

THE FLORIDA STATE UNIVERSITY
COLLEGE OF MUSIC

ASSESSMENT AND GRADING IN THE
BEGINNING BAND CLASSROOM

by

MICHAEL D. ANTMANN

A Thesis submitted to the
College of Music
in partial fulfillment of the
requirements for the degree of
Master of Music Education

Degree Awarded:
Summer Semester, 2007

The members of the Committee approve the thesis of Michael Antmann
on July 3, 2007.

Steven Kelly
Professor Directing Thesis

Patrick Dunnigan
Committee Member

Kimberly VanWeelden
Committee Member

The Office of Graduate Studies has verified and approved the above named
committee members.

TABLE OF CONTENTS

List of Tables.....	iv
Abstract.....	v
1. INTRODUCTION	1
2. REVIEW OF LITERATURE	5
Assessment in Education	5
Assessment in Music	7
Assigning Grades to Music Students	14
Assessment in Band	15
Summary	17
3. METHOD	19
4. RESULTS	21
5. DISCUSSION	36
APPENDICES	43
REFERENCES	53
BIOGRAPHICAL SKETCH	56

LIST OF TABLES

1. Mean, Median, and Mode Scores for Question 1	24
2. Chi-Square Analysis of Question 1	25
3. Mean, Median, and Mode Scores for Question 2	26
4. Chi-Square Analysis of Question 2	27
5. Mean, Median, and Mode Scores for Question 3	28
6. Chi-Square Analysis of Question 3	29
7. Mean, Median, and Mode Scores for Question 4	30
8. Chi-Square Analysis of Question 4	31
9. Frequency of Responses for Question 1	32
10. Frequency of Responses for Question 2	33
11. Frequency of Responses for Question 3	34
12. Frequency of Responses for Question 4	35

ABSTRACT

While much attention has been given to the need for assessment and accountability in education, little research has been done showing how to assess and grade music students effectively. There is very little research or information about the grading and assessment in the beginning band, a stage that might possibly be the most important concerning instrumental music. The study sought to determine how and what successful middle school band directors are assessing in beginning band, and how that information is being used when grading students. Surveys were sent to band directors at middle schools with “successful” middle school band programs. The survey asked the participants to rate how often they used various assessment tools, how often they used various assessment tools in determining a student’s grade, how important various categories are in a performance assessment, and how often various musical skills are assessed in a beginning band class. For this study, “successful” middle school band programs were defined as band programs who have received a “superior” rating at the Florida Bandmasters Association district concert band music performance assessment with at least one band each of the last four years ($N = 59$). Surveys were returned by 27 directors (45.8%). Results found that performing on instruments and reading/notating music were the most commonly assessed skills by the participants. The most commonly used assessment tool was found to be the individual playing test. The participants rated posture and air support as the two most important criteria in a beginning band student performance assessment. Finally, the participants rated playing tests, concert attendance, conduct/discipline, and participation as the most important factors in determining a student’s final grade.

CHAPTER 1 INTRODUCTION

The *No Child Left Behind* Act of 2001 and increasing state education requirements has led to increasing emphasis on assessment and accountability in recent years (U.S. Department of Education, 2004a). Accountability in determining teacher effectiveness and student learning has become a major element in education requirements. Consequently, it is becoming increasingly important for music teachers to accurately assess student learning. Assessment policies in many schools and school districts require all teachers to document student achievement through various forms, as well have plans to assign grades to these students in a fair and consistent manner (Duerksen, 1995). Schools use these grades to keep parents informed of students' progress by sending home report cards and progress reports throughout the year. Furthermore, many schools are now even posting student grades online so parents can constantly track their child's progress at school.

Many school improvement models include the identification of learning goals or standards as well as the frequent and ongoing assessment of those goals and standards (Duerksen, 1995). In addition, some state education agencies require "data-based school improvement programs" be in place, and that assessment data be used through the education system (Duerksen, 1995). In Florida, music programs use the Sunshine State Standards, which are based on the National Standards for music, to guide instruction and assessment of achievement. The Florida Music Educators Association is currently piloting a standardized test for the music Sunshine State Standards for fourth grade students (Orr, 2007). In 2007, this test was given to approximately 10% of fourth grade students in Florida. These school improvement models, when applied to music programs, present valuable ways to improve student learning and teacher effectiveness.

The *National Standards for Arts Education* suggest what every music student should be able to do, and is the basis for many state and local music

standards (MENC, n.d.). These standards are the foundation of the Florida Sunshine State Standards for music, which provide specific benchmarks that students should achieve for each grade level (MENC, n.d.). While these benchmarks are a valuable tool for teachers in planning curriculum and instruction, they also show the need for assessing student achievement. However, there is no complete standardized assessment tool for music, which prevents school districts from comparing music achievement nationally (LeCroy, 1999). Other subject areas have standardized tests and assessments, which further demonstrates the need for assessment in music education. Resources and funding are given to courses with standardized assessments (such as the Florida Comprehensive Achievement Test), and these courses often become the focus of the school day. Providing a fair and accurate assessment of a music student's mastery of the standards for music is essential to the argument for music education in a school curriculum.

In 1997, the National Assessment of Educational Progress (NAEP) assessed eighth-grade music students from throughout the United States on their ability to perform, create, and respond to music (Schneider, 2005). The NAEP music assessments do not provide individual students scores; the purpose of this test was to provide a broad, national overview of student achievement (Schneider 2005). Data from the NAEP showed that students' involvement in music was positively related to student musical achievement (White & Vanneman, 2000). While a positive outcome, this test only produced limited, baseline data (LeCroy, 1999).

The *No Child Left Behind Act* (NCLB) of 2001 mandates that fourth and eighth graders should be assessed in reading and math every other year (Schneider, 2005). In addition to mandatory assessments in math and reading, NCLB also permits voluntary assessment in other areas, including the arts (Schneider, 2005). By including music in standardized testing, this may provide opportunities for additional funding for music or justification for music in the curriculum.

The *No Child Left Behind Act* of 2001 puts a priority on assessment and accountability. “*No Child Left Behind* holds schools and school districts accountable for results. Schools are responsible for making sure each child is learning (U.S. Department of Education, 2004a).” According to the legislation, school systems should report on educational indicators “to better empower parents, educators, administrators, and schools to effectively address the needs of their children and students (U.S. Department of Education, 2004a).” Many states have developed standards for different subject areas, and these standards are what are being assessed in mandated standardized tests. In many cases, funding is tied to a school’s performance on these tests. While national standards, as well as local standards in many states have been established, in most cases standards are assessed by any form of test, making it impossible to compare school districts.

Some research has investigated the assessment and grading of music students. Lehman (1998) conducted an informal study showed that while most teachers assign grades (mostly) on performance-based criteria; some teachers included some non-performance factors. These factors included attendance, behavior, effort and attitude. The specificity of grading criteria also varied greatly. Lehman suggested music teachers use standards-based grading criteria. The distribution of grades in a music class using this system would differ from a math or history class, because music is an elective class; students who take music classes have an interest, and usually prior study in music, so average scores may be higher. Non-music criteria such as attendance, behavior, and attitude are not compatible with this form of grading. Furthermore, for grading to be most effective, the grade should only include the assessment of skills being taught in the class. For example, using the grade to punish a student for attendance or behavior may weaken the validity of a grade (Oosterhof, 2001). Also, grading effort, attitude, and behavior can be highly subjective, and difficult to assign a grade or point value.

There has been very little research concerning the assessment and grading of beginning band students. Assessment in beginning band is critical as

this is the time when fundamentals of band and music performance are developed. Assessment is a powerful tool that guides instruction, and if used effectively, can improve student learning. Furthermore, if national and state standards establish what a student should learn, then it is logical to assume that a student's grade should be representative of whether or not the student acquired those skills. According to a position paper released by MENC, "MENC: The National Association for Music Education recognizes that assessment, and the accountability that stems from the public dissemination of the results of assessment, are key components in building quality instructional programs (MENC, 2007)." Yet, little research on assessment in music programs, especially beginning bands, exist.

This study attempted to address MENC's concern for effective assessment in music classes. Specifically, the study sought to determine the grading and assessment practices and policies used in beginning bands by teachers of successful middle school band programs. This study addressed the follow questions in regards to assessment and grading in the beginning band class:

1. *What is being assessed in a beginning band in a successful middle school band program?*
2. *How are skills and knowledge being assessed in beginning band by directors of successful middle school band programs?*
3. *How is assessment being used to assign grades to students in beginning bands of successful middle school band programs?*

CHAPTER TWO RELATED LITERATURE

Assessment in Education

Assessment should serve as a useful and essential tool in the classroom. It can be used to evaluate student progress, set standards, guide instruction, and to communicate student progress to parents and administrators (Farrell, 1997). Assessment is most effective when it represents student outcomes that aligned with the goals of classroom instruction (Farrell, 1997).

For assessment to be effective it is first important to identify what information or skill is to be learned (Cope, 1996). Setting clear learning goals and objectives makes the job of assessment easier, and the assessment itself more effective. Accurate assessment helps provide better information for the student and teacher, as it helps the teacher evaluate past instruction and guide future instruction (Asmus, 1999). Before discussing assessment further, it is important to draw a distinction between measurement and assessment. Measurement simply measures an observable student performance or behavior (Oosterhof, 2001). It does not apply any knowledge a teacher may have. Assessment is accomplished by rating or judging someone or something based on a set of criteria or standards (Radocy, 1995), and requires prior knowledge of the person who is the assessor (Oosterhof, 2001)

Assessment is an essential part of effective classroom instruction. There are four types of assessment that are used: preliminary (or “placement”), diagnostic, formative, and summative (Goolsby, 1999). Preliminary assessment helps determine a student’s abilities or prior knowledge. Diagnostic assessment determines difficulties or roadblocks a student may have they may prevent learning. Formative assessment occurs during instruction and can indicate whether or not students or learning. Summative assessment involves the final product (concerts, exams, etc...) (Goolsby, 1999).

Preliminary assessment involves finding out what a student already knows, or what a student's prior abilities are. Sometimes referred to as "Placement" assessment, this can include band auditions, seating tests, or pre-tests (Goolsby, 1999). This type of assessment is used to give a starting point for student instruction.

Diagnostic assessment occurs throughout instruction. This type of assessment helps determine where learning difficulties exist (Goolsby, 1999). Through diagnostic assessment, a teacher can find roadblocks or issues they may prevent a student from learning. By fixing these problems, learning becomes easier for the student.

Formative assessment refers to assessment that should be happening constantly during instruction to guide future instruction (Goolsby, 1999). Formative assessment is normally informal, and can be accomplished in many ways. This can include monitoring student participation during class (Oosterhof, 2001).

Summative assessment refers to the process that usually assesses the final product (Goolsby, 1999). Summative assessments can include concerts, festivals, and other performances. Summative assessments also include standardized tests and final exams. These assessments can take place at the end of a year, semester, unit of instruction, etc... Summative assessment provides the basis for most grading systems (Oosterhof, 2001).

A checklist can be an effective form of summative assessment (Goolsby 1999). The checklist could include a list of skills to demonstrate during a period of time. Checklists can be used by the director to meet ensemble and individual goals. This can also be used as a tool to ensure students meet the national (or state) standards. Items can be placed on the checklist to meet each standard; these items do not need to be limited to playing/performance items. Music theory and writing assignments can be included.

In many assessment models, priority is placed on the assessment of knowledge and skills in real-life situations (Duerksen, 1995). This is referred to as authentic assessment. *Authentic assessment* occurs when behaviors are

being assessed in real-life applications (Radocy, 1995). Authentic assessment is very common in music due to the frequent use of performance assessments. While the type of assessment a teacher uses is important, the interpretation of the assessment is equally as important.

There are different ways to interpret assessments; the desired frame of reference will determine the interpretation to be used. Assessment interpretations include ability-referenced, growth-referenced, norm-referenced, and criterion-referenced (Oosterhof, 2001). Ability-referenced interpretation compares a student's performance to his/her maximum possible performance. Growth-referenced interpretation compares a student's performance to his/her previous performance (Oosterhof, 2001). Norm-referenced interpretation compares a student's performance to the performance of others (Oosterhof, 2001). Finally, criterion-referenced performance tells what a student can and can not do as demonstrated in his/her performance (Oosterhof, 2001).

Assessment in Music

Objective assessment in music is a difficult, if not an impossible task (Radocy 1995). Furthermore, according to Radocy, "there is no such thing as objective assessment of human endeavors." Measurements *can* be objective; Observing that a student demonstrated dynamic contrast is a measurement, and is objective (ex. *The student did or did not make a difference between the forte and piano dynamics*). However, the assessment of the same performance is subjective (ex. *The dynamic contrast demonstrated was or was not appropriate for the piece of music being played.*).

Individual assessment in the music classroom can also be a difficult task. Music classroom typically have larger student to teacher ratios, and assessment often must be completed one student at a time. This can cause a significant problem for a music teacher trying to assess individual students. Furthermore, performance assessments must be done in real-time, so students must be assessed in live performance, or on recording. It would be optimal for band students to meet in both large ensemble settings and small-group lessons, and

for music teachers to be provided with recording equipment (MENC, 2007).

The complex nature of music creates significant challenges for anyone who tries to assess it with any sort of specificity (Holahan, 1997). It is common, in the instructional setting, for music teachers to perform formative assessments of specific aspects of a student's performance (Holahan, 1997). Judges are often asked to assess musical performance using prepared rating instruments, which often times require the judges to "indicate levels of performance ability based solely on their personal judgment of quality and character" (Holahan, 1997). These rating scales, however, often do not provide specific information about the student performance (Holahan, 1997). Criteria-specific rating scales provide the assessor with specific statements about the students performance, and the assessor can choose the statement which best describes the student performance (Holahan 1997).

Music teachers have always used various assessment tools to evaluate aspects of the music program, and to make assessment results public. Large ensemble performance assessments are very common; in most cases, ensembles are evaluated using a set of performance criteria, and are usually assigned a rating. In Florida, these are referred to as "Music Performance Assessments." The Florida Bandmasters Association lists the following as the purpose for these assessments: "To provide opportunities for students and directors to perform in an environment which provides critical evaluation of its performance by noted experts in the field of band performance, to provide the opportunity for students and directors to perform for their peers in a formal concert setting, to provide a performance opportunity which will serve as a motivational goal for students and directors, to provide an opportunity for students and directors to hear performances of their peers and learn from hearing those performances, and to provide a goal which is so compelling that the preparation for attaining that goal becomes the vehicle for continued growth and to demonstrate students' abilities to apply musical fundamentals and concepts in an ensemble performance setting

(Florida Bandmasters Association, 2005).”

While the large-ensemble assessments can be an effective tool for evaluating the teacher and music program as a whole, it can be limited as far as providing specific information about individual students. This type of assessment can also be limited as it only provides assessment of certain aspects of music education (MENC, 2007). For assessment to be most reliable, it is important to use several different assessment tools (Oosterhof, 2001). It is also important for these assessments to cover all aspects of music education that are covered in the National Standards. In Florida, statewide assessments (standardized tests) are being piloted for use in the future (Orr, 2007). It is possible that the use of a standardized test for music would cause funding problems. This system would require additional funding for the standardized assessment, which may take away from funding used for the delivery of instruction. Without a standardized test, however, funding for music may be reallocated to subject areas that are currently being assessed. According to MENC, “some form of regular assessment of music programs should be adopted. The assessment should measure student learning across a range of standards representative of quality, balanced music curriculum, including not only responding to music but also creating and performing music. This assessment should serve the goal of educational accountability by providing data that can be included in the school- or district-level “report card” disseminated to the public as required by law (MENC, 2007).”

MENC provides the following guidelines for assessment:

1. *Assessment should be standards-based and should reflect the music skills and knowledge that are most important for students to learn.*

Assessment of student achievement should not be based on the skills and knowledge that are easiest to assess nor on those for which ready-made assessment devices are available. Instead, it should be based on the extent to which each student has met the standards established, and it should reflect the priorities of the instructional program. Assessment should not be based primarily on where the student ranks relative to a

particular class or group. It should be based on whether or not the student has met specific criteria. In these performance standards separate criteria have been established for basic, proficient, and advanced levels of achievement.

2. Assessment should support, enhance, and reinforce learning.

Assessment should be viewed by both students and teachers as a continuing, integral part of instruction rather than as an intrusion into--or interruption of--the process of learning. The assessment process should itself be a learning experience, and it should not be conducted or viewed as separate from the learning process. Students should regard assessment as a useful tool rather than as a source of fear or anxiety. They should use it as a means of further learning and as a means of measuring their own progress. When assessment tasks are designed to provide information concerning the extent to which students meet standards that have been established for them, teachers can adjust their instructional programs so as to be more effective.

3. Assessment should be reliable. Reliability refers to consistency. If an assessment is reliable, then another assessment of the same skills or knowledge will produce essentially the same results. For assessment to be reliable, every student must be assessed by identical procedures and the assessors must share the same levels of expectation so that a student's score does not depend on who is doing the scoring.

4. Assessment should be valid. Validity means that the assessment technique actually measures what it claims to measure. The mental processes represented by the scores correspond to the mental processes being assessed. No measurement instrument should be used to measure something that it was not designed to measure. If there is a mismatch between assessment strategies and the objectives of the curriculum, the assessment strategies are not valid for that curriculum.

5. *Assessment should be authentic.* Authentic assessment means that assessment tasks reflect the essential nature of the skill or knowledge being assessed. The student should actually demonstrate a music behavior in an authentic or realistic situation rather than merely answer written questions about it. For example, the ability to play the recorder should be assessed by having the student play the recorder, not by having the student answer test questions concerning fingerings, hand position, phrasing, and note-reading. Assessment does not need to be based on multiple-choice tests or even on paper-and-pencil tests, though those techniques have their uses. Portfolios, performance-based assessment, and other techniques of authentic assessment have been used successfully by music educators for many years; however, these techniques cannot by themselves solve the assessment problems facing educators. A portfolio is simply a collection of samples of a student's work taken periodically for a specific purpose throughout the instructional process. Those samples must still be assessed, and the assessment requires not only careful thought about what should go into the portfolio but also great care in developing suitable assessment strategies and appropriate scoring procedures. Assessment should take a holistic view of music learning. It should not concentrate on isolated facts and minutiae but should deal with broad concepts, "whole" performances, and complete works of music. Authenticity, like reliability, is a prerequisite to validity.

6. *The process of assessment should be open to review by interested parties.* Although assessment of music learning can best be carried out by qualified music teachers, it is important that students, parents, and the public be provided with sufficient information and help that they too can make judgments about the extent to which music learning is taking place in their schools. If their evaluations are faulty, it should be because of their lack of professional qualifications and not because of lack of information

concerning the assessment process. It is especially important that students know what they are to be assessed on, how they are to be assessed, and what criteria will be used to judge their achievement. When appropriate, they should be allowed to participate in developing the criteria by which their work will be assessed (MENC, n.d.).

One attempt at having a standardized music test was the National Assessment of Educational Progress (NAEP) Arts Assessment in 1997. The students participating in the NAEP assessment were assessed in their ability to create, perform, and respond (White & Vanneman, 2000). Creating referred to expressing feelings and thoughts through original music, much like improvisation. Performing referred to the performance of existing music, which showed the interpretive skills of the student. Responding refers to the student's ability to evaluate and analyze music (White & Vanneman, 2000). This test required the student to both write and sing or play instruments. The results of this test were basic, and found that female students performed better than male students, and in most cases, white students performed better than black and Hispanic students. The 1997 NAEP Arts Assessment also found that students who studied a musical instrument generally performed better on the music assessment than students who only received general music instruction (White & Vanneman, 2000).

Music teachers regularly make judgments about student performance and provide feedback to for the student to improve (Holahan, 1997). Assessment instruments that provide overall ratings to student or ensemble performance have been found to be reliable among different judges; however, these often do provide specific information about the performance (Holahan, 1997). Criteria-specific music performance rating scales provide a more accurate description of student performance (Holahan, 1997). These assessment tools ask the assessor to describe what is being played, and then assigns a score using a rubric (Holahan, 1997).

Performance assessments are a popular tool for music classrooms. Performance assessments are usually considered "authentic" assessments,

because they are assessing skills in a real-life setting (Oosterhof, 2001). When a band goes to festival, they are participating in a performance assessment; the band is performing a concert that is being assessed by adjudicators. This form of assessment can be equally as effective when applied to individual performance. One way to accomplish this is to have the student video or audio-tape themselves performing required exercises (Goolsby, 1999). It is important to have specific criteria for the assessment of these tapes.

The National Standards provide specific skills that every music student should acquire during their education. The national standards for music education were developed in 1994, and were a result of the *Goals 2000: Educate America Act*, passed by Congress, which identified the arts as a discipline in that American students should demonstrate competence (Byo, 1999). The National Standards are as follows: 1. Singing, alone and with others, a varied repertoire of music. 2. Performing on instruments, alone and with others, a varied repertoire of music. 3. Improvising melodies, variations, and accompaniments. 4. Composing and arranging music within specified guidelines. 5. Reading and notating music. 6. Listening to, analyzing, and describing music. 7. Evaluating music and music performances. 8. Understanding relationships between music, the other arts, and disciplines outside the arts. 9. Understanding music in relation to history and culture (MENC, 1994). While there is no national standardized assessment for these standards, using these standards as the basis for music curriculum can provide credibility for music programs.

Music teachers' ability and confidence to teach the national standards for music may be related to their college preparation for those standards. Music teachers receive significant training (in college) in the areas of music performance and music theory. In order to be better prepared to teach the standards, music teachers need more training in improvisation, composition, and music history (Byo, 1999). According to a study by Byo (1999), music teachers felt less able to teach six of the nine standards than their training and education indicated. Furthermore, this study showed that elementary music specialists did not feel qualified to teach the two integrated standards by themselves

(understanding the relationship between the music, the other arts, and disciplines outside the arts; understanding music in relation to history and culture).

Assigning Grades to Music Students

Often, the primary way a teacher communicates a student's academic progress is through a letter grade. With the emphasis on standards in recent years, it has become increasingly important to readjust grading systems to represent these standards. Consequently, for a student's grade to accurately represent a student's musical achievement, it should be based on the National (or local) standards for music education,

MENC conducted an informal survey about grading practices in 1997. Participation was voluntary, and 252 music teachers responded. The results revealed the following about grading practices: Most teachers reported using traditional letter grades (A, B, C, D, F). A large number of teachers reported that grading practices at their schools had been revised in recent years, and most were satisfied with the practices at their schools. The number of students that music teachers were responsible for assigning grades to ranged from 25 to 1,000. Finally, most music teachers (that responded) assigned grades based on performance-based criteria. There were some that assigned grades based on other criteria, such as attendance, effort, behavior, and attitude. Furthermore, some used precise criteria and point systems, while others used grading procedures that were imprecise ("Grading Practices in Music," 1998).

Attendance, effort, behavior, and attitude have long been an important part of music classes, and in many cases, have also been an important part of music grading practices. It is important, however, to separate non-musical criteria from the grading process ("Grading Practices in Music," 1998). Assigning grades to a student based on standards or skills not covered in the class would be inconsistent with grading practices in other subject areas. Effort, behavior, and attitude are difficult, if not impossible to grade objectively. Attendance can be graded objectively, but does not represent a student's understanding of music standards. There are many reasons that music teachers use non-musical criteria when determining student grades. With the

large numbers of students music teachers have it can be difficult to thoroughly and accurately assess all of them on musical criteria. Also, categories such as attendance, effort, and behavior are important to productive music rehearsals, so music teachers may feel that it is necessary to include them in grading practices ("Grading Practices in Music," 1998). An alternative could be to report them separately. Some schools and districts have a separate place on the report card to report behavior. For example, in the Leon County School District (Florida), teachers give a letter grade, as well as a citizenship score. The citizenship score is on a scale of one through four, four being the highest mark, and one being the lowest.

In 1991, McCoy performed a study that examined how high school band choral directors assigned grades to music students, and how this compared to the system proposed by the principals at their schools. Results showed that non-musical criteria were the most significant category in respect to the students' grades. The principals placed more weight on performance technique and cognitive criteria, and less weight on concert attendance and behavior than the music directors did.

Assessment in Band

As noted previously, obstacles instrumental music teachers have in assessing students include the high number of students music teachers have, and the amount of time it takes to thoroughly and accurately assess music performance. In 1993 Burgee studied the comparison between peer, faculty, and student self-evaluations of applied brass jury performances. The study also examined the difference between live performance assessments and assessments of the videotaped performance. Peer evaluations were shown to be accurate and reliable when compared to the faculty evaluations. Furthermore, the any prior knowledge a peer assessor had of the performer did not seem to effect the assessment. The self-evaluations, however, were shown to be unreliable, but students did not assess themselves consistently worse or better than faculty. This study also showed that videotape had little or no effect on the

brass performance assessments. Interestingly, the category that had the lowest mean discrepancy between videotaped and live performance assessments was tone quality (Bergee, 1993).

A second study was performed by Bergee in 1997 that expanded on the 1993 study. This study included brass, percussion, woodwind, and voice performances. Results of this study were consistent with those of the previous study. Peer-evaluations were found to be reliable when compared to evaluations done by faculty. Self-evaluations, however, were not consistent with the evaluations of peers and faculty. These results were consistent across performance medium (brass, woodwind, percussion, voice) (Bergee, 1997).

In research published in 2005, Hewitt performed a similar study on middle and high school instrumental music students. Similar to the results of the Bergee studies, Hewitt found no correlation between the self-evaluations of high school students, and those of expert musicians. This study also showed that middle school students did tend to overrate their own performances; high school students overrated their own performances as well, but not to the extent of the middle school students (Hewitt, 2005). Bergee's study indicated that some college students underrated their performances, although not consistently. This may indicate that students become harsher critics of themselves as they become more experienced.

Another commonly used tool of instrumental music teachers is the practice log or practice journal. A practice log simply requires the student to keep a list or log of the amount of time they practice. Some may even ask students to list what they practice. A 1972 study by Madsen examined the validity of practice journals or logs by college music students. The research found that most college students were not accurate when recording the amount of time they practiced, with many significantly overestimating or underestimating the amount of time practiced (Madsen, 2004).

Summary

Assessment is an essential tool in the music classroom. It can guide instruction and communicate student progress to parents and administrators. There are four types of assessment. Preliminary assessment is used to determine what a student knows prior to instruction. Diagnostic assessment can help determine difficulties a student has that could prevent further learning. Formative assessment is usually informal, occurs during instruction, and can help guide instruction. Summative assessment involves the final product, and most commonly used when assigning grades.

Authentic assessment occurs when knowledge or skills are being assessed in real-life situation. This type of assessment is common to music. One example of this is the performance assessment. In a performance assessment, a real-life situation (a music performance) is being assessed.

Assessments can be interpreted in different ways. Ability-referenced interpretation compares a student's performance to the best performance he/she is capable of. Growth-referenced interpretation compares a student's performance to his/her previous performance(s). Norm-referenced interpretation, which is commonly used with standardized tests, compares a student's performance to that of others. Criterion-referenced interpretation determines what a student can and can not do.

Individual assessment in music classes can be difficult due to the larger sizes of music classes. Furthermore, performance assessments can take a long time to assess, because they must be done in real-time. Some research, however, revealed that assessments of pre-recorded performances are accurate. To save time, teachers could have students submit recordings of performances to be assessed.

The National Standards for Music were developed in 1994 as a result of *Goals 2000: Educate America Act*. While this provided a comprehensive list of skills for music students to demonstrate, it provided no suggestion or method to assess them. Music teachers have come up with several ways to assess their students, but music still lacks a national standardized test to compare students

and school districts. Some attempts at a standardized test have been made in recent years. In 1997, the National Assessment of Educational Progress (NAEP) Arts Assessment assessed music students' abilities to create, perform, and respond. This test involved a small number of students, and provided baseline data to compare groups of students. In 2007, the Florida Music Educators Association (FMEA) piloted a standardized music test for 4th grade students.

In a publication by MENC ("Grading Practices in Music," 1998), suggestions are made concerning the grading of music students. MENC suggests limited grades to musical criteria. It is difficult, if not impossible, to grade great conduct, discipline, or attitude objectively. Research has shown, however, that many music teachers base grades on non-musical criteria.

There has been limited research concerning the assessment or grading of band students. There has been little or no research about the grading and assessment of beginning band students. This study addressed the following questions in regards to assessment and grading in the beginning band class:

- 1. What is being assessed in a beginning band in a successful middle school band program?*
- 2. How are skills and knowledge being assessed in beginning band by directors of successful middle school band programs?*
- 3. How is assessment being used to assign grades to students in beginning bands of successful middle school band programs?*

CHAPTER 3

METHOD

Based on previous research articles and investigations involving assessment methods and grading in music classes, a survey was constructed consisting of four items (see Appendix B). Item one listed eight various assessment tools or methods found on previous research. Respondents were asked to rate each tool or method on a scale of one to five as to how often they use these tools in their beginning band classes. Item two listed eight assessment tools that respondents were to rate from 1 (not important) to 5 (very important) regarding the importance of each item as it relates to the final/semester grades given in their classes. Item three listed twelve characteristics of instrumental performance that respondents were asked to rate from 1 (not important) to 5 (very important) regarding their importance in the performance assessment of a beginning band student. Item four listed nine skills that were taken from the Sunshine State Standards, and the respondents were asked to rate them on a scale of 1 (never) to 5 (very often) regarding to how often they assess these in beginning band classes.

The survey was then piloted by an expert in music research as well as other middle school band directors who were not included in this study. The survey was checked for clarity, as well as spelling and grammatical errors. The results indicated the survey instructions and questions were determined to be clear, easy to understand, and could be completed in three to five minutes. Approval was received from the Florida State University Institutional Review Board (IRB) for the survey and cover letter before any materials were sent out (Appendix D).

The subjects selected for this study were middle school band directors (N = 59) of successful middle school band programs from throughout the state of Florida. For this study, “successful middle school band programs” are defined as middle school bands at schools that have earned “Superior” ratings at the Florida Bandmasters Association’s Concert Band Music Performance

Assessment (MPA) each of the last four years (2003-2006), and have had the same band director each of those years (Florida Bandmasters Association, 2007). The “Superior” rating had to be the final rating, and only had to be achieved by one concert band from the school. The list was derived using data from the Florida Bandmasters Association website, <http://www.flmusiced.org/fba>. The band directors represented a diverse group of schools including different cultural and socio-economic backgrounds.

The surveys were sent to the band directors with a cover letter (Appendix A) stating why he/she was selected to complete the survey, as well as a postage-paid, pre-addressed envelope for returning the completed survey. The surveys were coded with numbers for each of the selected directors so it could be determined who had returned the survey. After ten days a second survey and postage-paid envelope were sent to the directors who had not yet responded. Finally, one last contact was made, by email, to directors who had not yet completed their surveys. Data were compiled from the surveys of responding directors (n = 27), and the mean, median, mode, and standard deviation was calculated for each item. A chi-square analysis was also performed for each item to determine how the numerical data differ from expected and actual received responses.

CHAPTER 4

RESULTS

Fifty-nine band directors in the state of Florida were identified as being the directors at middle schools with successful band programs. Twenty-seven surveys were returned at a return rate of 45.8%. Mean scores analysis of question one revealed that the most commonly used assessment tool in the beginning band classes of participating band directors was “individual playing tests” ($M = 4.667$). The mean scores for question one ranged from 1.852 to 4.667 (see Table 1). Other assessment tools that were revealed to be commonly used include “practice logs/journals” ($M = 3.148$), “student self-assessment/evaluation” ($M = 3.148$), “requirement checklists” ($M = 3.111$), “written tests” ($M = 2.926$), and “worksheets” ($M = 2.852$). The least-commonly used assessment tools were “computer music theory programs” ($M = 1.963$) and “portfolios” ($M = 1.852$).

Chi-square analysis (see Table 2) of question one revealed the results of most of the categories were significant ($p < .05$), meaning there was a general agreement among the participating directors about how often the given assessment tools are used. The assessment tools that had significant results include “individual playing tests” ($\chi^2 = 52.44$, $p < .000$), “portfolios” ($\chi^2 = 23.93$, $p < .000$), “written tests” ($\chi^2 = 12.07$, $p < .017$), “worksheets” ($\chi^2 = 12.82$, $p < .012$), and “computer music theory programs” ($\chi^2 = 16.89$, $p < .002$). The results for “requirement checklists” ($\chi^2 = .96$, $p < .915$), “student self-assessment/evaluation” ($\chi^2 = 4.30$, $p < .367$), and “practice journals/logs” ($\chi^2 = 5.78$, $p < .216$) were found to be non-significant, as the participants’ responses varied.

Mean scores analysis of question two showed that participating directors rated all of the categories as important in assigning grades to students (see Table 3). The categories rated most important were “playing tests/musicianship requirements” ($M = 4.740$), “participation” ($M = 4.740$), “concert attendance” ($M = 4.667$), “conduct/discipline” ($M = 4.370$), and “attendance (rehearsals)”

($M = 3.963$). The categories receiving the lowest mean scores were “written tests” ($M = 3.333$), “written assignments” ($M = 2.852$), and “practice journals/logs” ($M = 2.741$).

Chi-square analysis revealed that the results for most of the grading categories in question two were found to be significant ($p < .05$); there was a general agreement among participants as to the importance of the criteria in determining a student’s grade (see Table 4). Grading categories that had significant results include “playing tests/musicianship requirements” ($\chi^2 = 65.04$, $p < .000$), “participation” ($\chi^2 = 59.48$, $p < .000$), “attendance (rehearsals)” ($\chi^2 = 26.52$, $p < .000$), “concert attendance” ($\chi^2 = 58.00$, $p < .000$), and “conduct/discipline” ($\chi^2 = 43.93$, $p < .000$). Categories with results that were found to be non significant include “written tests” ($\chi^2 = 5.04$, $p < .722$), “practice journals/logs” ($\chi^2 = 7.63$, $p < .106$), and “written assignments” ($\chi^2 = 5.04$, $p < .284$); the participants did not agree on the importance of the criteria as it pertains to a student’s final grade.

Mean scores analysis of question three revealed that most of the criteria listed were rated important for performance assessments of beginning band students (see Table 5). Most of the listed criteria had mean scores higher than 4; these include “posture” ($M = 4.889$), “hand position” ($M = 4.778$), “note accuracy” ($M = 4.630$), “rhythmic accuracy” ($M = 4.741$), “articulation” ($M = 4.482$), “air support” ($M = 4.926$), “pulse/tempo” ($M = 4.482$), and “phrasing” ($M = 4.185$). The categories with the lowest mean scores were “dynamics (observance of markings)” ($M = 3.963$), “style” ($M = 3.630$), and “expression” ($M = 3.519$).

Chi-square analysis of question three found the results of most categories to be significant ($p < .05$), because most participants agreed on the importance of each category in a beginning band student performance assessment (see Table 6); these include “posture” ($\chi^2 = 8.33$, $p < .000$), “hand position” ($\chi^2 = 61.33$, $p < .000$), “note accuracy” ($\chi^2 = 51.70$, $p < .000$), “rhythmic accuracy” ($\chi^2 = 65.04$, $p < .000$), “articulation” ($\chi^2 = 36.15$, $p < .000$), “air support” ($\chi^2 = 89.48$, $p < .000$), “pulse/tempo” ($\chi^2 = 33.93$, $p < .000$), “dynamics (observance of markings)” ($\chi^2 = 13.56$, $p < .009$), “intonation” ($\chi^2 = 26.89$, $p < .001$), and “phrasing”

($\chi^2 = 19.48, p < .001$). The results of the categories “style” ($\chi^2 = 9.11, p < .058$) and “expression” ($\chi^2 = 5.41, p < .248$) were found to be non-significant; there was no clear consensus among participants as to the importance of these criteria in a beginning band performance assessment.

Mean scores analysis of question four revealed a wide variance in how often teachers assessed the given musical skills, with scores ranging from 1.667 to 4.852 (see Table 7). The standards that were rated as the most commonly assessed were “performing on instruments” ($M = 4.852$) and “reading/notating music” ($M = 4.630$). Other standards with were rated as being assessed often were “listening to/analyzing/describing music” ($M = 3.037$), “evaluating music and music performance” ($M = 3.667$), “understanding the relationship between music and other subjects” ($M = 3.185$), and “understanding the relationship between music and history/culture” ($M = 3.370$). The standards rated as the least commonly used were “composing/arranging music” ($M = 1.852$) and “improvisation” ($M = 1.667$).

Chi-square analysis of question four (see Table 8) revealed that the results for most of the given musical skills were significant ($p < .05$), as the participants were generally agreed on how often they assess these skills. These included “singing” ($\chi^2 = 14.30, p < .006$), “performing on instruments” ($\chi^2 = 80.59, p < .000$), “improvisation” ($\chi^2 = 28.00, p < .000$), “composing and arranging music” ($\chi^2 = 23.93, p < .000$), “reading and notation music” ($\chi^2 = 47.26, p < .000$), and “understanding the relationship between music and other subjects” ($\chi^2 = 10.22, p < .037$). The results for “listening to/analyzing/describing music” ($\chi^2 = 6.15, p < .188$), “evaluating music and music performance” ($\chi^2 = 6.89, p < .142$), and “understanding the relationship between music and history/culture” ($\chi^2 = 8.37, p < .079$) were found to be non-significant, as the participants’ did not agree on how often they assess these musical skills.

Table 1

Mean, Median, and Mode Scores to Question 1:

“Please rate how often you use the following assessment tools in your band classes.”

Item	Mean	Median	Mode	S.D.
Individual Playing Tests	4.667	5	5	.620
Requirement Checklists	3.111	3	3	1.423
Portfolios	1.852	1	1	1.231
Written Tests	2.926	3	3	.997
Worksheets	2.852	3	2	1.027
Student Self-Assessment/Eval	3.148	3	5	1.586
Computer Music Theory Programs	1.963	2	1	1.192
Practice Logs/Journals	3.148	4	1,5	1.657

Table 2

Chi-Square Analysis of Question 1:

“Please rate how often you use the following assessment tools in your band classes.”

Item	Chi-Square	df	Asymp. Sig.
Individual Playing Tests	52.44	4	.000
Requirement Checklists	.96	4	.915
Portfolios	23.93	4	.000
Written Tests	12.07	4	.017
Worksheets	12.82	4	.012
Student Self-Assessment/Eval	4.30	4	.367
Computer Music Theory Programs	16.89	4	.002
Practice Logs/Journals	5.78	4	.216

Table 3

Mean, Median, and Mode Scores to Question 2:
“Please rate the following items as to their importance in determining a student’s final grade in your class.”

Item	Mean	Median	Mode	S.D.
Playing tests/Musicianship Req.	4.740	5	5	.594
Participation	4.740	5	5	.526
Attendance (Rehearsals)	3.963	5	5	1.454
Written Tests	3.333	3	5	1.441
Practice Journals/Logs	2.741	3	1	1.655
Concert Attendance	4.667	5	5	.679
Conduct/Discipline	4.370	5	5	1.214
Written Assignments	2.852	3	3	1.433

Table 4

Chi-Square Analysis of Question 2:

“Please rate the following items as to their importance in determining a student’s final grade in your class.”

Item	Chi-Square	df	Asymp. Sig.
Playing tests/Musicianship Req.	65.04	4	.000
Participation	59.48	4	.000
Attendance (Rehearsals)	26.52	4	.000
Written Tests	2.07	4	.722
Practice Journals/Logs	7.63	4	.106
Concert Attendance	58.00	4	.000
Conduct/Discipline	43.93	4	.000
Written Assignments	5.04	4	.284

Table 5

Mean, Median, and Mode Scores for Question 3:

“Please rate the following items as to their importance in performance assessment of a beginning band student.”

Item	Mean	Median	Mode	S.D.
Posture	4.889	5	5	.320
Hand Position	4.778	5	5	.424
Note Accuracy	4.630	5	5	.688
Rhythmic Accuracy	4.741	5	5	.594
Articulation	4.482	5	5	.753
Air Support	4.926	5	5	.267
Pulse/Tempo	4.482	5	5	.700
Dynamics (Observance of Markings)	3.963	4	5	1.055
Intonation	4.333	5	5	.877
Style	3.630	4	4	1.079
Expression	3.519	4	3	1.221
Phrasing	4.185	4	5	.879

Table 6

Chi-Square Analysis of Question 3:

“Please rate the following items as to their importance in performance assessment of a beginning band student.”

Item	Chi-Square	df	Asymp. Sig.
Posture	8.33	4	.000
Hand Position	61.33	4	.000
Note Accuracy	51.70	4	.000
Rhythmic Accuracy	65.04	4	.000
Articulation	36.15	4	.000
Air Support	89.48	4	.000
Pulse/Tempo	33.93	4	.000
Dynamics (Observance of Markings)	13.56	4	.009
Intonation	26.89	4	.001
Style	9.11	4	.058
Expression	5.41	4	.248
Phrasing	19.48	4	.001

Table 7

Mean, Median, and Mode Scores for Question 4:

“Please rate the following items as to how often you assess them in your beginning band class.”

Item	Mean	Median	Mode	S.D.
Singing	2.370	2	1	1.621
Performing on Instruments	4.852	5	5	.456
Improvisation	1.667	2	1,2	.679
Composing/Arranging Music	1.852	2	2	.718
Reading/Notating Music	4.630	5	5	.629
Listening to/analyzing/ describing music	3.037	3	3	1.126
Evaluating Music and Music Performance	3.667	4	4	1.240
Understanding the relationship between music and other subjects	3.185	3	4	1.111
Understanding the relationship between music and history/culture	3.370	4	4	1.214

Table 8

Chi-Square Analysis of Question 4:
“Please rate the following items as to how often you assess them in your beginning band class.”

Item	Chi-Square	df	Asymp. Sig.
Singing	14.30	4	.006
Performing on Instruments	80.59	4	.000
Improvisation	28.00	4	.000
Composing/Arranging Music	23.93	4	.000
Reading/Notating Music	47.26	4	.000
Listening to/analyzing/ describing music	6.15	4	.188
Evaluating Music and Music Performance	6.89	4	.142
Understanding the relationship between music and other subjects	10.22	4	.037
Understanding the relationship between music and history/culture	8.37	4	.079

Table 9

Frequency of Responses to Question 1:

“Please rate how often you use the following assessment tools in your band classes.”

Item	1	2	3	4	5
Individual Playing Tests	0	0	2	5	20
Requirement Checklists	5	4	7	5	6
Portfolios	15	6	3	1	2
Written Tests	1	9	10	5	2
Worksheets	1	11	8	5	2
Student Self-Assessment/Eval	7	2	6	4	8
Computer Music Theory Programs	13	7	3	3	1
Practice Logs/Journals	8	2	3	6	8

Table 10

Frequency of Responses to Question 2:

“Please rate the following items as to their importance in determining a student’s final grade in your class.”

Item	1	2	3	4	5
Playing tests/Musicianship Req.	0	0	2	3	22
Participation	0	0	1	5	27
Attendance (Rehearsals)	3	2	4	2	16
Written Tests	4	4	6	5	8
Practice Journals/Logs	10	3	5	2	7
Concert Attendance	0	0	3	3	21
Conduct/Discipline	0	0	3	3	21
Written Assignments	7	3	9	3	5

Table 11

Frequency of Responses for Question 3:

“Please rate the following items as to their importance in performance assessment of a beginning band student.”

Item	1	2	3	4	5
Posture	0	0	0	3	24
Hand Position	0	0	0	6	21
Note Accuracy	0	0	3	4	20
Rhythmic Accuracy	0	0	2	3	22
Articulation	0	0	4	6	17
Air Support	0	0	0	2	25
Pulse/Tempo	0	0	3	8	16
Dynamics (Observance of Markings)	1	1	6	9	10
Intonation	0	1	4	7	15
Style	1	3	7	10	6
Expression	2	3	8	7	7
Phrasing	0	1	5	9	12

Table 12

Frequency of Responses for Question 4:

“Please rate the following items as to how often you assess them in your beginning band class.”

Item	1	2	3	4	5
Singing	13	4	2	3	5
Performing on Instruments	0	0	1	2	24
Improvisation	12	12	3	0	0
Composing/Arranging Music	9	13	5	0	0
Reading/Notating Music	0	0	2	6	19
Listening to/analyzing/ describing music	2	7	9	6	3
Evaluating Music and Music Performance	2	3	5	9	8
Understanding the relationship between music and other subjects	2	6	6	11	2
Understanding the relationship between music and history/culture	3	3	6	11	4

CHAPTER 5

DISCUSSION

The primary findings of this study demonstrate that the participating band directors at middle schools with successful band programs assess two of the skills that may be considered traditional for music, “performing on instruments” and “reading/notating music” very often, more so than the other standards. The skills that were found as the least-often assessed were “singing,” “improvisation,” and “composing/arranging music.” These skills may be considered less traditional in a band class. The assessment tool rated as the most commonly used was “individual playing tests.” The tools found as being the least commonly used in student assessment are “portfolios” and “computer music theory programs.”

When giving playing tests, responding directors rated all of the criteria in a beginning band music performance assessment highly, including posture, hand position, note accuracy rhythmic accuracy, articulation, air support, pulse/tempo, dynamics (observance of markings), intonation, style, expression, and phrasing. The results for style and expression, however, were found to be not significant; there was some discrepancy among directors as to their importance. Some participants rated style and expression as very important, while others rated them as not important for a beginning band music performance assessment.

Finally, directors rated “playing tests/musicianship requirements,” “concert attendance,” “participation,” and “conduct/discipline” the highest as they pertain to a student’s final grade in their beginning band class. This was contrary to suggestions made in related literature (“Grading Practices in Music,” 1998), but consistent with prior research (McCoy, 1991). Specifically, this study found the following:

“What is being assessed in a beginning band in a successful middle school band program?”

When presented with nine musical skills, directors overwhelmingly rated “performing on instruments” and “reading/notating music” as the most commonly assessed skills in their beginning band classes. These skills are the most closely related to traditional instrumental performance, which may be why they are rated highly. Also, in Florida, the only assessments performed by the Florida Bandmasters Association are performance assessments, which primarily assess the student’s performance and music literacy (Florida Bandmasters Association, 2007). This activity could further explain why these skills are regularly practiced and assessed. “Understanding the difference between music and other subjects” was also rated moderately high.

“Singing,” “improvisation,” and “composing/arranging music” were rated as the least-assessed standards in beginning band classes. It is possible that these skills were rated low because singing, improvisation, and composing/arranging music are not skills that are assessed by the Florida Bandmasters Association (FBA) Music Performance Assessment (Florida Bandmasters Association, 2007), which was the primary tool used in determining which band directors were included in this study. Furthermore, these skills may not be considered traditional for beginning band classes. Consequently, these concepts are likely to be practiced less, and thus assessed less often.

“Evaluating music and music performance,” “listening to/analyzing/describing music,” and “understanding the relationship between music and history/culture” were all rated moderately high. Chi-square analysis found that there was no consensus among directors for these three standards; the responses for these skills varied significantly between participants. Again, it is possible this finding was due to these skills not being assessed by FBA Music Performance Assessments (Florida Bandmasters Association, 2007). Also, some of the participants may teach their students to evaluate music performance and to listen to and analyze music to help them prepare for performance

assessments. This may be why some participants assess those skills more often. “Understanding the relationship between music and history/culture” may depend on the music being taught in class; some music may require that skill more than other music.

When asked to rate items as to their importance on a beginning band performance assessment, directors rated all of the given items highly. The highest rated items were “posture” and “air support.” The lowest rated items, although still rated as important, were “intonation,” “style,” “expression,” and “phrasing.” As stated previously, the participants were from middle school with successful middle school bands. It is possible that the participants hold them to the same standards on performance assessments as their older students. This could be the reason all of these categories were rated so highly. Also, all of the items in this question can be found on the adjudication form for the Florida Bandmasters Association Concert Band, Solo, and Ensemble Music Performance Assessments (Florida Bandmasters Association, 2007). The Concert Band Music Performance Assessment was part of the criteria to be included in this study, which may be another reason why all of the items were rated so highly. This finding suggests that the FBA Music Performance Assessment is a strong influence, even on beginning band instruction.

The results for “style” and “expression,” however, were revealed as not significant through chi-square analysis. The responses for these items varied, for no apparent reason. More research may be needed to clarify the difference between the participants who rated “style” and “expression” as important, and those who did not.

“How are skills and knowledge being assessed in beginning band by directors of successful middle school band programs?”

The responding band directors rated “individual playing tests” as the most commonly used assessment tool in their beginning band classes. This is consistent with the results from question four, which revealed that “performing on instruments” was the most commonly assessed of the given musical skills.

Individual playing tests, which are performance assessments, would be the most common method to assess a student's ability to perform on an instrument. "Written tests" and "worksheets" were rated moderately high as to how often they are used in the beginning band classroom. It is possible that these two assessment tools are used to assess reading and notating instruments, which was found in this study to be a commonly assessed skill. The lowest rated assessment tools were "portfolios" and "computer music theory programs."

The results for "computer music theory programs" were not unexpected, as there was little or no literature or research concerning their use in assessment or grading. More research should be done concerning the use of the computer programs to determine why they are not used. Research concerning the use of technology in assessment may answer this question. "Portfolios," however, were often cited in literature as a useful assessment tool, but was still rated as not being used often. It is possible, however, that portfolios are used less often because their assessment is more complicated and less objective, as assessment of music can vary from teacher to teacher (Goolsby, 1995).

Chi-square analysis revealed the results for "requirement checklists," "student self-assessment/evaluation," and "practice/journals" to be not significant. This is because the responses for these items varied among participants (see Table 9). Despite research, which revealed student self-assessment to be inconsistent (Hewitt, 2005), some directors indicated that they do use them often, although the results for this item varied among participants. The responses for "practice journals/logs" varied significantly. The mean score was 3.148, but the mode scores were both 1 and 5. Most directors indicated they used practice journals or logs either "very often" or "never." This could be because practice journals are meant to track the amount of time a student practices, so they would likely be turned in regularly. Participants who use them probably use them very often, and the other participants never use them at all. The results for "practice journals/logs" being not significant is consistent with the research that practice journals are not effective, as most students misrepresent the amount of time they spend practicing (Madsen, 2004).

“How is assessment being used to assign grades to students in beginning bands of successful middle school band programs?”

Responding directors rated “playing tests/musicianship requirements,” “participation,” “concert attendance,” and “conduct/discipline” as the most important items in determining a student’s final grade in their beginning band classes. “Attendance (rehearsals)” was also rated highly. It is possible that some directors feel that discipline is part of what is taught in a band program, which may explain these results. Furthermore, attendance and discipline are important to have productive band classes or rehearsals, which may be why the participants rated them as important when assigning student grades. These results are not consistent with related literature that suggests that grades in music classes would be most effective if they only include musical skills, and not non-musical items such as attendance, conduct, and participation (“Grading Practices in Music,” 1998). However, the results are consistent with previous research that shows music teachers believe items such as attendance and participation should factor into a student’s grade (McCoy, 1991).

Chi-square analysis of the question about grading criteria revealed the results for “written tests,” “practice journals/logs,” and “written assignments” to be non significant. This is because the responses for these items varied among the participants (see Table 10). The results for “written tests” and “written assignments” may vary because they can be used to assess several different skills, so their use would depend greatly on what skills the participant assesses, and how they assess them. The responses for “practice journals/logs” varied significantly, with the majority of the responses being 1 (not important) and 5 (very important). This is consistent with the results from the first question on the survey, where the directors’ responses varied significantly when asked how often they use practice journals.

Implications

The results of the study could serve as a guide for middle school directors to use in their beginning band classrooms. The findings provide a model of assessment and grading that is used in successful middle school band programs throughout Florida. Furthermore, the results of this research could help future teachers develop a system of assessment to use in their classrooms that could enhance instruction. Administrators and district music supervisors could also use these results in their school or district music programs to help young teachers struggling with assessment, or to develop a uniform assessment and grading policy among music programs.

This study revealed some common assessment and grading habits of successful teachers. These include the regular assessment of a student's ability to perform on instruments as well as to read and notate music, the assessment of musical skills and abilities during performance assessments, the regular use of performance assessments, and the use these performance assessments and/or musicianship requirements when determining student grading, as well as non-musical criteria such as attendance, participation, and conduct. Beginning band teachers should understand the importance of regular, effective assessment and grading practices in their classrooms, and should consider the trends found in this study to be a model for beginning band assessment and grading in successful middle school band programs.

There were several limitations that became apparent during this study. The first was the relatively small number of subjects chosen for the study. The criteria for inclusion was narrow; this was intended to include middle school band directors in Florida who have had significant success. While that goal was met, the criteria also limited the number of possible subjects.

Another limitation of this study concerning the criteria for inclusion is the definition used for a "successful middle school band." For this study, a successful middle school band was defined as one that has received a "Superior" rating each of the last four years (2003-2006), all with the same director. There are other indicators of successful programs, including size (student participation),

which could have been used. The study's criteria only include band programs that have been successful at group performance assessments.

Finally, one other limitation of this study was the size of the geographical region from which subjects were selected. Only directors of middle school bands from Florida were selected. Broadening the study to include middle school from other parts of the United States would have increased the number and diversity of subjects selected, and would have given a better indication of the assessment and grading trends in beginning bands nation wide.

The results found in this study could lead to further research that could be done concerning the grading and assessment of beginning band students. One suggestion would be to ask directors similar questions concerning the grading and assessment of their non-beginning classes. The study could examine the differences, if any, between the two types of classes, as well as what techniques work with one type of class, but not the other. Do successful middle school band directors use the same assessment and grading techniques on all of their classes, or do they differ depending on the level? Future research could also examine the relationship between assessment habits of middle school and high school band directors, or between band, choral, and orchestra directors.

Another suggestion would be to survey the beginning students of successful middle school directors, and examine their perception of assessment and grading in their band class in comparison to that of their directors. A common fear of middle school band directors is that if they give a student anything less than an "A" it could effect the enrollment in their program. It would be interesting to see the students' perception of assessment in their band class as it relates to their perception of assessment in their other classes.

APPENDIX A
LETTER

Michael Antmann
3934 Paces Court
Tallahassee, FL 32311

May 24, 2007

Dear Colleague:

My name is Michael Antmann, and I am the band director at Swift Creek Middle School in Tallahassee, FL. I am currently working on a research project for my thesis at FSU, and I need your help. You have been selected for this study because of your band's success at the FBA District Concert Band Music Performance Assessments in recent years.

Attached to this letter is a survey about assessment and assigning grades in Beginning Band. Please complete this survey at your earliest convenience, and return to me in the enclosed envelope. Please complete and return the survey by June 4, 2007. All surveys and information will be kept confidential (no names will be on the surveys, and names will not be used in the study), to the extent allowed by law. Surveys are being stored at my home, and will be destroyed as soon as the study is complete. Participation is completely voluntary, and there is no risk to you.

If you have any questions, please contact me at (850) 228-9123, or michaelantmann@gmail.com. You can contact my major professor, Dr. Steven Kelly, at (850) 487-4868, or the Institutional Review Board at (850) 644-9694.

Thank you,

Michael Antmann

APPENDIX B
SURVEY

Grading and Assessment in Beginning Band Survey

1. Please rate how often you use the following assessment tools in your band classes:
(1 = never, 5 = very often)

Individual Playing Tests	1	2	3	4	5
Requirement Checklists	1	2	3	4	5
Portfolios	1	2	3	4	5
Written Tests	1	2	3	4	5
Worksheets	1	2	3	4	5
Student Self-Assessment/Evaluation	1	2	3	4	5
Computer Music Theory Programs	1	2	3	4	5
Practice Logs/Practice Journals	1	2	3	4	5
Student Self-Evaluation	1	2	3	4	5

2. Please rate the following items as to their importance in determining a student's final grade **in your class**: (1 = not important, 5 = very important)

Playing tests/musicianship requirements	1	2	3	4	5
Participation	1	2	3	4	5
Attendance (Rehearsals)	1	2	3	4	5
Written tests	1	2	3	4	5
Practice Journals/Logs	1	2	3	4	5
Concert Attendance	1	2	3	4	5
Conduct/Discipline	1	2	3	4	5
Written Assignments	1	2	3	4	5

Please turn over...

3. Please rate the following items as to their importance in a performance assessment of a *beginning band student*: (1 = not important, 5 = very important)

Posture	1	2	3	4	5
Hand Position	1	2	3	4	5
Note Accuracy	1	2	3	4	5
Rhythmic Accuracy	1	2	3	4	5
Articulation	1	2	3	4	5
Air Support	1	2	3	4	5
Pulse/Tempo	1	2	3	4	5
Dynamics (Observance of markings)	1	2	3	4	5
Intonation	1	2	3	4	5
Style	1	2	3	4	5
Expression	1	2	3	4	5
Phrasing	1	2	3	4	5

4. Please rate the following items as to how often you assess them in your beginning band class: (1 = rarely/never, 5 = very often)

Singing	1	2	3	4	5
Performing on instruments	1	2	3	4	5
Improvisation	1	2	3	4	5
Composing/Arranging Music	1	2	3	4	5
Reading/Notating Music	1	2	3	4	5
Listening to/analyzing/describing music	1	2	3	4	5
Evaluating Music and Music Performance	1	2	3	4	5
Understanding the relationship between Music and other subjects	1	2	3	4	5
Understanding the relationship between Music and history/culture.	1	2	3	4	5

Thank you for completing the survey, please return in the postage-paid envelope by June 4, 2007!

APPENDIX C
SCHOOLS OF RESPONDING DIRECTORS

Schools of Responding Directors

Jim C. Bailey Middle School	Pensacola, FL
Ransom Middle School	Cantonment, FL
Mowat Middle School	Lynn Haven, FL
Raa Middle School	Tallahassee, FL
Deerlake Middle School	Tallahassee, FL
Fort Clarke Middle School	Gainesville, FL
Benito Middle School	Tampa, FL
Hill Middle School	Tampa, FL
Liberty Middle School	Orlando, FL
Martinez Middle School	Lutz, FL
Oak Grove Middle School	Clearwater, FL
Andrew Jackson Middle School	Titusville, FL
Jewett Academy	Lakeland, FL
Sleepy Hill Middle School	Lakeland, FL
St. Joseph School	Winter Haven, FL
Stuart Middle School	Stuart, FL
Independence Middle School	Jupiter, FL
Watson P. Duncan Middle School	Palm Beach Gardens, FL
Driftwood Middle School	Hollywood, FL
Pioneer Middle School	Cooper City, FL
Lyons Creek Middle School	Tamarac, FL
Ramblewood Middle School	Coral Springs, FL
Southwood Middle School	Miami, FL
Lawton Chiles Middle School	Miami, FL
Callahan Middle School	Callahan, FL
Gulf Middle School	Cape Coral, FL
Tavares Middle School	Tavares, FL

APPENDIX D
INSTITUTIONAL REVIEW BOARD APPROVAL

Office of the Vice President For Research Human Subjects Committee
Tallahassee, Florida 32306-2742
(850) 644-8673 .
FAX (850) 644-4392

APPROVAL MEMORANDUM

Date: 5/30/2007

To: Michael Antmann

Address: 3934 Paces Court

Dept.: MUSIC SCHOOL

From: Thomas L. Jacobson, Chair

Re: Use of Human Subjects in Research Grading and Assessment of Beginning Band Students

The application that you submitted to this office in regard to the use of human subjects in the proposal referenced above have been reviewed by the Secretary, the Chair, and two members of the Human Subjects Committee. Your project is determined to be Expedited per 45 CFR § 46.110(7) and has been approved by an expedited review process.

The Human Subjects Committee has not evaluated your proposal for scientific merit, except to weigh the risk to the human participants and the aspects of the proposal related to potential risk and benefit. This approval does not replace any departmental or other approvals, which may be required.

If you submitted a proposed consent form with your application, the approved stamped consent form is attached to this approval notice. Only the stamped version of the consent form may be used in recruiting research subjects.

If the project has not been completed by 5/22/2008 you must request a renewal of approval for continuation of the project. As a courtesy, a renewal notice will be sent to you prior to your expiration date; however, it is your responsibility as the Principal Investigator to timely request renewal of your approval from the Committee.

You are advised that any change in protocol for this project must be reviewed and approved by the Committee prior to implementation of the proposed change in the protocol. A protocol change/amendment form is required to be submitted for approval by the Committee. In addition, federal regulations require that the Principal Investigator promptly report, in writing any unanticipated problems or adverse events involving risk to research subjects or others.

By copy of this memorandum, the Chair of your department and/or your major professor is reminded that he/she is responsible for being informed concerning research projects involving human subjects in the department, and should review protocols as often as needed to insure that the project is being conducted in compliance with our institution and with DHHS regulations.

This institution has an Assurance on file with the Office for Human Research Protection.

The Assurance Number is IRB00000446.

Cc: Steven Kelly, Advisor HSC No. 2007.499

REFERENCES

- Asmus, E. P. (1999, September). Music assessment concepts. *Music Educators Journal*, 86(2), 19-24.
- Bergee, M. J. (1993, Spring). A comparison of faculty, peer, and self-evaluation of applied brass jury performances. *Journal of Research in Music Education*, 41(1), 19-27.
- Bergee, M. J. (1997, Winter). Relationships among faculty, peer, and self-evaluations of applied performances. *Journal of Research in Music Education*, 45(4), 601-612.
- Byo, S. L. (1999, Summer). Classroom teachers' and music specialists' perceived ability to implement the national standards for music education. *Journal of Research in Music Education*, 47(2), 111-123.
- Cope, C. O. (1996, July). Steps toward effective assessment. *Music Educators Journal*, 83(1), 39-42.
- Duerksen, G. L. (1995, Winter). Using planned assessment techniques to increase student growth in music education. *The Quarterly Journal of Music Teaching and Learning*, VI(4), 57-64.
- Fallis, T. L. (1994, November). National standards: what's next? *Music Educators Journal*, 81(3), 26-28+47.
- Farrell, S. R. (1997). *Tools for powerful student evaluation* (2nd ed.). Milwaukee: Meredith Music Publications.
- Florida Bandmasters Association. (2007). Retrieved June 28, 2007, from Florida Bandmasters Association Web site: <http://www.flmusiced.org/fba>
- Florida Bandmasters Association (2005). *Music performance assessments*. Retrieved May 4, 2007, from Florida Bandmasters Association Web site: http://www.flmusiced.org/FBA/mpa_results.htm
- Goolsby, T. W. (1999, September). Assessment in instrumental music. *Music Educators Journal*, 86(2), 31-35, 50.
- Grading practices in music. (1998, March). *Music Educators Journal*, 84(5), 37-40.
- Hewitt, M. P. (2005, Summer). Self-evaluation accuracy among high school and middle school instrumentalists. *Journal of Research in Music Education*, 53(2), 148-161.
- Lecroy, H. (1993, September). Assessment and strategic planning. *Music Educators Journal*, 86(2), 36-40.

- Madsen, C. K. (2004, Spring). A 30-year follow-up study of actual applied music practice versus estimated practice. *Journal of Research in Music Education*, 52(1), 77-88.
- McCoy, C. W. (1991, Fall). Grading students in performing groups: a comparison of principals' recommendations with directors' practices. *Journal of Research in Music Education*, 39(3), 181-190.
- MENC: The National Association for Music Education. (2007, April 23). *MENC position statement on health in music education*. Retrieved May 3, 2007, from MENC: The National Association for Music Education Web site: <http://www.menc.org/connect/surveys/position/assessment.html>
- MENC: The National Association for Music Education. (1994). *National standards for music education. The nine content standards, adopted in 1994, along with related links*. Retrieved May 27, 2007, from MENC Web site: <http://www.menc.org/publication/books/standards.htm>
- MENC: The National Association for Music Education. (n.d.). *Performance standards for music: grades preK-12*. Retrieved May 21, 2007, from MENC Web site: http://www.menc.org/publication/books/performance_standards/contents.html
- Oosterhof, A. (2001). *Classroom applications of educational measurement* (3rd ed.). Upper Saddle River, NJ: Prentice Hall, Inc. (Original work published 1994)
- Orr, C. S. (2007, March 6). *Florida elementary music assessment field test*. Retrieved June 28, 2007, from Florida Department of Education Web site: <http://www.fldoe.org/asp/k12memo/pdf/Phase1FloridaElementaryMusicAssessmentGrade4FieldTest.pdf>
- Radocy, R. E. (1995, Winter). Planning assessment in music education: it's not risk free. *The Quarterly Journal of Music Teaching and Learning*, VI(4), 19-25.
- Saunders, T. C., & Holahan, J. M. (1997, Summer). Criteria-Specific Rating Scales in the Evaluation of High School Instrumental Performance. *Journal of Research in Music Education*, 45(2), 259-272.
- Schneider, C. (2005, November). Measuring student achievement in the future based on lessons from the past: the NAEP Arts Assessment. *Music Educators Journal*, 92(2), 56-61.
- U.S. Department of Education. (2004a). *Executive summary of the No Child Left Behind Act of 2001*. Retrieved June 28, 2007, from U.S. Department of Education Web site: <http://www.ed.gov/nclb/overview/intro/execsumm.html>

U.S. Department of Education. (2004b). *The Elementary and Secondary Education Act*. Retrieved February 26, 2007, from <http://www.ed.gov/policy/elsec/leg/esea02/index.html>

White, S., & Vanneman, A. (2000, February). Student musical activities and achievement in music: NAEP 1997 Arts Assessment. *Educational Statistics Quarterly*, 1(4). Retrieved May 21, 2007, from http://nces.ed.gov/programs/quarterly/vol_1/1_4/3-esq14-b.asp

BIOGRAPHICAL SKETCH

Name: Michael David Antmann

Birthplace: Bay Shore, New York

Higher Education: Florida State University
Tallahassee, Florida
Major: Music Education
Degree: M.M.E. (2007)

Florida State University
Tallahassee, Florida
Major: Music Education
Degree: B.M.E. (2001)

High School: Nova High School
Davie, Florida
Graduated, 1997

Professional Experience: Swift Creek Middle School
Tallahassee, Florida
2001-2007
Band Director