lesson plans of 65 pre-service teachers. It was discovered that "only 13 out of 113 assessments requiring a rubric did contain one; of those 13, only 8 (approximately 7%) were complete" (p. 353). They also found the majority of assessments did not (a) report scoring methods or (b) provide complete information. This seems to suggest that even in the wake of formal coursework, preservice teachers were still unclear about grading and assessment. This echoes Bonner and Chen (2009), who concluded "candidates nearing the end of a preparation program are still very naïve about appropriate and ethical methods for evaluating their students" (p. 73).

Music education. Music teachers are not immune to the confusions regarding the meaning and interpretation of grades. Colwell (1998) concluded, "interviews with arts teachers about their use of assessment elicit the two extremes; they perceive that they either evaluate continually or not at all" (p. 30). Some music teachers reported the subjective nature of musical performance makes it difficult, maybe even inappropriate to assess (Asmus, 1999). As a result, music teacher grading policies are often made up of a hodgepodge of musical and non-musical criteria (Kotora, 2005; McCoy, 1991; Russell & Austin, 2010). Researchers have found that music teachers expressed a lack of clarity and frustration with regards to assessment (Kotora, 2005). This lack of clarity is illustrated in the research of Russell and Austin (2010), who reported a lack of consensus with regards to the calculation of music grades.

Colwell (1998) stated, "in arts education the argument is made that process is more important than the product. It is the joy of creating, of doing, of participating that is valued" (p. 30). This value on participation seems to have created a metaphorical monster regarding the use of non-musical criteria in grading policies. Results from a study by Russell and Austin (2010) identified that 91% of secondary music teachers were using attendance as a grading criterion with an average weight of 25%. It was also discovered that 93% of secondary music teachers were using attitude as a grading criterion with an
average weight of 27%. In total, 60% of grade weight was attributed to non-musical criteria. Kotora (2005) also found a heavy reliance on attendance and attitude. Among high school choral directors in Ohio, the three most common grading criteria were concert performances, student participation, and student attendance (Kotora, 2005). In a study that involved 28 pre-service music teachers, Wuttke and St. Pierre (2016) reported that 100% of participants had been graded using the hodgepodge of musical and non-musical criteria in their high school music ensembles. When asked to create their own hypothetical grading policies, it was reported that attendance and participation accounted for three out of the four most cited grading criteria (Wuttke & St. Pierre, 2016).

Even more problematic is the manner in which these attendance and attitude grades seemed to be calculated. According to research by Russell and Austin (2010), 67% of teachers who were using attendance as part of their grade calculation reported to have reduced final grades by one letter or more as a result of unexcused absences, and 30% reduced grades partially. While these results are not necessarily generalizable to the entire population, it does imply that a majority of music teachers who include attendance in their grade calculation may be doing so to punitively curb absenteeism from major performance events or rehearsals. While grading policies that punish absenteeism or reward attendance could curb absenteeism (Marburger, 2006), using grades as penalties for chronic absenteeism has not been shown to improve either attendance or academic performance (Moore, 2005). The attitudinal criteria are equally problematic as Russell and Austin (2010) discovered that as little as 10% of participants who used attitudinal measures as part of their grading procedures based those calculations on purely objective measures; 90% used either completely subjective measures or a combination of subjective and objective measures.

It is tempting to use the lack of training in proper assessment strategies as a scapegoat for this unfortunate trend in school assessment. For example, Colwell (1998) conducted a phone survey of arts teachers and discovered that "courses in testing and measurement were never required. Fewer than half of the colleges of education have a requirement in assessment, and in those that do, arts students are always exempt" (p. 33). Additional studies have also shown the lack of appropriate assessment training to be a contributing factor (Kotora, 2005; St. Pierre & Wuttke, 2015). However, Brookhart (1994) suggested "more training, by itself, will not cause grading practices to conform completely to recommendations" (p. 290). This has been confirmed by several studies (Bonner & Chen, 2009; Campbell & Evans, 2000; Russell & Austin, 2010), which indicated teachers continued to grade students contrary to expert recommendations even after they have participated in appropriate assessment training.

The Case for Standards-based Grading

To break the cycle of improper assessment and grading, it might be necessary to move away from traditional grading practices in favor of a new grading system. Toward the beginning of the standards-based movement in arts education, Brophy (1997, p. 25) posed that "given the multiple elements that make up music education, how can the progress of a student be represented by a single letter grade?" In general education, Brookhart (1994) also sounded the call for grading reform stating, "It would be helpful to explore the usefulness of a double grading system, marking achievement and effort separately and finding ways to take both seriously" (p. 297).

Standards-based Grading (SBG) could be the solution that satisfies the needs for grading reform expressed by those assessment experts. Asmus (1999) defined standardsbased assessment as, "assessment established from school, district, state, or national standards of content and performance in a subject" (p. 21). In addition, St. Pierre and Wuttke (2015) provided a definition based on a review of the literature that stated:

[SBG] is a way to provide students and parents with growth-producing feedback about classroom achievement in a reliable and valid way. More specifically, SBG allows for the evaluation of students based on how close their classroom performance is to meeting curriculum standards or specific learning targets... Additionally, SBG allows students to retake assignments until they demonstrate proficiency, uses current learning trends rather than an average of points to determine grades and employs rubrics as tools to compare student proficiency levels to curricular standards (p. 3; cf. Hanover Research, 2011).

Within the discipline of music education, the push for SBG seems to be a recent trend, but its roots in this field can be traced to the beginning of the standards movement itself. When the standards were published, it was stated that the standards were intended to form the "foundation for student assessment" (Consortium of National Arts Education Associations, 1994, p. 15). Colwell (1998) also championed the use of standards in assessment by saying, "the value of the standards is that they bring clarity to intentions and require attention to assessment" (p. 31). Other researchers have also demonstrated

how basing grades on the standards more closely aligns with the recommended practice of assessment experts stating:

if the purpose of grades is to document the degree to which students are learning what they are expected to learn, it follows logically that grades should report to what extent students are meeting the standards—national, state, or local—that represent the goals of the school district, (Lehman, 1998, p. 37).

As more school districts began to adopt standards-based curricula, researchers attempted to keep the focus on assessment. For example, Lehman (1998) believed "if the curriculum is standards-based, then student grading should be standards-based as well" (p. 37). As some recent research has found, Lehman's recommendation seems to have been ignored. For example, Russell and Austin (2010) surveyed music teachers and found that even when districts adopted standards-based curricula, secondary music teacher assessment practices remained unchanged. Those researchers also discovered that only 2% of school districts had adopted SBG models. More recently, St. Pierre and Wuttke (2015) surveyed 96 music teachers and found that 83% of teachers worked in districts with standards-based curricula, while approximately 40% of respondents were using SBG.

These findings are troubling as they demonstrate the adoption of standards-based curricula seems to have little effect on the grading practices of music teachers. However, as a call to action, "in an effort to emphasize achievement-based assessment and deemphasize the use of attendance and/or attitude to determine student grades in music,

27

standards-based curricula should be considered a 'point of departure' in formulating assessment strategies" (Russell & Austin, 2010, pp. 50–51).

This is an important consideration as it is possible the reason SBG has not caught on in music education is due to a lack of exposure. St. Pierre and Wuttke (2015) discovered that the most cited answers music teachers gave for not using SBG were a (a) lack of knowledge toward SBG and (b) lack of training in achievement-based assessment. However, it was also discovered that "almost all of the music teachers who had heard of SBG were using it to assess and grade their students" (St. Pierre & Wuttke, 2015, p. 7).

The Effect of Teacher Training

One possible way to maximize the prevalence of SBG in music classrooms might be to address it as a part of teacher training programs. This is consistent with the recommendations of other experts who have hypothesized that "the present inadequacy of most required arts offerings is due not to a lack of teacher competence but to a lack of emphasis on the use of assessment in preservice education" (Colwell, 1998, p. 29). Other researchers have shown that assessment training in teacher preparation can curb less than ideal grading practices (Bonner & Chen, 2009). Professional development, however, seems to have had little effect on teacher assessment strategies (Brookhart, 1994; Russell & Austin, 2010). Perhaps during the coursework of a teacher training program, professors "have the opportunity to break the cycle of 'teaching how you were taught'" (Ester, 1997, p. 26). McCoy deduced this of ensemble directors by stating, "if however, college and university directors adopt grading systems that evaluate student performance rather than attendance, those systems will probably 'trickle down' to the high school level" (McCoy, 1991, p. 189). It is important that methods professors stay current with trends in education as they may feel they are providing appropriate preparation, but their students may graduate feeling ill-prepared (Kotora, 2005).

To test the effect teacher preparation classes may have on pre-service teacher attitudes toward grading practices, Wuttke and St. Pierre (2016) measured the attitudes of 28 pre-service music teachers before and after training in SBG. In the weeks following training, pre-service teacher attitudes toward SBG increased significantly. As a further test, the researchers graded the participants using an SBG model for the entire semester; it was found that at the end of the course pre-service teacher attitudes toward SBG were maintained (Wuttke & St. Pierre, 2016).

Summary

Within achievement goal theory, there are four achievement goals in which humans orient themselves. They include (a) mastery-approach, (b) performanceapproach, (c) mastery-avoidance, and (d) performance-avoidance (Baranik et al., 2007; Elliot & McGregor, 2001; Finney et al., 2004; Miksza, 2009). They are domain specific (Baranik et al., 2007; Throndsen & Turmo, 2013), and each person has a goal orientation unique to each academic setting. In a classroom context, it is believed that goal-oriented teachers would most like to see their students espouse a mastery-approach (cf. Schraw & Aplin, 1998), which is linked to (a) highly intrinsic motivation (Elliot & Church, 1997; Elliot & Harackiewicz, 1996; Harackiewicz et al., 1997; Malmberg, 2006) and (b) longterm retention of material (Elliot & McGregor, 1999, 2001; Graham & Golan, 1991; Midgley et al., 2001). In turn, students can change their achievement goals based on their perceived values of the teacher in a classroom context (Ames, 1992; Ames & Archer, 1988).

To maximize student mastery goals, the practice of SBG is recommended. Previous research also suggests the practice of SBG become part of methods course training to (a) emphasize the importance of authentic arts assessment (Asmus, 1999; Colwell, 1998), (b) maximize the prevalence of mastery-based classrooms (Ames, 1992), and (c) end the tendency to teach as one was taught (Ester, 1997; McCoy, 1991).

Teacher goal orientations in the school environment are mostly stable, but there is evidence that they can be changed (Anderson et al., 1990; Malmberg, 2008). Given that exposure to SBG can change pre-service teacher attitudes toward grading policies (Wuttke & St. Pierre, 2016) the questions remain: (a) which achievement goal orientations are ideal for the anticipated use of SBG? and (b) which attitude scores are ideal for the anticipated use of SBG? With improved awareness, college methods professors might be able to change aspects of the methods courses to make them more ideal for the development of these types of pre-service teachers. The proposed method for answering these questions are presented in Chapter Three.

30

Chapter Three: Methodology

The primary purpose of this study was to examine the interplay between achievement goal orientation (AGO) and attitudes toward grading practices. A secondary purpose was to identify homogeneous groups of participants for further analysis. This study further investigated the effect of these groups on pre-service music teachers' attitudes toward grading practices as well as their anticipated use of standards-based grading (SBG). To accomplish this investigation, the following research questions were considered:

- 1. What are the descriptive statistical characteristics of the examined variables?
- 2. To what extent do AGQ scores correlate with the other variables?
- 3. What clusters emerge related to scores on the AGQ subscales?
- 4. What is the effect of AGO Profile on the other variables?
 - a. On attitude toward grading practices?
 - b. On anticipated use of SBG?

This chapter will describe the research design, variables, participants,

instrumentation, and procedures that were used to collect and analyze the data to answer each of the research questions.

Research Design

The present study involved an *ex-post facto* design, which examined several psychological variables that may influence whether pre-service music teachers anticipate using SBG models in their classrooms. This design was deemed appropriate, as the psychological variables, which were being examined, have existed in the participants prior to the current study. Through *ex-post facto* research, it is possible to examine relationships between these variables and subsequent behaviors (Leedy & Ormrod, 2009).

Although *ex-post* facto designs are non-experimental, they still allow researchers to test hypotheses. In order to answer the final research question, two null hypotheses were considered:

- 1. The AGO Profile would not be associated with attitude toward grading practices (GPAS) scores with an alternate hypothesis that the GPAS mean scores using AGO Profile as a predictor would differ significantly and be a better fit than the null model: H₀: $\mu_1 = \mu_2 = ... = \mu_k$
- 2. The probability of anticipating use of SBG would be the same regardless of AGO profile.

For the first hypothesis, the dependent variable was attitude toward grading practices. This is a continuous variable. The appropriate inferential method, in this case, is a one-way ANOVA, which will be explained in more detail in the following section. For the second hypothesis, the dependent variable was anticipated use of SBG. This is a dichotomous variable. Given that the AGO Profile and anticipated use of SBG are categorical variables, the appropriate inferential method would be an exact test (e.g., Chisquare, Fisher's exact).

Variables

There were several other variables considered. Each of the four AGOs was

measured as a continuous variable: (a) mastery-approach; (b) performance-approach; (c)

mastery-avoidance; and (d) performance-avoidance. Attitude toward grading practice was

also a continuous variable. Each of the variables in the present study, and their

classifications, can be found in Table 1.

Table 1

Classifications of Measured Variables

Variable Name	Variable	Abbreviation
	Туре	
Achievement Goal Orientation		
Performance Approach	Continuous	Perf_Ap_Total
Performance Avoidance	Continuous	Perf_Av_Total
Mastery Approach	Continuous	Mast_Ap_Total
Mastery Avoidance	Continuous	Mast_Av_Total
Attitude Toward Grading Practices	Continuous	GPAS_Total
Anticipated Use of SBG	Categorical	Use_SBG
AGO Profile (cluster membership)	Categorical	AGO Profile

The AGO Profile variable was calculated using a cluster analysis of the four AGO variables to obtain more thorough profiles of the participants involved. These profiles were then placed in homogeneous groups. Cluster analysis is an exploratory design that

does not make statistical inferences; the results are dependent on the variables selected and the similarity measures selected by the researcher, and as such, are typically not generalizable (Kent, 2015; Meyers et al., 2013). For the purpose of this study, a one-way ANOVA was employed in which the AGO Profile was used as the independent variable and attitude toward grading practices (GPAS) was used as the dependent variable. Cluster analysis was necessary prior to ANOVA because participants could (a) score high on measures of more than one achievement goal orientation (Midgley et al., 2001; Pintrich, 2000), and (b) pursue more than one goal (Eison, 1979; Midgley et al., 2001; Pintrich, 2000; Wentzel, 1989). As such, considering each variable separately would not provide complete information about the profiles of the participants with regards to the dependent variable. More specifically, this research sought to group participants based on a combination of four variables rather than any individual variable. For this purpose, cluster analysis was preferred. As Mooi and Sarstedt (2011) wrote, "an object in a certain cluster should be as similar as possible to all the other objects in the same cluster, it should likewise be as distinct as possible from objects in different clusters" (p. 238-239).

Subsequent Analysis

With the clusters in place, the AGO Profile could be used as a categorical variable in a subsequent analysis (Meyers, Gamst, & Guarino, 2013). For the purpose of this study, the subsequent analysis was a one-way ANOVA as attitude toward grading practices is a continuous variable. ANOVA compares the means of different groups using a quantitatively measured dependent variable (Meyers et al., 2013; D. S. Moore, McCabe, & Craig, 2012). Therefore, ANOVA was the appropriate statistical test in this case. An alpha level of $p \le .05$ was used to test the first null hypothesis listed above. Performing an ANOVA after a cluster analysis is a research method that has been used by Bae and Lee (2015) and Odoom (2016).

The next analysis determined whether anticipated use of SBG changed according to participant AGO Profile. Given that both variables were categorical, an exact test was needed to investigate the second null hypothesis. Chi-square is the most common of these tests, however, as will be seen in the following section, the sample size for the current study was too small to use that analysis. In instances of small sample size, statisticians recommend Fisher's exact test of independence:

Use Fisher's exact test of independence when you have two nominal variables and you want to see whether the proportions of one variable are different depending on the value of the other variable... Fisher's exact test is more accurate than the chi-square test or *G*-test of independence when the expected numbers are small... use Fisher's exact test when the total sample size is less than 1000 (McDonald, 2014, p. 77)

Participants

Participants. Participants were undergraduate pre-service music teachers in their third-year of study. This year of study was selected based on previous theory. Malmberg (2008) discovered that preservice teacher goal orientations oscillate over the course of their teacher training. These AGOs were at their peak during the third year of teacher training. For this reason, third year undergraduates were selected as the target sample population.

Sampling. The third-year pre-service music teachers who participated in this study were enrolled in state-approved teacher preparation programs at institutions of higher education within the United States. These participants were obtained using a combination of probability and non-probability sampling techniques. First, 50 certification-granting music schools in the Mid-Atlantic region of the United States were randomly identified, and the first round of participants was selected from those institutions. This is a probability sampling technique known as *cluster sampling* where potential participants exist in clusters (i.e., institutions), then those clusters are randomly selected for sampling (Leedy & Ormrod, 2009). This sampling technique is very sensitive to the characteristics of the population (Creswell, 2011). Cluster sampling was not feasible to make a list of every state-approved teacher preparation program (Leedy & Ormrod, 2009).

Once the institutions were identified, participants were obtained using *snowball sampling*. Music education professors at each of the institutions were directly contacted by the researcher and asked to forward the research materials to their third-year preservice music teachers. This eliminated the necessity to calculate the return rate, and reduced the generalizability of the results (Creswell, 2011). Emerson (2015) advised:

The best way to reduce the influence of uncontrolled factors is to use random sampling, in which study participants are randomly identified from the population of people who meet the criteria for inclusion in the study. Random sampling is, however, generally far too expensive and cumbersome for researchers to accomplish (p. 166).

This potential loss in generalizability was not problematic as cluster analysis is an exploratory design that does not make statistical inferences; the results are dependent on the variables selected and the similarity measures selected by the researcher. As such, they are typically not generalizable. (Kent, 2015; Meyers, Gamst, & Guarino, 2013). Cluster solutions are most helpful when used with other analyses (Kent, 2015).

Data collection using this combination of cluster sampling and snowball sampling took place over the course of approximately four months. Despite the longevity of this data collection period, and the direct contact with music education professors at the various institutions, a very small number of participants were obtained using this sampling method (n = 12). It was therefore necessary to consider other sampling methods. To accomplish this, the researchers used a form of purposeful sampling called *homogeneous sampling* where one "purposefully samples individuals or sites based on membership in a subgroup that has defining characteristics" (Creswell, 2011, Kindle Locations 6179-6180). In this case, the subgroup was third-year undergraduates in a music teacher preparation program. For this sampling procedure, the researchers employed the National Association for Music Education (NAfME) Research Assistance Program. This program allowed the researchers to send the research materials directly to third-year undergraduate students enrolled in music education licensure programs via indirect access using the e-mail transmission platform provided by NAfME. This data

collection took place over the course of one month and yielded an additional set of participants (n = 46). The total number of participants in the current study was N = 58. Measures

Achievement goal orientations. This study used the Attitude Toward Learning and Performance in College This Semester questionnaire (Finney et al., 2004), which is a modified version of the Achievement Goal Questionnaire (AGQ) developed by Elliot and McGregor (2001). This instrument, as modified by Finney et al., (2004) was specifically designed to measure four distinct AGOs in a general academic context.

As part of the validation process for this instrument, a factor analysis was conducted. Four factors were confirmed, which accounted for 81.5% of the total variance. Reliability for each of the scales was measured using Cronbach's Alpha, and reliability scores ranged from .83 to .92. As the developers concluded, "in sum, the CFA and reliability data clearly indicate that the four achievement goal measures represent empirically separable and internally consistent variables" (Elliot & McGregor, 2001, p. 507).

When the AGQ was developed in 2001, items measuring mastery-approach, performance-approach, and performance-avoidance goals were selected from other achievement goal questionnaires (e.g., Elliot & Church, 1997), and new items were created for mastery-avoidance goals. After a series of pilot studies, 12 items were selected: three for each achievement goal (Elliot & McGregor, 2001). For this scale, "participants indicated the extent to which they thought each item was true of them on a 1 (not at all true of me) to 7 (very true of me) scale" (Elliot & McGregor, 2001, p. 503). Within the context of the present study, one potential problem with the original AGQ was that it is classroom-specific. Given that it cannot be said for certain whether all the participants at any of the participating institutions would have a specific course together in which the instrument can be administered, it was best to examine AGOs for the entire semester. Measuring the achievement goal orientations in this manner would make the variable academic domain-specific, rather than classroom-specific.

A better to way to measure this domain-specific variable would be through the administration of the Attitude Toward Learning and Performance in College This Semester questionnaire (Finney et al., 2004). This is a modified version of the AGQ that specifies to the responder that information was related to coursework over the current semester rather than any specific course being taken. A confirmatory factor analysis found that "all standardized pattern coefficients had values greater than .50, with the majority having values at or greater than .70" (Finney et al., 2004, p. 356). Reliability was measured using Cronbach's Alpha and all alpha levels were between .68 and .88 (Finney et al., 2004).

Attitude Toward Grading Practices. To measure this variable, the researcher used the Grading Practice Attitude Survey (GPAS) developed by Wuttke and St. Pierre (2016). The instrument contains 23 prompts with a possible range of scores from 0–230. Higher overall scores indicate a preference for achievement-based grading models, while lower scores represent a preference for traditional (i.e., hodgepodge) grading schemes. Validity information was not reported, but high reliability was reported (r =.84). The instrument employs a 10-point rating scale where 10 indicates complete agreement and zero indicates complete disagreement (Wuttke & St. Pierre, 2016). Eleven of the 23 items are negatively worded, and reverse scoring was employed (i.e., higher scores indicate preference for hodgepodge grading and must be reversed for instrument consistency). Wuttke and St. Pierre (2016), however, did not make this explicitly clear in publication.

Anticipated use of SBG. The GPAS includes a dichotomous (Yes/No) item asking participants whether or not they plan on using standards-based grading (SBG) in their future classrooms even if they are not required to do so by their school districts (Wuttke & St. Pierre, 2016). The responses to this item will provide the data needed to measure the dependent variable of anticipated use of SBG.

Procedure

Prior to data collection, approval was obtained through the Institutional Review Board to conduct research on human subjects. This study was one of minimal risk to the participants and the harm or discomfort anticipated as part of this study did not exceed that of daily life or which is typically encountered during psychological examinations or academic tests.

Research materials were then distributed to music education professors at each of the 50 randomly selected institutions. Professors were directly contacted by the researchers and asked to forward the research materials to their third-year pre-service music teachers, or college juniors. Data collection in this manner took place over the course of four months. For a variety of reasons, all of which were beyond the scope of this study, this data collection method yielded only 12 participants (approximately 0.25 participants per school). The data collection process continued by sending research materials to third-year undergraduate students enrolled in music education licensure programs via indirect access using the e-mail transmission platform provided by NAfME. An email was sent three times to the collegiate members of NAfME (N = 11,934). This email was opened and read a total of 12,411 times, indicating that a certain number of collegiate members opened and read the email more than once. This method of data collection lasted for one month and yielded an additional set of participants (N = 46). This resulted in a response rate of less than 0.37%. It is not known how many of the collegiate members who did not respond were third-year undergraduates and how many were ineligible for participation. The total number of participants collected for the current study was N = 58.

Research materials included copies of the Achievement Goal Questionnaire (AGQ), Grading Practices Attitude Survey (GPAS), and a video on standards-based grading in music education. The video provided a rationale for, definitions of, and practical uses for standards-based grading in the music classroom. The video was included to remove lack of knowledge regarding SBG as a confounding variable as St. Pierre and Wuttke (2015) determined that the most common reason why music teachers were not using SBG was due to participants' lack of awareness toward the topic. Thus, responses on the GPAS could be considered representative of participants' informed attitudes toward grading practices free from the influence of lack-of-understanding.

The items comprising the GPAS and AGQ measures were combined into one form so participants would be able to complete both measures at the same time. These measures were made available to participants electronically via Google Forms to facilitate data collection and protect the anonymity of the participants. This internet-based program randomizes the question order, and prevents participants from submitting multiple times. The information is automatically entered into a spreadsheet to facilitate analysis. Print copies were made available, but no participants took advantage of that medium.

Summary

The present study involved a quantitative *ex-post facto* investigation of the effects of AGO on the pre-service teacher's attitude toward grading practices and anticipated use of SBG models in their future classrooms. Participants (N = 58) were third-year music education majors enrolled in teacher training programs at certification-granting institutions throughout the United States. The sample was obtained using a combination of probability and non-probability sampling. Research materials included (a) a modification of the Achievement Goal Questionnaire (AGQ) (Finney et al., 2004), (b) the Grading Practice Attitude Survey (GPAS) (Wuttke & St. Pierre, 2016), and (c) an informational video that outlined standards-based grading in the music classroom. The researchers contacted music education professors at 50 institutions and asked them to distribute the research materials to participants to protect their anonymity from the researcher. This yielded a low return rate over the course of four months (N = 12). Participants were further gathered via indirect email as part of the NAfME Research Assistance Program. The response rate was still extremely low (less than 0.37%) and total participants for this project were N = 58.

42

Participants watched the informational video, and then completed the AGQ and GPAS, which were made available electronically via an internet-based data collection resource. Once collected, the scores on the AGQ were entered into a cluster analysis to locate patterns in the responses of the participants and homogeneously group participants together. AGO Profile was considered a new independent variable. A one-way ANOVA compared AGO Profile to the dependent variable of attitude toward grading practices to determine if the mean GPAS scores among the homogeneously grouped clusters differed significantly. A Fisher's exact test was conducted to determine if the probability of anticipating use of SBG changes with AGO Profile. Results of these analyses will be reported in Chapter Four.

Chapter Four: Data and Analysis

The primary purpose of this study was to examine the interplay between achievement goal orientation (AGO) and attitudes toward grading practices. A secondary purpose was to identify homogeneous groups of participants for further analysis. This study further investigated the effect of these groups on pre-service music teachers' attitudes toward grading practices as well as their anticipated use of SBG. To accomplish this investigation, the following research questions were considered:

- 1. What are the descriptive statistical characteristics of the examined variables?
- 2. To what extent do AGQ scores correlate with the other variables?
- 3. What clusters emerge related to scores on the AGQ subscales?
- 4. What is the effect of AGO Profile on the other variables?
 - a. On attitude toward grading practices?
 - b. On anticipated use of SBG?

The following chapter will explain the statistical analyses and report the findings with regards to each of the four research questions. To answer the first research question, data derived from the observed variables were evaluated to (a) describe the characteristics of the sample, (b) remove any univariate and/or multivariate outliers, and (c) check assumptions for normality. To answer the second research question, a correlation analysis was conducted to examine the interrelationships between the observed variables. Next, a cluster analysis was conducted to group participants based on their AGQ responses to answer the third research question. Finally, to answer the fourth research question, an analysis of variance (ANOVA) and a Fisher's exact test were conducted to determine if significant differences existed among participants' AGO Profile profiles according to the observed variables.

Research Question One: Descriptive Statistics

Item Reverse Scoring. Eleven of the 23 items comprising the GPAS are

negatively worded. As such, these items were reverse coded (i.e., higher scores indicate

preference for hodgepodge grading and must be reversed for instrument consistency).

Wuttke and St. Pierre (2016), however, did not make this explicitly clear in publication.

Table 2 displays the original item number and wording of the eleven items that were

reverse coded.

Table 2

Item Abbreviation	Wording
GPAS03	Both practice logs and music performance tests should be used
	to measure student proficiency.
GPAS08	Grading that focuses on averaging all of the scores received on
	tests and other assignments indicates that the student has learned
	the material.
GPAS10	Students with good behavior that participate in class regularly
	will receive the highest academic grades.
GPAS11	Test results are final, there are no "do-overs" regardless of
	whether or not the student is happy with their grade.
GPAS13	Due to the subjective nature of the art form, student achievement
	in music is best measured on a case-by-case basis.

Reverse-Coded GPAS Item Numbers and Exact Wording

Item Abbreviation	Wording
GPAS14	Assigning academic grades based upon judgments of character
	and teacher expectations is more often than not, a clear-cut
	representation of a student's abilities.
GPAS15	The purpose for assigning academic grades is to show the
	"whole picture" of the student including, but not limited to:
	effort, achievement, participation, attendance, compliance, and
	behavior.
GPAS16	Students who do not finish assignments should be issued a zero
	in the grade book because it will teach the student to work
	harder on their next assignment.
GPAS18	Grading "on a curve" is good practice because it is a more
	effective way to show parents and administrators where each
	student scored in comparison with each other.
GPAS20	Extra credit that rewards students for their additional effort is
	sensible practice and provides a meaningful representation of
	musical achievement.
GPAS23	Because everyone has different opinions about musical likes and
	dislike, music performance tests are not a legitimate measure of
	student ability, achievement, and proficiency in music.

Note. As these items are currently worded, higher scores indicate affinity for grading practices associated with "hodgepodge" grading rather than SBG. These items were reverse coded to be consistent with the other test items.

The presence of outliers can adversely affect the results of a statistical analysis (Meyers et al., 2013). Outliers can be univariate (viz. occurring on one variable) or multivariate (viz. occurring on multiple variables). To objectively determine the presence of multivariate outliers, the researcher calculated the Mahalanobis distance (D^2) of each case. This D^2 statistic was evaluated using a chi-square table with the degrees of freedom equal to the number of variables and a stringent alpha level of .001. The cut-off D^2 statistic was determined to be 20.52. A multiple linear regression analysis was conducted with the case ID number as the dependent variable and five independent variables

resulting in D^2 statistics for each of the cases. None of the cases had a D^2 statistic larger than 20.52. Therefore, the data set did not include any multivariate outliers.

To identify univariate outliers, the scores representing the GPAS and AGQ subscales were standardized through the calculation of *z*-scores. As a general rule, *z*-scores that lie outside the realm of ± 2.5 standard deviation units should be considered outliers (Meyers et al., 2013). Two cases were found to be univariate outliers according to the Mast_Ap_Total variable and were considered possible candidates for elimination. Cohen, Cohen, West, and Aiken (2003) recommend that if there are only a few outliers (1-2%) it is best to leave them in the data set. Due to the small sample size representing the current study (N = 58), these outliers represent approximately 3% of the data set.

Prior to removing these cases, the skewness and kurtosis of all the variables were examined to see if these outliers were affecting the variable distribution. As defined by Kim (2013), "Skewness is a measure of the asymmetry and kurtosis is a measure of 'peakedness' of a distribution" (p.52). These can be evaluated by dividing the statistic by its standard error and comparing the values to a standard normal table of *z*-scores (Meyers et al., 2013). The calculated scores are presented in Table 3.

Table 3

Skewness and Kurtosis Values for Continuous Variables (N = 58)

47

Variable	Skewness		Ku	irtosis	Z-1	z-scores	
	Statistic	Std.	Statistic	Std.	7.1	Zkurtosis	
	Statistic	Error	Statistic	Error	2.skewness		
Mast_Ap_Total	-1.37	0.31	-2.36	0.62	-4.35	-3.81	
Mast_Av_Total	-0.01	0.31	-0.88	0.62	-0.03	-1.42	
Perf_Av_Total	-0.34	0.31	-0.44	0.62	-1.09	-0.71	
Perf_Ap_Total	-0.35	0.31	-0.69	0.62	-1.12	-1.11	
GPAS_Total	-0.30	0.31	-0.36	0.62	-0.96	-0.58	

One of the variables (Mast_Ap_Total) displayed inappropriate distribution under these circumstances. This provided further evidence that the univariate outliers were going to be problematic in future analyses that included the Mast_Ap_Total variable. Considering the *z*-scores and skewness/kurtosis indices, two cases were removed from analyses. In sum, two cases were removed as either univariate or multivariate outliers thus reducing the total sample size (N = 56). The skewness and kurtosis values for the data set post removal of the outliers are presented in Table 4. The new data set shows all skewness and kurtosis values within the parameters of ±1, which some statisticians say demonstrates appropriate distribution (Meyers et al., 2013).

Table 4

Variable	Skewness		Ku	Kurtosis		z-scores	
	Statistic	Std.	Statistic	Std.	F .	-	
	Statistic	Error	Statistic	Error	Zskewness	Zkurtosis	
Mast_Ap_Total	-0.632	0.31	-0.70	0.63	-1.98	-1.12	
Mast_Av_Total	-0.025	0.31	-0.88	0.63	-0.08	-1.40	
Perf_Av_Total	-0.355	0.31	-0.38	0.63	-1.11	-0.60	
Perf_Ap_Total	-0.415	0.31	-0.64	0.63	-1.30	-1.02	
GPAS_Total	-0.213	0.31	-0.43	0.63	-0.67	-0.68	

Skewness and Kurtosis Values for Continuous Variables After Outlier Removal (N = 56)

Descriptive Statistics. The sample was comprised of a sample (N = 56) of thirdyear undergraduate pre-service music teachers. It should be noted that participants' demographic data were not collected as part of this study.

Achievement goal orientations (AGO). AGO variables were calculated using items from the AGQ. Each item was scored using a seven-point scale. These scores were added together and the following variables were created:

- Performance-Approach total score (Perf_Ap_Total)
- Performance-Avoidance total score (Perf_Av_Total)
- Mastery-Approach total score (Mast_Ap_Total)
- Mastery-Avoidance total score (Mast_Av_Total)

The means and standard deviations for each of these variables are displayed in

Table 5. These variables all had skewness and kurtosis statistics demonstratingappropriate distribution (see Table 4).

Table 5

Variable	Minimum	Maximum	Mean		Standard Deviation
	Statistic	Statistic	Statistic	Std. Error	Statistic
Mast_Ap_Total	12	21	18.23	0.35	2.59
Perf_Ap_Total	3	21	13.96	0.63	4.72
Perf_Av_Total	3	21	12.98	0.58	4.34
Mast_Av_Total	5	21	12.66	0.63	4.70

Descriptive Statistics for AGO Measures

In each of the AGO subscales, at least one participant possessed the highest possible score (21). In the Mast_Ap_Total subscale the lowest obtained score was 12. This indicates that this sample of participants had reasonably high Mastery-Approach orientation (M = 18.23, SD = 2.59). The other three AGO subscales exhibited wide variability.

Attitudes toward grading practices. This variable was calculated using the 23 items from the GPAS. Each item was scored on a ten-point scale. These scores were added together and the new variable (GPAS_Total) was created. The minimum score was 121, and the maximum score was 197. Overall, composite scores representing the GPAS_Total variable (M = 152.93, SEM = 2.27, SD = 17.02) exhibited adequate variability.

Anticipated use of SBG. This was a dichotomous response item on the GPAS. Most of the participants provided affirmative Use_SBG responses [Yes (n = 49), No (n = 7)].

Research Question Two: Correlation Analysis

Variable Correlations. The following variables exist outside of the cluster analysis: (a) Achievement Goal Orientation, (b) Attitude Toward Grading Practices, (c) Anticipated Use of SBG, (d) AGO Profile. Pearson product-moment correlation coefficients were calculated for each (see Table 6). A statistically significant correlation was found between Perf_Av_Total and Perf_Ap_Total ($r = .71, p \le .01$). In addition, total score on the GPAS shared statistically significant correlations ($p \le .01$) with the following variables: (a) a weak negative correlation with Perf_Ap_Total (r = ..35), (b) a moderate negative correlation with Perf_Av_Total (r = -.47), (c) a weak positive correlation with Mast_Ap_Total (r = .37), and (d) a weak-to-moderate positive correlation with Use_SBG (r = .39). All absolute coefficient values were interpreted according to Evans (1996).

Table 6

Correlations Between Measured Variables

Measure	1	2	3	4	5
Perf_Ap_Total					
Perf_Av_Total	.71**				
Mast_Ap_Total	11	23			
Mast_Av_Total	.08	.14	13		
GPAS_Total	35**	47**	.37**	13	
Use_SBG	20	17	.06	.09	.39**

***p* ≤ .01

GPAS Factor Structure. The GPAS is a relatively new instrument that previously had been used in only one peer-reviewed research study (Wuttke & St. Pierre, 2016). In that study, the researchers reported test-retest reliability on the instrument (r =.84), although factor analysis procedures were not conducted. For the current study, an exploratory factor analysis for the GPAS was performed on the data from 56 pre-service music teachers. An oblique rotation strategy was employed due to component correlations. These correlations can be seen in

Table 7.

Table 7

Component	1	2	3	4	5	6	7
1							
2	-0.04						
3	0.19	0.19					
4	0.27	0.20	0.21				
5	0.20	0.03	-0.05	0.00			
6	0.08	0.15	0.02	0.18	-0.03		
7	0.22	0.05	-0.19	0.06	0.26	0.02	
8	0.12	0.05	-0.04	-0.16	0.20	-0.06	0.11

GPAS Component Correlation Matrix

Due to the small sample size, a problematic variables-to-cases ratio was anticipated. The Kaiser–Meyer–Olkin (KMO) measure of sampling adequacy for the first analysis was .48, which is considered "unacceptable" for analysis (Kaiser, 1974, p. 35). Bartlett's test of sphericity was significant (p < .001) demonstrating appropriate correlations between the variables for analysis. This first analysis found eight factors with eigenvalues greater than 1.00, which combined to explain 69.51% of the total variance. Item 8 (Grading that focuses on averaging all of the scores received on tests and other assignments indicates that the student has learned the material) and Item 18 (Grading 'on a curve' is good practice because it is a more effective way to show parents and administrators where each student scored in comparison with each other) failed to load meaningfully (< .45) on any of the factors and were removed from further analyses. The next run extracted eight factors and the total variance explained increased to 72.11%. Item 17 (Teachers should allow students to finish incomplete assignments so long as the initial assignment was submitted on time) failed to load meaningfully and was removed from the instrument. The analysis was run again and eight factors were once again extracted. The KMO measure increased to .48 (unacceptable) and the total variance explained increased to 74.05%. Item 1 (Grading should focus on the material that has or has not been learned rather than on accumulating points to reach a certain average or total) and Item 16 (Students who do not finish assignments should be issued a zero in the grade book because it will teach the student to work harder on their next assignment) loaded together strongly forming a unique factor. These items seemed to be related regarding the use of gradebook "points," but the inclusion of this eighth factor seemed to confuse interpretation. As such, these two items comprising the eighth factor, were removed from the instrument.

Finally, a principal component analysis with an oblique rotation solution was performed on the remaining 18-items of the GPAS. The KMO measure of sampling adequacy increased to .50, which is classified as "miserable," but no longer "unacceptable" (Kaiser, 1974, p. 35). Bartlett's test of sphericity was again significant (p< .001) demonstrating appropriate correlation between the variables for analysis. The analysis extracted seven factors with eigenvalues greater than 1.00, which together explained 71.21% of the total variance. A summary of the analyses can be found in Table 8.

Table 8

Summary of Factor Solutions

GPAS Items	Factors	Total Variance Explained	KMO Measure	Classification
All 23 Items	Eight	69.51%	.48	Unacceptable
Without 8, 18	Eight	72.11%	.47	Unacceptable
Without 8, 17, 18	Eight	74.05%	.48	Unacceptable
Without 1, 8, 16, 17, 18	Seven	71.21%	.50	Miserable

Note. Bartlett's test of sphericity significant (p < .001) for all analyses

The seven factors indicate the GPAS is measuring seven subscales (specific

attitudes) toward grading practices. These subscales are:

- Acceptance of Grading Musical Ability (eigenvalue = 3.23)
- Acceptance of Grading Subjectively (eigenvalue = 2.35)
- Acceptance of Allowing Retakes (eigenvalue = 2.06)
- Acceptance of Grading Practice Logs (eigenvalue = 1.83)
- Acceptance of Grading Extra Credit (eigenvalue = 1.65)
- Acceptance of Grading Good Citizenship (eigenvalue = 1.24)
- Acceptance of Achievement-Only Grading (eigenvalue = 1.17)

Item 6 (Grading student behavior is acceptable only when that grade is kept separate from the academic grade) most strongly loaded with Acceptance of Achievement-Only Grading (.57), however it also loaded with Acceptance of Grading Extra Credit (-.50). This indicates that, to this group of respondents, giving students grades based on behavior is a form of extra credit. A complete list of GPAS items and their factor loadings is provided in Table 9. When the GPAS was first published, reliability was reported using test-retest correlation method, and the reliability coefficient was reasonably high (r = .84). To further test the reliability, Cronbach's coefficient alpha was calculated for the GPAS in the current study (.61). This statistic is an indicator of the consistency with which participants complete the measure (i.e., internal consistency). Higher alpha levels are preferred, but levels around .65 are considered acceptable for research purposes; alpha levels below that are problematic (Meyers et al., 2013). Cronbach's alpha was calculated for each of the seven subscales as well. These results are presented in Table 9. There were several problematic alpha levels. This could be due to the small sample size, but it could also indicate that the instrument itself is problematic and should be more clearly rewritten if used in future research to facilitate consistent responses from participants.

Table 9

Acceptance of:	Item Number and Exact Wording	Pattern Coefficient	α
Grading Musical Ability	12. The best way to assess student musical achievement is by comparing what students know and can do with criteria they should be expected to master	.91	.74
	05. Scores should be recorded by type, such as tests, homework, or performance tests in order to clearly identify the areas of concern for each student	.66	
	04. Although musical preferences are subject to opinion, performance tests can be constructed so that they fairly and accurately measure student ability, achievement, and proficiency in music	.65	

Summary of GPAS Subscales, Items, Pattern Coefficients, and Cronbach's Coefficient a

Acceptance of:	Item Number and Exact Wording	Pattern Coefficient	α
	23. Because everyone has different opinions about musical likes and dislike, music performance tests are not a legitimate measure of student ability, achievement, and proficiency in music	.88	
Grading Subjectively	13. Due to the subjective nature of the art form, student achievement in music is best measured on a case-by-case basis	.74	.67
	15. The purpose for assigning academic grades is to show the "whole picture" of the student including, but not limited to: effort, achievement, participation, attendance, compliance, and behavior	.69	
Allowing Retakes	19. After a low score on a performance test, the student should be allowed multiple opportunities to retake it to demonstrate mastery	91	.74
Retakes	11. Test results are final, there are no "do- overs" regardless of whether or not the student is happy with their grade	.85	
	tests should be used to measure student proficiency	81	
Grading Practice Logs	22. Grades from music performance tests should be used to measure student proficiency; homework (such as practice logs), should not	.76	.55*
Grading Extra	20. Extra credit that rewards students for their additional effort is sensible practice and provides a meaningful representation of musical achievement	81	64
Credit	07. A student's academic grade is skewed by receiving extra credit for effort and does not reflect true musical achievement	.81	
	participate in class regularly will receive the highest academic grades	.92	
Grading Good Citizenship	14. Assigning academic grades based upon judgments of character and teacher expectations is more often than not, a clear- cut representation of a student's abilities	.55	.58*

Acceptance of:	Item Number and Exact Wording	Pattern Coefficient	α
	21. Basing academic grades on teacher expectations and character judgments may cause the students to perform in ways that do not accurately reflect their abilities	73	
Achievement- Only Grading	02. Other factors, such as effort and attitude are essential, but should not be indicative of the students who have the highest academic grades	70	.60*
	09. The intent for issuing an academic grade is to report how well a student has mastered what has been taught	70	
	06. Grading student behavior is acceptable only when that grade is kept separate from the academic grade	.57	

Note. *Alpha levels problematic (Meyers et al., 2013)

AGQ Factor Structure. As mentioned in Chapter Three, the AGQ used in this study has been previously validated in previous research. However, a confirmatory factor analysis was conducted to determine if items in this data set would load like those previously reported. The Kaiser–Meyer–Olkin (KMO) measure of sampling adequacy was .70, which Kaiser (1974) classified as "middling" (p. 35). Bartlett's test of sphericity was significant (p < .001) demonstrating appropriate correlation between the variables for analysis. The extraction is presented in Table 10.

Table 10

AGQ Confirmatory Factor Analysis: Total Variance Explained

Component	Extraction Sums of Squared Loadings		
	Eigenvalue	Cumulative % of Variance	
Performance-Approach	3.96	33.03	
Mastery-Avoidance	2.46	53.53	
Mastery-Approach	1.96	69.84	
Performance-Avoidance	0.88	77.20	

The four factors explained 77.20% of the total variance, although one factor had an eigenvalue < 1.00. The items loaded in like manner to previous studies (i.e., Elliot & McGregor, 2001; Finney et al., 2004) with this exception; item 5 (The fear of performing poorly is what motivates me) was the only item that strongly loaded on the Performance-Avoidance scale and the other performance-avoidance items loaded with the Performance-Approach scale. This factor solution is most likely due to the "middling" sampling adequacy coupled with the strong positive correlation between Perf_Ap_Total and Perf_Av_Total (r = .71). More research would be needed prior to making any claims regarding whether this factor solution suggests an AGO trend unique to music. Reliability for these scales was measured using Cronbach's Alpha and the results are presented in Table 11.

Table 11

Summary of AGQ Subscales, Items, Pattern Coefficients, and Cronbach's Coefficient a

Subscale	Item Number and Exact Wording	Pattern Coefficient	α
Mastery- Approach	12. The most important thing for me this	.88	77
	courses as thoroughly as possible		.//

Subscale	Item Number and Exact Wording	Pattern Coefficient	α
	10. Completely mastering the material in my	.81	
	courses is important to me this semester 11. I want to learn as much as possible this	.81	
	semester		
	08. I worry that I may not learn all that I possibly	.92	
Mastery- Avoidance	09. I am definitely concerned that I may not learn all that I can this semester	.89	.89
	07. I am afraid that I may not understand the content of my courses as thoroughly as I'd like	.85	
	03. I want to do better than other students this	.92	
Performance-	01. My goal this semester is to get better grades	.91	
Approach	06. My goal this semester is to avoid performing	.88	.90
	02. It is important for me to do well compared to other students this semester	.85	
	04. I just want to avoid doing poorly compared to other students this semester	.55	
Performance- Avoidance	05. The fear of performing poorly is what motivates me	.98	

Research Question Three: Cluster Analysis

Even though analysis of the AGQ did not extract factors in exactly the same manner as previous studies, there was evidence to suggest that may have been the case had there been a more robust sample size (see Elliot & McGregor, 2001; Finney et al., 2004). It was therefore determined that all four subscales should be used in the cluster analysis. Cluster analysis was used to examine psychological patterns that existed among the participants with regards to AGO.

Clustering Procedure. The 56 participants had valid scores on four AGO

measures: (a) Mastery-Approach, (b) Mastery-Avoidance, (c) Performance-Approach,
and (d) Performance-Avoidance. These values were standardized (*z*-scores) to interpret levels (e.g., high, average, low, etc.) within this population of respondents. The *z*-scores were analyzed in a hierarchical cluster analysis employing Minkowski distance with Ward's method. This clustering procedure has been found to be one of the most accurate for continuous data sets (Fonseca, 2013). A five-cluster solution was determined to provide the best fit based on a visual examination of the dendrogram (Figure 1) and the coefficient change in the agglomeration schedule (Table 12). The large changes in the coefficient indicate the clusters are dissimilar. The cluster solution emerges when the values become small (Norušis, 2012). Therefore, the agglomeration schedule should be examined from the bottom. There are no additional large coefficient decreases after five clusters.



Figure 1. Dendrogram using Ward's linkage method.

Table 12

Agglomeration Schedule Using Minkowski Distance with Ward's Linkage Method

Stage	Coefficient
55	72.83
54	60.13
53	50.97
52	45.78
51	40.73
50	38.05
49	35.53

The *z*-scores were then analyzed by the k-means cluster procedure to determine group membership. Convergence was reached in four iterations. Univariate ANOVAs signified that the four AGO scales differed significantly in each cluster (all *ps* < .001). Final cluster centers with number of cases in each cluster are presented in Table 13. Cluster sizes were comparable, with a range of n = 8-16. An alternative three-cluster solution (Table 14) was examined due to the large coefficient drops in the agglomeration schedule at stages 55–53. The number of cases in each group was appropriate, but the mean cluster centers did not provide as much information about the group members as the five-cluster solution. For these reasons, the five-cluster solution was determined to be the best fit for this data set.

Table 13

Five-Cluster Solution with Cluster Names and Final Cluster z-Score Means

	Non- Mastery- Oriented	Non- Performance- Oriented	Performance- Oriented	Eclectic	Mastery- Oriented
	(<i>n</i> = 12)	(<i>n</i> = 8)	(<i>n</i> = 10)	(<i>n</i> = 16)	(<i>n</i> = 10)
Mast_Ap_Total	-1.57	.15	.41	.44	.64
Mast_Av_Total	.44	99	.77	91	.94
Perf_Av_Total	.37	-1.49	.74	.39	62
Perf_Ap_Total	.27	-1.48	.73	.66	92

Table 14

Alternative Three-Cluster Solution with Final Cluster z-Score Means

	1	2	3
	(<i>n</i> = 23)	(<i>n</i> = 12)	(<i>n</i> = 21)
Mast_Ap_Total	.50	.46	81
Mast_Av_Total	39	48	.70
Perf_Av_Total	.53	-1.28	.15
Perf_Ap_Total	.74	-1.46	.02

Naming the Clusters. Cluster 1 was labeled *Non-Mastery-Oriented* because the cases in this group had the lowest Mast_Ap_Total scores of any group and only slightly higher than average scores on the other three measures. Participants in Cluster 2 reported (a) the lowest scores on both performance-oriented measures and (b) average scores on Mast_Ap_Total and (c) low scores on Mast_Av_Total. Since this group contained the lowest performance-based orientation scores, it was named *Non-Performance-Oriented*. Cluster 3 contained those with the highest scores on both performance-based scales. It was therefore labeled *Performance-Oriented*. Cluster 4 included those with slightly higher-than-average Mast_Ap_Total, low Mast_Av_Total, and above average to moderate performance scores. This group was not exceptionally high or low on any one

orientation, and was labeled *Eclectic*. Cluster 5 was labeled *Mastery-Oriented* in that it was classified by the highest scores on both mastery-based orientation scales and low scores on both performance-based orientation scales.

Research Question Four: The Effect of AGO Profile on Other Variables

AGO Profile and Attitudes toward Grading Practices. The principal components analyses of the GPAS found between seven and eight subscales measuring attitudes toward specific grading practices. It is difficult, however, to make any statements regarding the construct validity of the instrument due to the extremely poor KMO measures of sampling adequacy (.48–.50). The GPAS_Total variable was therefore kept unaltered with the knowledge that it was measuring participant attitudes toward several grading practices the acceptance/rejection of which can point to an acceptance/rejection of SBG.

The five cluster groups were compared to GPAS_Total to find any mean differences among clusters with regards to attitudes toward grading practices. Higher scores on the GPAS are thought to indicate preference for grading practices closely aligned with SBG, and lower scores are thought to indicate preference for grading practices closely aligned with hodgepodge or traditional grading. The five clusters produced significantly different GPAS_Total scores, F(4, 51) = 2.94, p = .03, $\eta^2 = .19$. As such, 19% of the variance in GPAS_Total can be accounted for by AGO Profile.

Levene's test for homogeneity of variances was significant, F(4, 51) = 3.48, p = .01. The null hypothesis of homogeneity of variance was rejected. Tamhane's T2 test was

run as a post hoc analysis because Tamhane's T2 is the most conservative post hoc for unequal variances (Meyers et al., 2013). These post hoc results indicated that the Mastery-Oriented AGO Profile scored significantly ($p \le .03$) higher (M = 162.20, SD =10.56) than the Non-Mastery-Oriented AGO Profile (M = 143.33, SD = 14.99) and the Performance-Oriented AGO Profile (M = 145.90, SD = 10.72). This suggests that having high mastery-based goal orientations in addition to low performance-based goal orientations makes one more agreeable to grading practices typically associated with SBG.

AGO Profile and Anticipated Use of SBG. AGO Profile (AGO Profile) was compared to Use_SBG to determine if there any of the clusters were more likely than the others to anticipate using SBG in their future teaching environments. Given that both AGO Profile and Use_SBG were categorical variables an exact test was recommended. This data set indicated a 2 x 5 table where two is the number of Use_SBG anticipations and five is the number of cluster groups. The sample size (N = 56) indicated that Fisher's exact test was most appropriate because the test is more accurate than other tests for sample sizes less than N = 1000 (McDonald, 2014). The null hypothesis stated that all five cluster groups would be equally likely to anticipate using SBG when they fully enter the profession. Outcome results indicated no significant difference in Use_SBG proportion among the five cluster groups (3.29, p = .51). These results indicate that that each cluster group was equally likely to anticipate using SBG. The 2 x 5 table used for this analysis is displayed in Table 15

Table 15

			AG	O Profile			
		Non- Mastery- Oriented	Non- Performance- Oriented	Performance- Oriented	Eclectic	Mastery- Oriented	Total
Use_ SBG	No	1	2	2	2	0	7
	Yes	11	6	8	14	10	49
	Total	12	8	10	16	10	56

2 x 5	Table c	of Use ,	SBG	Cases	Compared	to AGO	Profile

Summary

This chapter presented a series of statistical analyses to answer four research questions:

- What are the descriptive statistical characteristics of the examined variables? (RQ1)
- 2. To what extent do AGQ scores correlate with the other variables? (RQ2)
- 3. What clusters emerge related to scores on the AGQ subscales? (RQ3)
- 4. What is the effect of AGO Profile on the other variables?
 - a. On attitude toward grading practices? (RQ4a)
 - b. On anticipated use of SBG? (RQ4b)

To answer RQ1 data were analyzed to describe the mean, standard error, standard deviation, minimum, maximum, skewness, and kurtosis for each of the measured variables. The data-cleaning process for identifying and removing univariate and multivariate outliers was also described. In total, two cases were removed as univariate

outliers on the Mast_Ap_Total variable bringing the total participants for this study to N = 56.

To answer RQ2 Pearson product-moment correlation coefficients were computed for each of the measured variables. A significant correlation was found between Perf_Av_Total and Perf_Ap_Total ($r = .71, p \le .01$). Significant correlations were also found between GPAS_Total and the following variables: (a) Perf_Ap_Total (r = ..35, p < .01), (b) Perf_Av_Total (r = ..47, p < .01), and (c) Mast_Ap_Total (r = ..37, p < .01). In addition, Use_SBG shared a significant correlation with GPAS_Total ($r = ..39, p \le .01$).

An exploratory factor analysis was conducted on the GPAS. Given that no previous factor solutions were presented in the literature, a series of solutions was presented in this study. Ultimately, a seven-factor structure, which accounted for 71.21% of the total variance, was adopted. These seven factors were (a) Acceptance of Grading Musical Ability, (b) Acceptance of Grading Subjectively, (c) Acceptance of Allowing Retakes, (d) Acceptance of Grading Practice Logs, (e) Acceptance of Grading Extra Credit, (f) Acceptance of Grading Good Citizenship, and (g) Acceptance of Achievement-Only Grading. Due to the inadequate sample size, the EFA results should be interpreted with caution.

The AGQ had been previously validated (see Elliot & McGregor, 2001; Finney et al., 2004), so a confirmatory factor analysis was run on the current data set. The CFA found four factors, which accounted for 77.20% of the total variance. Most of the performance-based items loaded together and only one item (The fear of performing poorly is what motivates me) was found to measure Perf_Av_Total. This result was most

likely due to sampling inadequacy (KMO = .70) as well as the strong correlation between Perf_Ap_Total and Perf_Av_Total (r = .71). Internal consistency was measured using Cronbach's Alpha and the results ($\alpha = .77$ —.90) were deemed appropriate for research purposes.

For the third research question, a hierarchical cluster analysis using Minkowski distance with Ward's linkage method was conducted on the scores representing the AGQ. Five homogeneous groups were identified: (a) Non-Mastery-Oriented; (b) Non-Performance-Oriented; (c) Performance-Oriented; (d) Eclectic; and (e) Mastery-Oriented.

The fourth research question was comprised of two sub-questions. To answer the first sub-question, a one-way ANOVA to determine if differences exist between the AGO Profiles on their attitudes toward grading practices. Significant mean differences were found between the groups [F(4, 51) = 2.94, p = .03, $\eta^2 = .19$]. The Mastery-Oriented AGO Profile scored significantly ($p \le .03$) higher (M = 162.20, SD = 10.56) than the Non-Mastery-Oriented AGO Profile (M = 143.33, SD = 14.99) and the Performance-Oriented AGO Profile (M = 145.90, SD = 10.72). To answer the second sub-question, a Fisher's exact test was conducted to determine if any AGO profiles reported higher frequency of anticipated use of SBG. The results were no significant difference (p = .51) between AGO profiles. Further implications of these findings and recommendations for future research are presented in Chapter Five.

Chapter Five: Discussion and Recommendations

Student assessment is an extremely important task for all teachers, although the assignment of student grades may not be the most glamorous part of a music teacher's job description. The process often involves several obstacles including time, class size, number of students, lack of planning time, lack of appropriate training, and lack of understanding (St. Pierre & Wuttke, 2015). For these reasons, and perhaps others, music teachers have historically graded their students using non-musical criteria (e.g., attendance, effort, behavior, participation, etc.), which could reside outside the approved curriculum. Standards based grading (SBG) has been suggested as a way to combat this, but there is still professional hesitance regarding its adoption (St. Pierre & Wuttke, 2015)

The purpose of this study was to examine the interplay between achievement goal orientation (AGO) and attitudes toward grading practices to identify homogeneous groups of participants for further analysis. It was theorized that the achievement goals to which teachers orient themselves might affect teacher preference for certain grading practices. These grading practices would point to a rejection or acceptance of the principle tenets of standards-based grading (SBG). To accomplish this investigation, the following research questions were considered:

 What are the descriptive statistical characteristics of the examined variables? (RQ1)

- 2. To what extent do AGQ scores correlate with the other variables? (RQ2)
- 3. What clusters emerge related to scores on the AGQ subscales? (RQ3)
- 4. What is the effect of AGO Profile on the other variables?
 - a. On attitude toward grading practices? (RQ4a)
 - b. On anticipated use of SBG? (RQ4b)

Participants for this study were third-year undergraduate pre-service music teachers enrolled in state-approved teacher licensure programs. Participants were recruited using a combination of probability and non-probability sampling techniques. Data collection took place over the course of six months, and materials were distributed to a sizeable number of prospective participants (N > 12,000). However, despite the long collection period and large number of dispersed materials, the response rate was low (< 0.40%). The total number of participants was N = 58. Participants completed two measures (AGQ and GPAS) that provided data on six variables including four AGOs, attitude toward grading practices, and anticipated use of SBG in future teaching. These variables were then used in subsequent analyses in order to answer the research questions.

Discussion

RQ1. The two variables of interest here were Mast_Ap_Total and GPAS_Total. The highest possible score for Mast_Ap_Total was 21 and the mean score was 18.23. This indicates that most of the participants had moderately high to very high masteryapproach AGO. Even the minimum score reported for this variable (12) was higher than the scale midpoint. Responses on the GPAS were varied, but the minimum recorded score was 121. This value is also higher than its scale midpoint. The highest recorded GPAS_Total score was 197. The mean and standard deviation for the GPAS in this study were lower than the statistics reported by Wuttke and St. Pierre (2016), who first used the instrument (M = 166.60, SD = 36.60). However, data from both studies indicated that all participants, even those who did not anticipate using SBG, have aspects of the grading system with which they agree and disagree.

Thirteen percent of participants (n = 7) in the present study indicated that they would not use SBG if not required to do so. Of those, 100% (n = 7) had GPAS_Total scores lower than the mean (M = 152.93). These results are similar to those reported by Wuttke and St. Pierre (2016) who found that 25% (n = 7) of sampled pre-service music teachers said they would not use SBG in future teaching and of those, 86% (n = 6) had GPAS scores lower than the mean score (166.60). This might suggest that GPAS scores may distinguish between those who have an affinity for SBG and those who do not.

RQ2. Pearson product-moment correlation coefficients were computed for each of the measured variables. A strong positive correlation was found between Perf_Av_Total and Perf_Ap_Total. Total score on the GPAS shared significant correlations with (a) Perf_Ap_Total, (b) Perf_Av_Total, (c) Mast_Ap_Total, and (d) Use_SBG. Among the AGQ variables, there was one significant correlation between Perf_Ap_Total and Perf_Av_Total. This result is remarkedly different from previous studies using the AGQ in which every variable significantly correlated (Finney et al., 2004) or almost every variable significantly correlated (Elliot & McGregor, 2001). Also of interest is Use_SBG was the only variable to share a significant correlation with GPAS_Total. As GPAS scores increased, so did the frequency of affirmative responses in Use_SBG. Unpublished raw data from Wuttke and St. Pierre (2016) showed a moderate positive correlation between GPAS score and anticipated use (r = .53, p < .01). The exact nature of this relationship is uncertain as the GPAS contains as many as eight subscales. Future research could test these subscales as predictors of Use_SBG. A preliminary analysis of the current data set showed that two of the subscales exhibited significant correlations with Use_SBG: (a) Acceptance of Grading Musical Ability $(r = .32, p \le .05)$, and (b) Acceptance of Allowing Retakes (r = .43, p = .001). These results should only be interpreted as guidance for future research due to small sample size and instrument error.

RQ3. A cluster analysis of AGQ data for 56 participants identified five unique AGO Profiles: (a) Non-Mastery-Oriented, (b) Non-Performance-Oriented, (c) Performance-Oriented, (d) Eclectic, and (e) Mastery-Oriented. A three-cluster solution was a possible alternative, but large coefficient drops in the agglomeration schedule suggested that the five-cluster solution was more appropriate for this data set.

RQ4a. A one-way analysis of variance (ANOVA) was conducted to determine if significant differences existed among AGO Profiles according to GPAS_Total The Mastery-Oriented AGO Profile was found to exhibit significantly higher GPAS_Total scores than both the Non-Mastery-Oriented and Performance-Oriented AGO Profiles. This suggests that neither high mastery-based AGQ scores nor low performance-based AGQ scores by itself are enough to affect GPAS_Total. Rather, it is the combination of

both high mastery-based orientations with low performance-based orientations that might lead to increased acceptance of grading practices typical of SBG. This finding is contrary to previous research in which comparisons between high mastery/low performance and high mastery/high performance always favor the high mastery/high performance group (Pintrich, 2000). SBG represents a departure from traditional points-based gradeaveraging systems. This could be contrary to the preferences of performance-approach and performance-avoidance AGOs as it is difficult to objectively compare oneself to others when there is no quantitative scale. SBG compares students to standards, not to other students.

RQ4b. The proportions of Use_SBG were compared to AGO Profile using Fisher's exact test to determine if any of the AGO Profiles had a higher concentration of affirmative Use_SBG responses. The test did not reveal statistical significance, which could be attributed to the low number of negative Use_SBG responses. The small sample size also could have affected this result, however, Wuttke and St. Pierre (2016) found that, out of a sample of 28 undergraduates, seven participants did not anticipate SBG use. This proportion is decidedly different from the one found in the current study.

Implications

Maintaining a proper balance between musical instruction and authentic assessment can be a difficult task. As such, some teachers might prefer their own approaches when assessing student achievement (St. Pierre & Wuttke, 2015). However, SBG is becoming much more common and many school districts are requiring teachers to use SBG. St. Pierre and Wuttke (2015) found that 36% of sampled music teachers were using SBG because they were required to do so. However, even with proper training, music teachers might be reluctant to accept all aspects of SBG. In fact, the administration of the GPAS provided insight into each individual hodgepodge preferred by participants.

No two teachers are alike, and as such, no two grading systems are alike. The large number of factors extracted from the GPAS in addition to the under-representation of participants who did not anticipate using SBG, especially among scores lower than the mean, point to a "cafeteria-style" approach to SBG. Specifically, when using SBG, music teachers might select the parts of SBG with which they agree and ignore the parts with which they disagree. If this is the reality, standards-based grading will unfortunately become just another hodgepodge grading scheme.

It is possible the reason SBG has not caught on in music education is due to a lack of exposure. However, the current results suggest that SBG may not have caught on because some members of the profession are psychologically predisposed to reject it. When individuals orient themselves toward performance goals, a large amount of their social and cognitive motivation is tied to their ability to compare their performance to others either by (a) displaying competence or (b) avoiding the display of incompetence. Therefore, one might be reluctant to accept a grading system that values comparison to standards over comparison to people. In sum, hesitation to adopt all aspects of SBG might be tied to philosophical preference for performance goals.

This presents a tremendous challenge for teacher-training programs. Does music education need to lower preference for performance-based goal orientations in a profession that instructs in a performance-based art form? This is further complicated by a popular culture that seems to value performance goals as they relate to music as evidenced by the number of televised singing competitions. To accomplish this, teacher trainers may need to emphasize the domain-specificity of achievement goals to help preservice teachers understand that performance goals might be appropriate in the studio or ensemble, but not when leading a classroom.

Limitations

While the present study could make some inferences regarding the psychological antecedents to grading practices, several limitations must be addressed. First, despite generalizability not being a primary focus of this study, the sample size was inadequate to make any generalizations about the population. Additionally, the small sample size limited the number of statistical tests that could be conducted on the data set. To perform the factor analyses, an important test to help establish the construct validity of the measurement instruments, statisticians recommend a sample of between 200 and 1,000. A sample 50 participants is considered very poor for such an analysis (Meyers et al., 2013). A second limitation involved construct problems with the GPAS. The proportionally large number of extracted factors, extremely low KMO statistics, and low Cronbach's coefficient α values indicated potential problems with using the GPAS as an indicator of preference for SBG. These two limitations are related, however, as the extent to which the small sample size affected the reliability and factor solution remains unknown.

Upon reflection, some potential reasons for the small sample size have been identified. Participants were required to use an email address to access the research materials, and the study was limited to third-year undergraduates. The email address login process was used to protect the results from duplicate submissions. It is possible that, in the opinion of some potential respondents, this represented a time-consuming extra step that caused them not to participate. Additionally, while there is the theoretical grounding for the isolation of third-year undergraduates, the lack of direct sampling access for this population subset hindered the attainment of a meaningful number of participants. In short, this attempt to meaningfully add to the research base ended up providing a shallow examination of the topic by going too narrow too fast.

Additionally, the use of snowball sampling proved problematic for this study. Not all methods professors have similar philosophies regarding undergraduate research, and at least one professor reported feeling uncomfortable asking students to respond to the research materials. To avoid putting professors in situations where they will have to make philosophical choices of this nature, it is recommended a more direct sampling method should be employed.

The GPAS lived up to its name by providing information related to seven extracted factors that each demonstrate an "acceptance" of a common grading practice. However, as it is currently constructed, the GPAS reported information about which hodgepodge participants preferred rather than their acceptance of standards-based grading. In short, this was a Grading Practices Attitude Survey, but perhaps not in the manner the original researchers had hoped.

Problems with the GPAS potentially point to the underlying complexity regarding the slow acceptance of SBG by the music education community. There are many facets to

75

SBG that must be more carefully considered when attempting to measure attitudes toward its use. The large number of factors extracted from the GPAS and the underrepresentation of participants who did not anticipate using SBG, especially among those who scored below the mean, point to a "cafeteria-style" approach to SBG where music teachers select the parts of SBG with which they agree and ignore the parts with which they disagree.

Recommendations for Future Research

The first recommendation for future study would involve a replication of the current study with a larger sample size. Recommendations for possible ways to accomplish this include removal of the need for a log-in to access the research materials, and the expansion of the sample to include more participants than just third-year undergraduates. It is further recommended that investigation of pre-service music teacher AGOs should start more broadly, followed by longitudinal studies to determine if the same AGO patterns exist among music educators.

The use of Achievement Goal Theory in music is an interesting topic that has only recently started to receive research attention. It is recommended that additional studies, which investigate musician AGO Profiles be conducted to determine if patterns such as those found in the present study are typical to the profession, or unique to this small sample.

SBG is a multi-faceted approach to grading and therefore potentially cannot be measured using only one scale. Future research endeavors could involve a reworking of the GPAS to measure these different facets in a more complete and reliable manner. Additionally, the wide range of scores and high number of affirmative Use_SBG responses point to an adopted "cafeteria-style" approach to SBG where users adopt the parts they like, reject the parts they do not, but still self-report using SBG. This version of SBG would look nothing like the SBG that has been previously recommended, and future research could explore the extent to which this is happening in music education.

It has been proposed that "a period of memory consolidation plays an important role in shaping attitude and may influence pre-service teachers as they move toward acceptance of SBG practices" (Wuttke & St. Pierre, 2016, pp. 420–421). Thus, is it recommended that future participants be asked to watch the video one-week prior to completing the GPAS and AGQ. This will allow for memory consolidation, which could solidify informed grading practice attitudes.

Finally, it is recommended that future research in this area of inquiry should be designed to predict which future teachers will use SBG in their classrooms. The GPAS correlated with Use_SBG, but more information is needed regarding the exact relationship between GPAS sub-scales, AGO Profile, and anticipated use of SBG. This will help music teacher-training programs to do everything they can to help future educators use appropriate grading schemes in their teaching careers.

Conclusion

The purpose of this study was to examine the interplay between achievement goal orientation (AGO) and attitudes toward grading practices to identify homogeneous groups of participants for further analysis. It was discovered that the unique combination

77

of (a) high mastery-based AGOs and (b) low performance-based AGOs provided the psychological predisposition best suited for the adoption of SBG.

This chapter has provided a detailed discussion of the research findings, described the immediate implications for both teacher-training programs and the music classroom, described its limitations, and suggested improvements. Recommendations for future research were provided, and new questions were posed to the field to grow the knowledge base as it relates to teacher assessment of student learning.

It has been documented for decades that teachers sometimes ignore the recommendations of assessment experts when creating grading policies. Teacher hesitance is not a new phenomenon, and this research has taken an important step in the field of music education to identify certain reasons for a teacher's hesitation to assess using achievement-only criteria.

MCB.3 The student will demonstrate vocal technique	es and choral skills
Performance-based score:	Things to Improve:
Exceeds the standard	 Using proper posture for choral singing;
Meets the standard	Using breathing techniques that support
Approaching the standard	vocal production;
Below the standard	Identifying vocal anatomy
□ Far Below the standard	Developing vocal independence
No attempt	Developing vocal agility and range
-	Using correct intonation
	Blending with other singers on the same
	vocal part
	Using proper diction (i.e., pure vowel
	sounds, diphthongs, and consonants)
	□ Singing an assigned vocal part in a small
	group
	Singing music literature from memory
	□ Following a choral octavo
MCB.4 The student will sing expressively	
Performance-based score:	Things to Improve:
Exceeds the standard	□ Singing correct tempo
Meets the standard	Performing correct rhythms
Approaching the standard	Interpreting dynamic markings
Below the standard	Demonstrating expressive phrasing
□ Far Below the standard	Responding to basic conducting patterns
No attempt	and interpretive gestures
	Using facial and physical expressions that
	reflect the mood and style of the music
MCB.6 The student will demonstrate collaboration a	nd concert etiquette as a performer
Performance-based score:	Things to Improve:
Exceeds the standard	Cooperating and collaborating as a singer
Meets the standard	during rehearsal
Approaching the standard	
Below the standard	
□ Far Below the standard	
No attempt	
MCB.1 The student will read and notate music	
Performance-based score:	Things to Improve:
Exceeds the standard	□ Identifying, defining, and using standard
Meets the standard	notation
Approaching the standard	Echoing, reading, and notating rhythmic
Below the standard	patterns
□ Far Below the standard	□ Sight-singing eight-measure melodic
No attempt	patterns, while maintaining a steady beat

Appendix A: An Example of a Standards-Based Grading Instrument

Appendix B: Achievement Goal Questionnaire

Achievement Goal Questionnaire (Finney et al., 2004)

Seven-point rating scale with 1 (not at all true of me) to 7 (very true of me)

Item

My goal this semester is to get better grades than most of the other students It is important for me to do well compared to other students this semester. I want to do better than other students this semester I just want to avoid doing poorly compared to other students this semester The fear of performing poorly is what motivates me My goal this semester is to avoid performing poorly compared to other students I am afraid that I may not understand the content of my courses as thoroughly as I'd like I worry that I may not learn all that I possibly could this semester I am definitely concerned that I may not learn all that I can this semester Completely mastering the material in my courses is important to me this semester I want to learn as much as possible this semester The most important thing for me this semester is to understand the content in my courses as thoroughly as possible

Appendix C: Grading Practice Attitude Survey

Grading Practice Attitude Survey (GPAS) (Wuttke & St. Pierre, 2016)

Item Number	Exact Wording
Item 1	Grading should focus on the material that has or has not been learned rather than on accumulating points to reach a certain average or total
Item 2	Other factors, such as effort and attitude are essential, but should not be indicative of the students who have the highest academic grades
Item 3	Both practice logs and music performance tests should be used to measure student proficiency
Item 4	Although musical preferences are subject to opinion, performance tests can be constructed so that they fairly and accurately measure student ability, achievement, and proficiency in music
Item 5	Recording scores by type, such as tests, homework, or performance tests in order to clearly identify the areas of concern for each student
Item 6	Grading student behavior is acceptable only when that grade is kept separate from the academic grade
Item 7	A student's academic grade is skewed by receiving extra credit for effort and does not reflect true musical achievement
Item 8	Grading that focuses on averaging all of the scores received on tests and other assignments indicates that the student has learned the material
Item 9	The intent for issuing an academic grade is to report how well a student has mastered what has been taught
Item 10	Students with good behavior that participate in class regularly will receive the highest academic grades
Item 11	Test results are final, there are no "do-overs" regardless of whether or not the student is happy with their grade

Ten-point rating scale with 0 (disagree) to 10 (agree) Item

Item Number	Exact Wording
	The best way to assess student musical achievement is by
Item 12	comparing what students know and can do with criteria they
	should be expected to master
Itom 13	Due to the subjective nature of the art form, student achievement
Item 15	in music is best measured on a case-by-case basis
	Assigning academic grades based upon judgments of character
Item 14	and teacher expectations is more often than not, a clear-cut
	representation of a student's abilities
	The purpose for assigning academic grades is to show the "whole
Item 15	picture" of the student including, but not limited to: effort,
	achievement, participation, attendance, compliance, and behavior
	Students who do not finish assignments should be issued a zero
Item 16	in the gradebook because it will teach the student to work harder
	on their next assignment
Item 17	Teachers should allow students to finish incomplete assignments
	so long as the initial assignment was submitted on time
	Grading "on a curve" is good practice because it is a more
Item 18	effective way to show parents and administrators where each
	student scored in comparison with each other
	After a low score on a performance test, the student should be
Item 19	allowed multiple opportunities to retake it to demonstrate
	mastery
	Extra credit that rewards students for their additional effort is
Item 20	sensible practice and provides a meaningful representation of
	musical achievement
	Basing academic grades on teacher expectations and character
Item 21	judgments may cause the students to perform in ways that do not
	accurately reflect their abilities
	Grades from music performance tests should be used to measure
Item 22	student proficiency; homework, (such as practice logs) should
	not
T	Because everyone has different opinions about musical likes and
Item 23	dislike, music performance tests are not a legitimate measure of
	student ability, achievement, and proficiency in music
	Based upon what I now know about Standards-Based Grading, I
Use SBG	plan on using this system in my own classroom when I teach
	even if it is not already required by the school district (Yes/No)

Appendix D: Consent Form

Personal Goal Orientation and Grading Practice Attitude in Pre-service Music Teachers: Predicting Anticipated Use of Standards-based Grading

INFORMED CONSENT FORM

RESEARCH PROCEDURES

This research is being conducted to examine music education majors' definitions of success as well as their attitudes toward grading to see if those aspects can predict how you will grade in your future teaching. More specifically, can they predict whether or not you will use standards-based grading? If you agree to participate, you are asked to (a) watch the video; (b) complete the survey. Watching the video will take about 11 minutes, and completing the survey should take no more than five minutes. Therefore, participation is anticipated to take approximately 16 minutes. The information that you provide is related to your ideas regarding success in academics and your attitudes toward grading and assessment. You will not be contacted again in the future. Information is completely anonymous.

To make sure that each participant completes the survey only once, it will be necessary to use an email address to login. However, this is not associated with your responses and all information is completely anonymous. A print copy of the survey can be made available if you prefer.

RISKS

There are no foreseeable risks for participating in this research.

Project Number: 908766-2

BENEFITS

There are no benefits to you as a participant other than to further research in helping to inform all music teacher training programs on how best to inform the future generation of music teachers about appropriate and authentic assessment and grading in the music classroom.

CONFIDENTIALITY

The data in this study will be confidential. Names and other identifiers will not be placed on surveys or other research data. While it is understood that no computer transmission can be perfectly secure, reasonable efforts will be made to protect the confidentiality of your transmission.

PARTICIPATION

Your participation is voluntary, and you may withdraw from the study at any time and for any reason. If you decide not to participate or if you withdraw from the study, there is no penalty or loss of benefits to which you are otherwise entitled. There are no costs to you or any other party. You must be 18 years of age or older to participate.

CONTACT

This research is being conducted Nathan St. Pierre, a doctoral candidate in music education at George Mason University. He may be reached via email at <u>nstpierr@gmu.edu</u>. His faculty advisor is Dr. Brian C. Wuttke (bwuttke@gmu.edu; 703-933-1381) (investigator's phone number). You may contact either for questions or to report a research-related problem. You may contact the George Mason University Office of Research Integrity & Assurance at 703-993-4121 if you have questions or comments regarding your rights as a participant in the research.

This research has been reviewed according to George Mason University procedures governing your participation in this research.

CONSENT

By clicking below, I agree that I have read this form, all of my questions have been answered by the research staff, and I agree to participate in this study.



Office of Research Integrity & Assurance IRB: For Official Use Only

Page 1 of 1

Appendix E: Recruitment Letter for Participants

Dear Future Teacher,

My name is Nathan St. Pierre and I am a doctoral candidate in music education at George Mason University. I am collecting data for research I am conducting as part of my dissertation and I would like to invite you to participate.

The purpose of this research project is to examine music education majors' definitions of success as well as their attitudes toward grading to see if those aspects can predict how you will grade in your future teaching. More specifically, can they predict whether or not you will use standards-based grading.

Included with this letter is a short video on standards-based grading and a brief survey. If you agree to participate, you are asked to (a) watch the video located at https://youtu.be/OaDIfcn7lfE; and then (b) complete the survey located at https://goo.gl/forms/5ntNqhoHHDpp7C4X2. Watching the video will take about 11 minutes, and completing the survey should take no more that 5 minutes. Therefore participation is anticipated to take approximately 16 minutes. The information that you provide is related to your ideas regarding success in academics and your attitudes toward grading and assessment. You will not be contacted again in the future. Information is completely anonymous.

To make sure that each participant completes the survey only once, it will be necessary to use an email address to login. However, this is not associated with your responses and all information is completely anonymous. A print copy of the survey can be made available if you prefer. The present study is one of minimal risk (i.e., not beyond that of daily life or which is typically encountered during academic tests). The biggest inconvenience you will experience will be time commitment.

Participation is voluntary and there are no negative consequences should you withdraw. There are no extrinsic benefits to you. However, by participating, you will be helping to inform all music teacher training programs on how best to inform the future generation of music teachers about appropriate and authentic assessment and grading in the music classroom.

If you have questions, I can be reached via email at mstpierr@gmu.edu. My faculty advisor is Dr. Brian C. Wuttke (bwuttke@gmu.edu; 703-993-1381). You may contact the George Mason University Office of Research Integrity & Assurance at 703-993-4121 if you have questions or comments regarding your rights as a participant in the research. Thank you for your help, and I hope that you will participate in this research.

With kind regards,

Nathan St. Pierre <u>nstpierr@gmu.edu</u> IRBNet number: 908766-1



Project Number: 908766-2 of Research Integrity & Assurance IRB: For Official Use Only

Page 1 of 1

Appendix F: IRB Decision Letter



Office of Research Development, Integrity, and Assurance

Research Hall, 4400 University Drive, MS 6D5, Fairfax, Virginia 22030 Phone: 703-993-5445; Fax: 703-993-9590

DATE:	March 10, 2017
TO:	Brian Wuttke
FROM:	George Mason University IRB
Project Title:	[908766-3] Personal Goal Orientation and Grading Practice Attitude in Pre- service Music Teachers: Predicting Anticipated Use of Standards-Based Grading
SUBMISSION TYPE:	Amendment/Modification
ACTION: DECISION DATE:	DETERMINATION OF EXEMPT STATUS March 10, 2017
REVIEW CATEGORY:	Exemption category #2

Thank you for your submission of Amendment/Modification materials for this project. The Institutional Review Board (IRB) Office has determined this project is EXEMPT FROM IRB REVIEW according to federal regulations.

Please remember that all research must be conducted as described in the submitted materials.

Please note that any revision to previously approved materials must be submitted to the IRB office prior to initiation. Please use the appropriate revision forms for this procedure.

If you have any questions, please contact Katie Brooks at (703) 993-4121 or kbrook14@gmu.edu. Please include your project title and reference number in all correspondence with this committee.

Please note that all research records must be retained for a minimum of five years, or as described in your submission, after the completion of the project.

Please note that department or other approvals may also be required to conduct your research.

GMU IRB Standard Operating Procedures can be found here: <u>http://oria.gmu.edu/1031-2/?</u> _ga=1.12722615.1443740248.1411130601

This letter has been electronically signed in accordance with all applicable regulations, and a copy is retained within George Mason University IRB's records.

- 1 -

Generated on IRBNe

References

- Ames, C. (1992). Classrooms: Goals, structures, and student motivation. Journal of Educational Psychology, 84(3), 261–271. https://doi.org/10.1037/0022-0663.84.3.261
- Ames, C., & Archer, J. (1988). Achievement goals in the classroom: Students' learning strategies and motivation processes. *Journal of Educational Psychology*, 80(3), 260–267. https://doi.org/10.1037/0022-0663.80.3.260
- Anderson, D. B., Anderson, A. L. H., Mehrens, W., & Prawat, R. S. (1990). Stability of educational goal orientations held by teachers. *Teaching and Teacher Education*, 6(4), 327–335. https://doi.org/10.1016/0742-051X(90)90025-Z
- Asmus, E. P. (1986). Student beliefs about the causes of success and failure in music: A study of achievement motivation. *Journal of Research in Music Education*, *34*(4), 262–278. https://doi.org/10.2307/3345260
- Asmus, E. P. (1994). Motivation in music teaching and learning. *The Quarterly Journal of Music Teaching and Learning*, *5*(4), 5–32.
- Asmus, E. P. (1999). Music assessment concepts. *Music Educators Journal*, 86(2), 19–24. https://doi.org/10.2307/3399585

Bae, H., & Lee, Y. (2015). Do the causal effects of internal collaboration and external collaboration improve performance in supply chains? *Journal of International Logistics and Trade*, 13(1), 73–97. https://doi.org/10.24006/jilt.2015.13.1.004

Baranik, L. E., Barron, K. E., & Finney, S. J. (2007). Measuring goal orientation in a work domain: Construct validity evidence for the 2 x 2 framework. *Educational* and Psychological Measurement, 67(4), 697–718. https://doi.org/10.1177/0013164406292090

- Bonner, S. M., & Chen, P. P. (2009). Teacher candidates' perceptions about grading and constructivist teaching. *Educational Assessment*, 14(2), 57–77. https://doi.org/10.1080/10627190903039411
- Brookhart, S. M. (1991). Grading practices and validity. *Educational Measurement: Issues and Practice*, *10*(1), 35–36. https://doi.org/10.1111/j.1745-3992.1991.tb00182.x
- Brookhart, S. M. (1994). Teachers' grading: Practice and theory. *Applied Measurement in Education*, 7(4), 279–301. https://doi.org/10.1207/s15324818ame0704_2
- Brophy, T. S. (1997). Reporting progress with developmental profiles. *Music Educators Journal*, 84(1), 24–27. https://doi.org/10.2307/3399082
- Brunson, B. I., & Matthews, K. A. (1981). The Type A coronary-prone behavior pattern and reactions to uncontrollable stress: An analysis of performance strategies, affect, and attributions during failure. *Journal of Personality and Social Psychology*, 40(5), 906–918. https://doi.org/10.1037/0022-3514.40.5.906

Buluş, M. (2011). Goal orientations, locus of control and academic achievement in prospective teachers: An individual differences perspective. *Educational Sciences: Theory and Practice*, 11(2), 540–546.

Butler, R. (2007). Teachers' achievement goal orientations and associations with teachers' help seeking: Examination of a novel approach to teacher motivation. *Journal of Educational Psychology*, 99(2), 241–252. https://doi.org/10.1037/0022-0663.99.2.241

Butler, R., & Shibaz, L. (2008). Achievement goals for teaching as predictors of students' perceptions of instructional practices and students' help seeking and cheating. *Learning and Instruction*, 18(5), 453–467.

https://doi.org/10.1016/j.learninstruc.2008.06.004

- Campbell, C., & Evans, J. A. (2000). Investigation of preservice teachers' classroom assessment practices during student teaching. *The Journal of Educational Research*, 93(6), 350–355. https://doi.org/10.1080/00220670009598729
- Carlton, J. P. (1992). A case study analysis of parents', teachers', and students' perceptions of the meaning of grades: Identification of discrepancies, their consequences, and obstacles to their resolution (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9306051).
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). Applied multiple regression/correlation analysis for the behavioral sciences. (3rd ed.). Hillsdale, NJ: Lawrence Erlbaum.

- Colwell, R. (1998). Preparing student teachers in assessment. *Arts Education Policy Review*, 99(4), 29–36. https://doi.org/10.1080/10632919809600780
- Consortium of National Arts Education Associations. (1994). *National standards for arts* education: Dance, music, theatre, visual arts. Reston, VA: MENC.
- Creswell, J. W. (2011). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (4th ed.). Boston, MA: Pearson.
- Cross, L. H., & Frary, R. B. (1999). Hodgepodge grading: Endorsed by students and teachers alike. *Applied Measurement in Education*, 12(1), 53–72. https://doi.org/10.1207/s15324818ame1201_4
- Droe, K. L. (2013). Effect of verbal praise on achievement goal orientation, motivation, and performance attribution. *Journal of Music Teacher Education*, 23(1), 63–78. https://doi.org/10.1177/1057083712458592
- Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychologist*, *41*(10), 1040–1048. https://doi.org/10.1037/0003-066X.41.10.1040
- Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review*, 95(2), 256–273. https://doi.org/10.1037/0033-295X.95.2.256
- Eison, J. A. (1979). The development and validation of a scale to assess differing student orientations towards grades and learning (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8005386).

- Elliot, A. J., & Church, M. A. (1997). A hierarchical model of approach and avoidance achievement motivation. *Journal of Personality and Social Psychology*, 72(1), 218–232. https://doi.org/10.1037/0022-3514.72.1.218
- Elliot, A. J., & Harackiewicz, J. M. (1996). Approach and avoidance achievement goals and intrinsic motivation: A mediational analysis. *Journal of Personality and Social Psychology*, 70(3), 461–475. https://doi.org/10.1037/0022-3514.70.3.461
- Elliot, A. J., & McGregor, H. A. (1999). Test anxiety and the hierarchical model of approach and avoidance achievement motivation. *Journal of Personality and Social Psychology*, 76(4), 628–644. https://doi.org/10.1037/0022-3514.76.4.628
- Elliot, A. J., & McGregor, H. A. (2001). A 2 × 2 achievement goal framework. *Journal of Personality and Social Psychology*, 80(3), 501–519. https://doi.org/10.1037/0022-3514.80.3.501
- Elliott, E. S., & Dweck, C. S. (1988). Goals: An approach to motivation and achievement. *Journal of Personality and Social Psychology*, 54(1), 5–12. https://doi.org/10.1037/0022-3514.54.1.5
- Emerson, R. W. (2015). Convenience sampling, random sampling, and snowball sampling: How does sampling affect the validity of research? *Journal Of Visual Impairment & Blindness*, 109(2), 164–168.
- Ester, D. P. (1997). Choral methods and the CMP model. *Journal of Music Teacher Education*, 6(2), 26–28. https://doi.org/10.1177/105708379700600209
- Evans, J. D. (1996). *Straightforward statistics for the behavioral sciences*. Brooks/Cole Publishing Company.

- Finney, S. J., Pieper, S. L., & Barron, K. E. (2004). Examining the psychometric properties of the achievement goal questionnaire in a general academic context. *Educational and Psychological Measurement*, 64(2), 365–382. https://doi.org/10.1177/0013164403258465
- Fonseca, J. R. S. (2013). Clustering in the field of social sciences: that is your choice. *International Journal of Social Research Methodology*, 16(5), 403–428. https://doi.org/10.1080/13645579.2012.716973
- Frary, R. B., Cross, L. H., & Weber, L. J. (1993). Testing and grading practices and opinions of secondary teachers of academic subjects: Implications for instruction in measurement. *Educational Measurement: Issues and Practice*, *12*(3), 23–30. https://doi.org/10.1111/j.1745-3992.1993.tb00539.x
- Graham, S., & Golan, S. (1991). Motivational influences on cognition: Task involvement, ego involvement, and depth of information processing. *Journal of Educational Psychology*, *83*(2), 187–194. https://doi.org/10.1037/0022-0663.83.2.187
- Harackiewicz, J. M., Barron, K. E., Carter, S. M., Lehto, A. T., & Elliot, A. J. (1997). Predictors and consequences of achievement goals in the college classroom:

Maintaining interest and making the grade. *Journal of Personality and Social Psychology*, 73(6), 1284–1295. https://doi.org/10.1037/0022-3514.73.6.1284

- Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, 39(1), 31–36. https://doi.org/10.1007/BF02291575
- Kent, R. A. (2015). Analysing quantitative data: Variable-based and case-based approaches to non-experimental datasets. SAGE Publications Ltd.
- Kim, H.-Y. (2013). Statistical notes for clinical researchers: assessing normal distribution
 (2) using skewness and kurtosis. *Restorative Dentistry & Endodontics*, 38(1), 52–54. https://doi.org/10.5395/rde.2013.38.1.52
- Kotora, E. J. (2005). Assessment practices in the choral music classroom: A survey of Ohio high school choral music teachers and college choral methods professors. *Contributions to Music Education*, 32(2), 65–80.
- Leedy, P. D., & Ormrod, J. E. (2009). *Practical research: Planning and design* (9th ed.). Pearson.
- Lehman, P. R. (1998). Grading practices in music. *Music Educators Journal*, 84(5), 37–40. https://doi.org/10.2307/3399129
- Maehr, M. L., Pintrich, P. R., & Linnenbrink, E. A. (2002). Motivation and achievement.
 In R. Colwell & C. P. Richardson (Eds.), *The new handbook of research on music teaching and learning* (pp. 348–372). New York, NY: Oxford University Press.
- Malmberg, L.-E. (2006). Goal-orientation and teacher motivation among teacher applicants and student teachers. *Teaching and Teacher Education*, 22(1), 58–76. https://doi.org/10.1016/j.tate.2005.07.005

Malmberg, L.-E. (2008). Student teachers' achievement goal orientations during teacher studies: Antecedents, correlates and outcomes. *Learning and Instruction*, 18(5), 438–452. https://doi.org/10.1016/j.learninstruc.2008.06.003

Marburger, D. R. (2006). Does mandatory attendance improve student performance? *The Journal of Economic Education*, 37(2), 148–155. https://doi.org/10.3200/JECE.37.2.148-155

- Marlatt, J. H. (2004). Attribution theory as an influence on the musical lives of children (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3125542).
- McCoy, C. W. (1991). Grading students in performing groups: A comparison of principals' recommendations with directors' practices. *Journal of Research in Music Education*, 39(3), 181. https://doi.org/10.2307/3344718
- McDonald, J. H. (2014). *Handbook of biological statistics* (3rd ed.). Baltimore, MD: Sparky House Publishing.
- Meyers, L. S., Gamst, G. C., & Guarino, A. J. (2013). *Applied multivariate research:Design and interpretation* (2nd ed.). Los Angeles, CA: SAGE Publications, Inc.
- Middleton, M. J., & Midgley, C. (1997). Avoiding the demonstration of lack of ability:
 An underexplored aspect of goal theory. *Journal of Educational Psychology*, 89(4), 710–718. https://doi.org/10.1037/0022-0663.89.4.710
- Midgley, C. (Ed.). (2002). *Goals, goal structures, and patterns of adaptive learning*. Mahwah, N.J: Routledge.

- Midgley, C., Kaplan, A., & Middleton, M. (2001). Performance-approach goals: Good for what, for whom, under what circumstances, and at what cost? *Journal of Educational Psychology*, 93(1), 77–86. https://doi.org/10.1037/0022-0663.93.1.77
- Midgley, C., Kaplan, A., Middleton, M., Maehr, M. L., Urdan, T., Anderman, L. H., ...
 Roeser, R. (1998). The development and validation of scales assessing students' achievement goal orientations. *Contemporary Educational Psychology*, 23(2), 113–131. https://doi.org/10.1006/ceps.1998.0965
- Midgley, C., Maehr, M. L., Hruda, L. Z., Anderman, E., Anderman, L., Freeman, K. E.,
 ... Urdan, T. (2000). *Manual for the patterns of adaptive learning scales*. Ann
 Arbor, MI: University of Michigan.
- Miksza, P. (2009). An investigation of the 2x2 achievement goal framework in the context of instrumental music. In *Research Perspectives: Thought and Practice in Music Education* (pp. 81–100). Charlotte, NC: Information Age Publishing, Inc.
- Moore, R. (2005). Attendance: Are Penalties More Effective Than Rewards? *Journal of Developmental Education*, 29(2), 26–32.
- Norušis, M. (2012). *IBM SPSS Statistics 19 statistical procedures companion*. Upper Saddle River, NJ: Addison Wesley.

 Odoom, R. (2016). Brand-building efforts in high and low performing small and medium-sized enterprises (SMEs). *Journal of Small Business and Enterprise Development*, 23(4), 1229–1246. https://doi.org/10.1108/JSBED-05-2016-0067

- Pedersen, S., & Williams, D. (2004). A comparison of assessment practices and their effects on learning and motivation in a student-centered learning environment. *Journal of Educational Multimedia and Hypermedia*, 13(3), 283–306.
- Pintrich, P. R. (2000). Multiple goals, multiple pathways: The role of goal orientation in learning and achievement. *Journal of Educational Psychology*, 92(3), 544–555. https://doi.org/10.1037/0022-0663.92.3.544
- Richards, J. (2014). An old chestnut revisited: Teachers' opinions and attitudes toward grading within a competency based training framework. *International Journal of Training Research*, *12*(3), 182–191.

https://doi.org/10.1080/14480220.2014.11082040

- Russell, J. A., & Austin, J. R. (2010). Assessment practices of secondary music teachers. Journal of Research in Music Education, 58(1), 37–54. https://doi.org/10.1177/0022429409360062
- Sandene, B. A. (1997). An investigation of variables related to student motivation in instrumental music (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9811178).
- Schraw, G., & Aplin, B. (1998). Teacher preferences for mastery-oriented students. *The Journal of Educational Research*, 91(4), 215–221. https://doi.org/10.1080/00220679809597546
- St. Pierre, N. A., & Wuttke, B. C. (2015). Standards-based grading practices among practicing music educators: Prevalence and rationale. *Update: Applications of Research in Music Education*. https://doi.org/10.1177/8755123315604468
- Throndsen, I., & Turmo, A. (2013). Primary mathematics teachers' goal orientations and student achievement. *Instructional Science*, 41(2), 307–322. https://doi.org/10.1007/s11251-012-9229-2
- West, C. (2013). Motivating music students: A review of the literature. Update: Applications of Research in Music Education, 31(2), 11–19. https://doi.org/10.1177/8755123312473611
- Wuttke, B. C., & St. Pierre, N. A. (2016). The effects of instruction and experience on preservice teacher attitudes toward accepting a standards-based grading paradigm.
 In Timothy S. Brophy, J. Marlatt, & G. K. Ritcher (Eds.), *Connecting practice, measurement, and evaluation: selected papers from the Fifth International Symposium on Assessment in Music Education* (pp. 415–425). Chicago, IL: GIA Publications, Inc.

Biography

Nathan A. St. Pierre is the chorus teacher at Ronald Reagan Middle School in Haymarket, VA. He holds a Master of Education Degree in Kodàly Music Education from Loyola University Maryland as well as a Bachelor of Music Degree from The Hartt School, University of Hartford.