THE IMPACT OF RUBRIC USE IN ELEMENTARY GENERAL MUSIC

by

INSUN KO

(Under the Direction of Mary Leglar)

ABSTRACT

The purpose of the study was to determine if the use of a teacher-authored rubric improves elementary students' comprehension and retention of material related to singing and playing an instrument. The quasi-experimental design employed a pretest, a posttest, and a retention test. The subjects were 608 students enrolled in kindergarten through grade five attending a suburban school in the southeastern part of the United States. Intact classes (four classes at each grade level) were randomly assigned to an experimental (n=300) or a control (n=308) group. The treatment period consisted of four consecutive lessons for each class of students. All groups were instructed on the same material within a month of time. The experimental groups were given a copy of a rubric and the contents were discussed before beginning the instruction. The control group received instruction without the use of a rubric. A retention test was administered to each group a month after their completion of the posttest. An independent two-sample t-test was used to determine if there was a significant difference between the mean scores of the posttest and the retention test. Results of the t-test revealed that scores of the experimental group were significantly higher than those of the control group for both variables (singing and playing). In addition, an informal student survey indicated that the students reacted favorably to using the rubric. It was concluded that the use of a rubric in music

classes at the elementary level could be an effective tool for improving and assessing student performance.

INDEX WORDS: Music education, Rubrics, Retention, Student Achievement, Performance

THE IMPACT OF RUBRIC USE IN ELEMENTARY GENERAL MUSIC

by

INSUN KO

BM, University of Georgia, 2005

MM, University of Georgia, 2007

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF EDUCATION

ATHENS, GEORGIA

2014

THE IMPACT OF RUBRIC USE IN ELEMENTARY GENERAL MUSIC

by

INSUN KO

Major Professor:	Mary Leglar
Committee:	Stephen Valdez
	Roy Legette

Electronic Version Approved:

Julie Coffield
Interim Dean of the Graduate School
The University of Georgia
December 2014

DEDICATION

This dissertation is dedicated to my Heavenly Father,

my husband, Tong-Seok,

our children, Elissa and Ethan,

and my parents, Shin-Ho and Eun-Sook Baik,

with love and gratitude

for your unfailing support and encouragement.

ACKNOWLEDGMENTS

I would like to take this opportunity to thank my major professor, Dr. Mary Leglar, for continuing to encourage me and for working many late hours to help me refine my ideas and revise my report. I would also like to thank my committee, Dr. Stephen Valdez and Dr. Roy Legette, for reading my drafts and providing directions and advice.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION

With the passing of the Goals 2000: Educate America Act, the arts were written into federal law as a core academic subject in K–12 public schools. The arts maintained their status under the No Child Left Behind (NCLB) Act, which placed greater emphasis on assessment and accountability. Ever since NCLB became law, assessment has become one of the most pervasive aspects on the American educational landscape.

In most school districts, assessment has become an important issue related to standards-based education. Assessment plays an important role in instructional practices because it places additional focus on the objective or purpose of the learning and provides a measure of progress based on particular goals and standards. Although assessment serves a summative purpose, the main intent of assessment in education is to follow students' improvement throughout a given program, evaluate their needs, and support their ongoing progress (Audrey, 2006; Leonhart, 2005; Taggart & Phifer, 1998). It is assumed that, to be effective facilitators, teachers should know what students have learned and be able to evaluate their understanding.

Ideally, each school district has established objectives based on the National Standards for Music Education. In addition, to determine whether students are meeting these objectives, each school district ideally has developed assessment methods that identify what students have learned in music classes. Further, it is important that the assessment, regardless of method, is consistent and fair for all students; in effect, all students should be evaluated with the same scale and criteria.

The use of rubrics provides a consistent, fair, and effective way to focus on important aspects of a task. The use of teacher-made rubrics in particular benefits both teachers and students. Teacher-made rubrics reflect the individual instructor's decisions about learning objectives, evaluation criteria, and goals for skill mastery. They thus clarify and define, for the students as well as the teacher, the teacher's expectations of what he or she wants students to achieve through a particular lesson (Brian, 2012; Eppink, 2002; Leonhart, 2005; Stiggins, 2005). Teacher-made rubrics have also been found to enrich and facilitate the assessment process for both the instructor and the students. Since rubrics give clear indications of what students need to accomplish, they are aware of what is being assessed, and the grading process is no longer a mystery to them (Asmus, 1999; Hall, 1997; Koops, 2008; Smith, 2001; Stauffer, 1999). Researchers have also suggested that rubrics can be used in elementary music classes to develop a set of guidelines for clearly assessing student performance (Asmus, 1999; Brophy, 2000; Stauffer, 1999).

Efficiency is another benefit derived from the use of rubrics. Once created, rubrics can serve as effective time savers. Most music educators would agree that one of the greatest challenges in teaching general music classes is assessing large numbers of students in short periods of time (Levi, 2005; Scott, 2001). Hale and Green (2009) recommended that teachers assessing music students use "the simplest procedure that is adequate to get the job done" and strongly supported the use of rubrics.

Purpose and Need for the Study

Although several studies (Brinko, 1993; Gordon, 2002; Leonhart, 2005) have shown that rubrics have a positive impact on students' academic achievement in general classroom settings, very few have examined the impact of rubrics on the achievement levels of elementary students specifically in the context of the general music classroom.

Using a quasi-experimental pretest and posttest design, the primary purpose of this study was to determine whether a teacher-made rubric aided student comprehension and retention of concepts and skills related to singing and playing an instrument. The following questions guided the study:

1. Is there a significant difference between the immediate pretest and posttest scores of students who use a rubric throughout the instructional period (experimental group) and students who do not (control group)?
2. Is there a significant difference in comprehension and retention test scores between the experimental and control groups one month after the completion of the treatment period?
3. Are there significant differences in score gains between upper- and lower-grade elementary students?
4. Do students react favorably to the use of a rubric?

In summary, the findings of the study sought to clarify: (1) the academic benefit of implementing rubrics in elementary general music classroom settings, (2) the effectiveness of rubrics on increasing students' retention test scores, (3) the effect of the use of rubrics at various grade levels, and (4) student perception of rubric effectiveness.

Delimitations of the Study

The research was conducted in one elementary school located in the southeastern United States. The school serves students in kindergarten through fifth grade, with four classes/sections at each grade level. The treatment period was limited to four consecutive 45-minute lessons, with assessment taking place on the fifth day.

Methodology

Of the 649 students enrolled in the school, 608 participated in the study, which employed a quasi-experimental design with pre-, post-, and retention tests. The teacher/researcher (TR) randomly selected two classes from each grade level to form the experimental group, with the remaining two forming the control group. All students were given a pretest to assess level of knowledge prior to the study. A copy of the assessment rubric, serving as the treatment, was distributed and explained to all students in the experimental groups. The control groups received identical instruction, over the same period of time, without the rubric. All students were given the posttest at the end of their instructional period. The posttest was administered again to all groups as a retention test four weeks after their instructional period was completed. The instructional period consisted of four consecutive 45-minute lessons, with assessment taking place on the fifth day. Tables 1 and 2 display the research design.

Table 1

Research Design

	Pre-treatment	Treatment	End of Treatment	Retention Test (4 weeks after end of treatment)
Experimental Group	Pretest	RUBRICS	Posttest	Follow-up test
Control group	Pretest	X	Posttest	Follow-up test

Table 2

Comparison Between Upper and Lower Grades

	Pretest score	Posttest score	Posttest-Pretest
Lower Grade (K-2)			
Upper Grade (3-5)			

The rubric, based on the Georgia Performance Standards, was created by the TR and reviewed by six certified and experienced music educators. The identified rubric tasks were evaluated by the TR after each assessment event—pretest, posttest, and retention test. An independent two-sample t-test was used to determine if there was a significant difference between

the mean scores of the posttest and the retention test. In addition, an informal student perception survey was administered to ascertain student reaction to using the rubric.

CHAPTER 2

REVIEW OF RELATED LITERTURE

The literature reviewed in this study was gathered from scholarly articles, books, and educational journals related to rubrics. In particular, the following topics were chosen for review: (1) benefits of using rubrics in the general music classroom, (2) features of a good-quality rubric (3) types of rubrics, and (4) rubric construction.

Benefits of Using Rubrics in the General Music Classroom

Rubrics offer many benefits in the educational classroom setting. Several researchers have shown that rubrics improve students' academic achievement in an academic environment (Andrade, 2000, 2005; Bolton, 2006; Schafer, Swanson, Bene & Newberry, 2001). Goodrich (1996) discusses the effects of rubrics, concluding that rubrics support students' learning. Burbridge (1998) suggests that the use of rubric strategies adds a needed depth to assessment in music, asserting that when rubrics are used, the quality of instruction increases. Many music educators would agree that one of the greatest challenges in teaching general music classes is to assess large numbers of students in short periods of time. Rubrics reduce the time teachers spend on assessment because students already know expectations and are able to do well the first time. Using several rubrics of completed assignments, teachers are able to keep a complete record of each student's progress without much extra effort. Rubrics are easily understood at a glance; they are concise and digestible. Levi (2005) concluded that rubrics save time, and have the potential to become an effective part of the teaching and learning process.

Using rubrics as an assessment makes it possible to give prompt feedback. Giving feedback as soon as possible after task completion is most effective in helping students make positive changes in their subsequent work (Rucker & Thomson, 2003). Instructional rubrics provide students with more informative feedback about their strengths and the areas in need of improvement than traditional forms of assessment. Students can also draw their own conclusions about the weaknesses in their work and set out their own plans for improvement as well. A well-written rubric—one that describes the kinds of mistakes students tend to make, as well as the ways in which their work excels—gives them valuable information. In brief, students can learn more about their progress from a rubric than they can learn from a grade.

Ainsworth and Christinson (1998) argue that using a rubric as an assessment system provides many benefits to students, teachers, parents, and administrators alike.

Benefits to Students

1. Expectations are clearly defined and understood by all.
2. The grading process is no longer a mystery to students.
3. Involving students in setting standards results in their understanding the characteristics of quality work.
4. The level of personal responsibility for learning grows.
5. The quality of individual and group work increases over time.
6. The unit content is constantly reviewed through the peer- and self-assessment process.

Benefits to Teachers

1. Clarifying instructional objectives provides structure for students.
2. Focused questions make instructional choices easier.

3. Instructional objectives, instructional activities, and the performance task are aligned.
4. Rubrics provide the teacher with useful feedback about instruction.

Benefits to Parents

1. Parents understand the requirements of the assignment and can better assist their child in its completion.
2. They know ahead of time how their child's work will be assessed.
3. They understand how their child's grade was determined.
4. They know what content their child is actually learning.

Benefits to Administrators

1. Communication among teachers regarding priorities of the instructional program is improved.
2. Schoolwide standards and grade-level expectations are developed.
3. The overall quality of the school's academic performance increases.
4. Consistent standards are established schoolwide for instruction and assessment programs.

Wesolowski (2012) found that rubrics are able to provide the following:

1. clear levels of accomplishment by defining tangible measures of individual achievement;
2. clear indications of what students need to accomplish in the future to improve their individual performance;
3. a learner-centered approach to performing, learning, and assessing;
4. a bridge between student learning and teacher expectations;
5. versatility in adapting to meet the needs of a specific curriculum, student age, ability, style of music, and type of ensemble;

6. a valid and reliable form of individualized assessment and documentation of teacher accountability;
7. a quantitative means for evaluating and scoring qualitative, performance-based tasks;
8. a means for clearly implementing content standards and course objectives into the assessment process;
9. valuable information for parents on their child's progress and needs for improvement.

Features of a Good-Quality Rubric

Assessing and evaluating student work has become more exacting with the shift of focus to authentic learning experiences based on state and national standards. Research on the effects of using rubrics instructionally in the classroom has led to increased emphasis on assessment practices. A good-quality rubric can be used as an assessment to follow students' improvement throughout the program, evaluate their needs, and support their ongoing progress.

According to Arter and Chappuis (2006), a good-quality rubric must be understandable, be aligned with standards and learning targets it is intended to measure, be illustrated with samples of student work, be concise, be stated in a way students can understand, be easy to use, be worded in a positive manner; match the assignment/task, define various levels of performance, and include the same features across various levels of performance. Moreover, a good-quality rubric contains detailed feedback and gives all the necessary details about how and where the assignment did or did not achieve its goal (Stevens & Levi, 2005). Brinko (1993) found that feedback was most effective when it contained as much information as possible rather than simply evaluating the level of the work.

Types of Rubrics

Rubrics may assume several different forms. It is common practice to distinguish between *holistic* and *analytic* rubrics and between *task-specific* and *general rubrics*.

Holistic and analytic rubrics

The distinction between holistic and analytic rubrics is discussed by Gordon (2002), Quinlan (2006), and Wesolowski (2012). If all parts of the instrument are considered together to arrive at a single judgment of product worth, the rubric is considered to be *holistic*. Holistic rubrics provide a single score. The evaluator matches the descriptors of the scale to his or her overall impression of the performance. Generally, a holistic rubric is written in a manner that is generic and simple enough to adapt to other performance situations. An advantage of holistic rubrics is that they are easy and fast to use. However, they do not provide detailed information on an overall performance assessment. The score of the holistic rubric will not provide the student with specific feedback on the teachers' choice of grade.

A rubric is considered to be *analytic* if the performance rating is derived by looking at each of its relevant component parts. An analytic rubric contains more than one dimension of evaluative criteria. The multiple criteria are matched with multiple descriptors and the teacher's feedback, and scoring is based on each of these individual dimensions. Because of the assessment by multiple criteria, the analytic rubric provides more information than does the holistic rubric. A benefit of analytic rubrics is the wealth of specific, individualized assessment information that can be of great value to students, parents, and teachers.

Task-specific and general rubrics

Task-specific rubrics are those specific to the performance task (Arter & Chappuis, 2007). There is a unique scoring guide tailored to each individual task. Task-specific rubrics contain the answer to a problem, or explain the reasoning students are supposed to use, or list facts and concepts students are supposed to mention. General rubrics are those that can be used to judge quality across similar tasks. The descriptions of performance are general, so students learn general qualities and not isolated, task-specific features.

Constructing a Rubric

Before creating a rubric, the teacher needs to take the time to reflect on what he or she wants from the students, why the assignment was created, what happened the last time it was given, and what the teacher's expectations are (Hart, 1994). Constructing a rubric requires reflection on overall class objectives, the assignment itself, its purposes, the task objectives, and students' prior knowledge, as well as the teacher's own previous experience with this type of assignment.

After the reflection stage is completed, the learning objectives should be identified—the essential knowledge and understanding students need to acquire based on district guidelines and state frameworks. Lessons are then designed around the learning objectives and a performance task is selected that accurately measures student performance in relation to the objectives. Hart (1994) writes, "Designing a performance task is a challenge. Good tasks grow out of the curriculum. They are feasible in terms of available time and resources. They are inviting to both teachers and students. And the results can be scored and reported in ways that satisfy students, teachers, parents, administrators, as well as district or state testing directors (p. 41)."

The performance task having been selected, the next step is to determine the type of rubric appropriate for assessment of that task (Gordon, 2002; Quinlan, 2006; Wesolowski, 2012). Teachers need to be aware of whether they are teaching a new skill or further developing a previously taught skill. It is necessary to consider the complexity of the skill and how many component parts are to be assessed. If there is a certain level of overlap, a holistic rubric may be a better fit. An analytic rubric can be used when there is diversity in the learning outcomes.

After the teacher determines the type of rubric to be used, the next stage is to define the range and degrees of proficiency of performance scale levels. There are unlimited labels that can serve to categorize the levels of proficiency achieved by the students, for example, (1) beginning, (2) developing, (3) accomplished, and (4) exemplary. The majority of performance assessment rubrics tend to contain four categories (Brophy 2000; Hickey 1998). Most students should be at level 3, a few at levels 4 and 2, and in rare circumstances a couple of students at level 1. Levels 1 and 2 should not be used as negative feedback.

Writing a Rubric

Odegaard (2009) recommends approaching the writing of a rubric in a series of steps:

1. Level 3 is written first (expected performance). The remaining levels are written in this order: 4 next, then level 2 and level 1 last.
2. Each rubric criterion should focus on a single concept or skill. Combining several items to evaluate in one section makes assessment confusing.
3. The most important criterion is listed at the top of the rubric, with other criteria following in order of decreasing importance.
4. Positive statements encourage lower-level students to work toward mastery.

5. Criteria that count mistakes should not be included.
6. Also to be avoided are subjective words such as “good,” “acceptable,” or “poor.”
7. It is advisable to write a common sentence stem that can be changed for varying levels.
8. The criteria should be numbered.

The next stage is to define appropriate task expectations and meaningful descriptors for each criterion performance level. The descriptors should be written as clearly and concisely as possible. The descriptors should be detailed enough to limit subjectivity yet concise enough to avoid confusion. Edwin Gordon states that the more descriptors that are included for each dimension, the more reliable the rubric will become (Odegaard, 2009).

The final stage is to choose an appropriate scoring scale with clearly defined cut points. It is important to construct a logical and easy-to-understand scoring guide that correlates to the entirety of the performance assessment. There are generally three methods to choose from for grade construction: (1) assign each scale level a point value, and sum each descriptor to a determined scoring guide; (2) assign each scale level a point value while weighting each descriptor according to importance within the total assessment; or (3) assign a letter or numeric value for each level. In each case, a clearly defined scoring method must be labeled and maintained throughout the assessment process.

There are five models of rubric construction (Anderson, 1998).

1. The Presentation Model: This is the most commonly used rubric construction model. In the Presentation Model, the teacher does all the work and makes all the major decisions. The teacher determines the weight that will be given to each dimension, decides on a scale, and describes the highest and the lowest level of performance. The teacher begins by

passing out the rubric before the students have started the assignment and asking them to read it.

2. The Feedback Model: The Feedback Model differs from the Presentation Model only in that it can be changed through student feedback. Before the teacher finalizes the rubric, the students are presented with a completed rubric and are given the option to make edits, offer ideas, and ask questions.
3. The Pass-the-Hat Model: The Pass-the-Hat Model gives the students a maximum amount of flexibility and creativity in developing task expectations for a grading rubric, while allowing the teacher to retain considerable control over the final product. In this model, the teacher does not create a rubric in advance but helps the students to create part of their own rubric during class time. The students start with the teacher-created assignment and list possible expectations for this assignment. The teacher then groups these expectations into dimensions, labels them, and applies these to the rubric grid.
4. The Post-it Model: The Post-it Model gives greater control to the students, who create not simply some of the descriptions of the dimensions, but the dimensions themselves. The Post-it Model is begun in the same way as the Pass-the-Hat Model. The teacher distributes the Post-its and asks students to write down two or three things they think should define an excellent fulfillment of the assignment. The teacher, however, does not collect and group these expectations into dimensions. Instead, the teacher helps students organize them by putting related items together in the same area. Once all the groups are completed, the teacher collects them and creates the final rubric.
5. The 4X4 Model: This model allows for student input at all stages of the rubric construction process. The teacher's role is limited to setting the assignment, explaining

what the finished rubric will look like in a generic sense, and facilitating the creation of the rubric by the students. The students fully participate in all stages of creating the final rubric.

Summary

The implementation of rubrics as an assessment tool can enlighten the teaching and learning processes, and provide a foundation for understanding the focus of assessment. Good-quality rubrics assist teachers in defining complex learning targets, ensure that judgments about student work are consistent over time, aid in planning instruction, and give descriptive feedback to students. Rubrics benefit students by orienting them to what they are expected to accomplish, efficiently helping them analyze and evaluate their own products and performances. Finally, rubrics may be used to communicate educational goals precisely and concretely to teachers, students, parents, and administrators.

CHAPTER 3

METHODOLOGY

The problem addressed in this study was to determine if the use of a teacher-made rubric makes a positive impact on students' academic achievement in the elementary general music classroom. While commonly used assessment techniques, such as pencil and paper tests, group performance adjudications, portfolio entries, informal observations, textbook series music tests, and participation draw focus to the learning objective, the rubric, unlike many other measures, assesses on a performance level. A good rubric provides clear descriptions of what students need to accomplish in the context of a specific activity.

Participants

The study site, a K-5 school located in the southeastern United States, enrolled 626 students at the time of the study. The student ethnicity distribution was 68.9% Hispanic, 24.4% African-American, 3.9% White, 1.8 Multiracial, and 0.9% Asian. Ninety-three percent were eligible for free or reduced price lunch, 11% receive special needs services, 70% are English second-language learners, and 48% qualify for ESOL instruction. Participants included gifted students, inclusion students, and regular education students. The age range of the students was 5-11 years.

Of the total student population of 626, the number of participants was reduced to 608 due to student absences and transfers. Intact classes were divided into two groups: an experimental group, aware of the assessment rubric (Group A, $n = 300$); and a control group, unaware of the

rubric (Group B, *n* = 308). The experimental group consisted of 57 kindergarten students, 58 first grade students, 44 second grade students, 53 third grade students, 51 fourth grade students, and 37 fifth grade students. The control group consisted of 60 kindergarten students, 59 first grade students, 45 second grade students, 50 third grade students, 47 fourth grade students, and 47 fifth grade students. Table 3 shows the number of participants and total enrollment by grade level.

Table 3

Study Participants and School Enrollment by Grade Level

Grade Level	Experimental Group (Participants/School Enrollment)	Control Group (Participants/School Enrollment)
Kindergarten	57/58	60/60
1st grade	58/61	59/61
2nd grade	44/46	45/47
3rd grade	53/56	50/52
4th grade	51/53	47/48
5th grade	37/37	47/47
Total	**300**/311	**308**/315

The pretest, conducted by the TR, required the students to sing and play a song. The song, chosen by the TR, was unfamiliar to almost all of the students; the selection was different for each grade level but the same for both groups within each grade level. The scoring instruments for the pretest was the same TR-authored rubrics that were used for scoring the posttest. No significant differences were found between the experimental and control groups on pretest scores.

Following analysis of the pretest, all students received singing/playing instruction on the appropriate selection for 45 minutes a day for four days. For the experimental group only, the teacher distributed copies of the rubrics and explained expectations prior to each lesson. During the lesson students were allowed to consult the rubrics as they wished. Students in the control group received the same musical instruction but did not receive copies or an explanation of the rubric. The posttest, scored by the rubrics, was administered on the fifth day.

Materials and Instruments

The assessment instruments included a TR-authored rubric for the singing test and another for the instrument-playing test, both of which underwent critical review by six experienced certified elementary music teachers. The achievement targets of the instructional rubrics were based on the National Standards for Music Education: (NS1) Singing, alone and with others, a varied repertoire of music; and (NS2) Performing on instruments, alone and with others, a varied repertoire of music.

Repertoire was selected from *Making Music*, published by Silver Burdett, and *Kodály Today* by Houlahan and Tacka. Selections were made on the basis of melodic difficulty and rhythmic content. (See Appendices A and B.) The Kodály method was chosen as the vocal methodology because of its sequential presentation of pitch and rhythm concepts and its emphasis on the skills of listening, singing, and reading. Songs and rhythm/pitch requirements for each grade level appear in Appendix A. Accompaniment patterns, employing harmonic techniques such as drones (bourdons) and ostinatos, were taught using Orff instruments. All students were tested on their ability to play short accompaniment patterns according to their instructional level. (See Appendix C.)

The rubrics for singing and playing instruments were administered as the pre-, post-, and retention tests. These assessments were aligned with National Achievement Standards: Students sing independently, on pitch and in rhythm, with appropriate timbre, diction, and posture, and maintain a steady tempo (Standard 1a); Students perform easy rhythmic, melodic, and chordal patterns accurately and independently on rhythmic, melodic, and harmonic classroom instruments (Standard 2b).

Five assessment criteria were employed on the singing rubric: melodic accuracy, rhythmic accuracy, vowel and consonant pronunciation, tone and breath support, and posture. The rubric for instrumental performance used four assessment criteria: melodic accuracy, rhythmic accuracy, posture, and ensemble precision. Both rubrics used a rating scale of 1-4 accompanied by verbal descriptors of the performance levels for each criterion: excellent, very good, good, and beginning proficiency. (See Appendix D.)

Data Collection, Processing, and Analysis

Data collection began upon approval from the University of Georgia Institutional Review Board. A letter of consent was sent to parents asking permission for their child's participation in the study. (See Appendix F.) The letter described the purpose of the research, procedures for collecting data, and provisions for confidentiality.

In the interest of controlling extraneous variables due to environment, the same room was used for all groups during instructional and testing periods and while directions were given. The TR administered the pretest to both the experimental and the control groups. All groups subsequently received 45 minutes of instruction per day for four consecutive days dedicated to learning to sing the assigned song and play the instrument accompaniment patterns. Prior to the

instructional sequence, the purposes and application of the rubric assessment were explained to the experimental groups and copies were distributed. The rubric was also displayed for the experimental groups on an interactive whiteboard during lessons. The rubric was not distributed or explained to the control groups.

As appropriate to the instructional level, the following activity sequence was followed in presenting singing lessons to both groups. Students: (1) listened to the song while tapping the beat; (2) learned the song by rote, one phrase at a time, using unstructured hand movements to outline melodic contour or, as level-appropriate, the Curwen hand signs to indicate exact pitches; (4) sang the song using rhythm syllables and then clapped the rhythm. Throughout each lesson, correct posture, breath support, and tone quality were emphasized, with mistakes in these areas corrected as they occurred. Students practiced the song in class for 20 minutes each of the four days. They were not asked to practice outside of class. The posttest was administered on the fifth class day.

For the final assessment, each student was given the starting pitch and asked to sing the song in front of the class without accompaniment. The TR scored each test by circling mistakes on the music as the student sang the song and then assigning each error to a rubric category: melodic accuracy, rhythmic accuracy, text, breath and tone support, and posture. (See Appendix D.)

With adjustments made for instructional levels, the following basic activity sequence was used in presenting the instrumental lessons to both groups. Prior to the playing lessons, the experimental group was given a rubric with the description of each criterion; the control group received no rubric. After receiving basic technical instruction on the instrument (body position, mallet grasp, striking movement), students (1) tapped the designated rhythm pattern (patchen); (2)

visually followed the melodic pattern displayed on the board as TR pointed to the notation; (3) using index fingers, practiced playing motions without instrument contact; and (4) transferred playing motions to prepared instruments (unneeded bars removed) and practiced playing in ensemble. Students practiced 20 minutes per day in class for four consecutive days. As with the singing instruction, the posttest was administered on the fifth day of class. The playing posttest followed the same format as the singing test, except that the students performed in groups of four. The TR circled mistakes on the music score as they occurred, and subsequently assigned each error to a rubric category: notes, rhythmic accuracy, posture, and ensemble precision. (See Appendix D.)

Four weeks after the posttest was administered, all groups were asked to take an unannounced retention test (identical to the posttest). The students were told at this time that the objective of the experimental study was to see whether incorporating a rubric into the music class promoted delayed retention learning. They were asked to do their best and told that their scores did not affect their grades. Following the retention test, students in grades three through five completed the survey on student perception of rubric effectiveness. (See Appendix E.)

In this study, statistical data were used to represent academic achievement. The singing and playing test measured student achievement in music. The two-sample t-test was used to determine whether a statistical significance existed between the two means. Statistical significance was established at .05 level of probability to ensure all findings were statistically significant and relevant for use in the research findings.

CHAPTER 4

FINDINGS

The purpose of this study was to determine whether the use of a teacher-made rubric as an instructional tool aids student comprehension and retention of materials related to singing and playing an instrument. Students were divided into two groups, the experimental group and control group. Data were collected using a quantitative pretest-posttest control group design. This chapter presents the results of the analyses performed to address the following research questions:

1. Is there a significant difference between the immediate pretest and posttest scores of students who use a rubric throughout the instructional period (experimental group) and students who do not (control group)?
2. Is there a significant difference in comprehension and retention test scores between the experimental and control groups one month after the completion of the treatment period?
3. Are there significant differences in score gains between upper- and lower-grade elementary students?
4. Do students react favorably to the use of a rubric?

Singing Test Results

For the singing data set, 626 students took part in the project initially. However, 18 were absent from the school on one or both of the test days, and another 18 students were familiar with the song they were given before the experiment started. Therefore 590 observations—299 in the

control group and 291 in the experimental group—were used for the statistical analysis. Table 4 shows the number of students involved in the study per grade level.

Table 4

Number of Participants Involved in Singing Test

Grade		
Grade	N	Percent
1	106	17.97
2	82	13.90
3	103	17.46
4	98	16.61
5	84	14.24
K	117	19.83

Group		
Group	N	Percent
Control	299	50.68
Experimental	291	49.32

A two sample independent t-test was used to determine if there was a significant difference between the two mean scores in the posttest. The values of the posttest mean scores, *t*-value, *df* and *p*-values for sing are displayed in Table 5. The mean posttest score of students in control group is $\overline{X}$=15.9398, and the posttest score for the experimental group is $\overline{X}$=16.7629. Figure 1 presents graphs and box plots showing the distribution of posttest scores for the control group and for the experimental group on the item of the scale dealing with rubric uses.

The output also gives the estimated difference of -0.8231, which is to say, the control group is expected to be 0.8231 points lower than the experimental group in the posttest singing scores. According to this data, it can be concluded that there is a significant difference in the posttest score between the experimental and control groups, in favor of the experimental group.

Table 5

Comparison of Two Independent Sample *t*-test Scores for Singing Test

Group	N	Posttest Mean	*t*-value	df	*p*-value
Control	299	15.9398	-4.02	588	.0001
Experimental	291	16.7629			
Diff (1-2)		-0.8231			

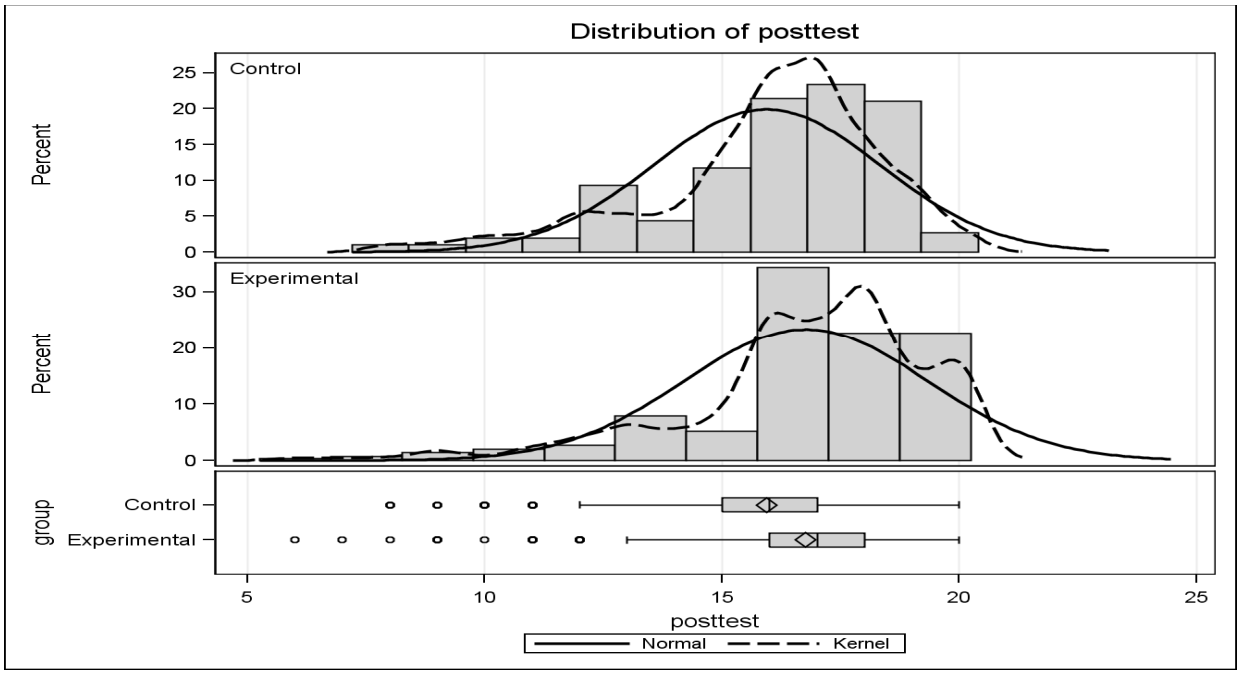


Figure 1. Distribution of Posttest Scores.

Following the analysis of the mean scores, confidence intervals were used to calculate a measure of the variability in the data. The upper and lower bound 95% confidence interval data shows a statistical significance in the average posttest singing scores between control and experimental groups. The 95% confidence interval is (-1.2250, -0.4212), i.e., with 95% confidence there is a true difference in the average posttest singing scores between control and experimental groups. The minus sign indicates that the control group is lower in average scores. Therefore, when comparing the two groups, the experimental group outperformed the control group. (See Table 6.)

Table 6

Two Independent Sample *t*-test

Group	*N*	Posttest Mean	StdDev	Std Error	Min	Max	95% CL Std Dev	
Control	299	15.9398	2.4045	0.1391	8.0000	20.0000	2.2260	2.6144
Experimental	291	16.7629	2.5654	0.1504	6.0000	20.0000	2.3726	2.7927
Diff (1-2)		-0.8231	2.4852	0.2046			-1.2250	-0.4212

Similarly, for the second research question, a two sample independent *t*-test was used. Research question two addressed whether any significant difference would be found in the average retention scores between the control and experimental groups. Table 7 is the retention score descriptive by the control and the experimental group. The specific data include group, *N*, and retention mean scores, *t*-value, *df* and *p*-value. The mean score of students' retention score in control group was $\overline{X}$=13.8361, and the retention score in experimental group was $\overline{X}$=15.1340. Figure 2 presents graphs and box plots showing the distribution of retention scores for the control group and for the experimental group on the item of the scale dealing with rubric uses.

The output also gives the estimated difference of – 1.2979, which is to say, the control group is expected to be 1.2979 points lower than the experimental group in the retention singing scores. This result shows that the treatment used on the experimental group enhanced students' achievement. According to this data, it can be concluded that there is significant difference in the retention score between the experiment and control groups.

Table 7

T-test for Variable Retention for Singing Test

Group	N	Retention Mean	t-value	df	p-value
Control	299	13.8361	-5.73	588	.0001
Experimental	291	15.1340			
Diff (1-2)		-1.2979			

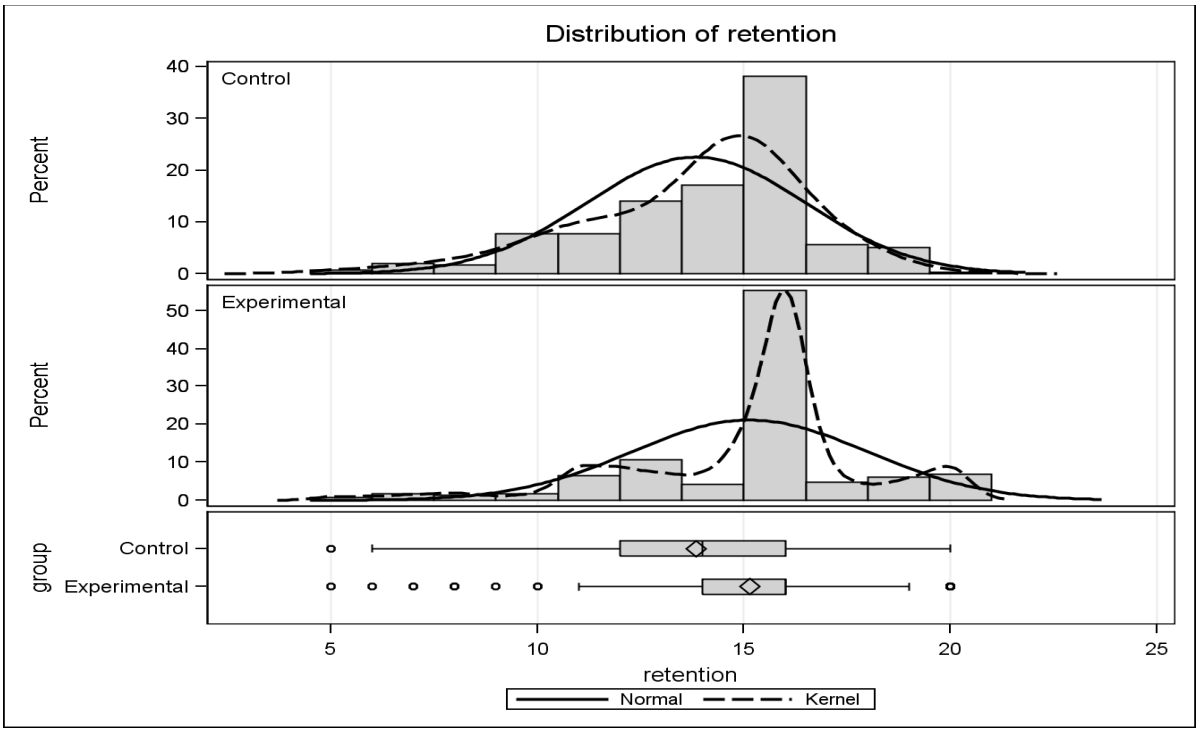


Figure 2. Distribution of Retention Test.

The upper and lower bound 95% confidence interval data shows a statistical significance in the average retention singing scores between the control and experimental groups. In comparing the two group data, the 95% confidence interval was -1.7426 and -0.8532 (see Table 8). This indicates that the experimental group scored higher than the control group.

Table 8

Two Independent Sample *t*-test

Group	N	Retention Mean	StdDev	Std Error	Min	Max	95% Interval Lower Bound	Confidence for mean Upper Bound
Control	299	13.8361	2.6565	0.1536	5.0000	20.0000	13.5338	14.1385
Experimental	291	15.1340	2.8417	.01666	5.0000	20.0000	14.8062	14.8062
Diff (1-2)		-1.2979	2.7494	0.2264			-1.7426	-0.8532

For the third research question (Are there significant differences in score gains between upper- and lower-grade elementary students?) a linear model was used for the analysis. Grades K-2 were categorized as lower grades and 3-5 as upper grades. Table 9 shows that that there were 285 upper grade students (48.31% of the total population) and 305 lower grade students 51.69%).

Table 9

Number of Students in Lower and Upper Grades

Grades	*N*	Percent
Upper	285	48.31
Lower	305	51.69

Along with the linear model for upper and lower grades, the combination of different levels of factor grade and group were analyzed (See Table 10). The means for the lower grades were 15.7933333 for the experimental group and 15.6709677 for the control group; for the upper grades the means were 17.7943262 for the experimental group and 16.2291667 for the control group. Although the experimental group showed more improvement than the control group in both upper and lower grades, differences were significant only for upper-grade students (Figure 3). For lower grade students, with *p*-value 0.6536, there was no significant difference in posttest singing scores between control and experimental groups (see Table 11). Thus it can be said that the use of a rubric as an instructional tool produced greater gains in singing scores for upper grade students.

Table 10

Model for Singing Posttest by Grade and Group

Grades	Group	Posttest MEAN
Lower	Experimental	15.7933333
Lower	Control	15.6709677
Upper	Experimental	17.7943262
Upper	Control	16.2291667

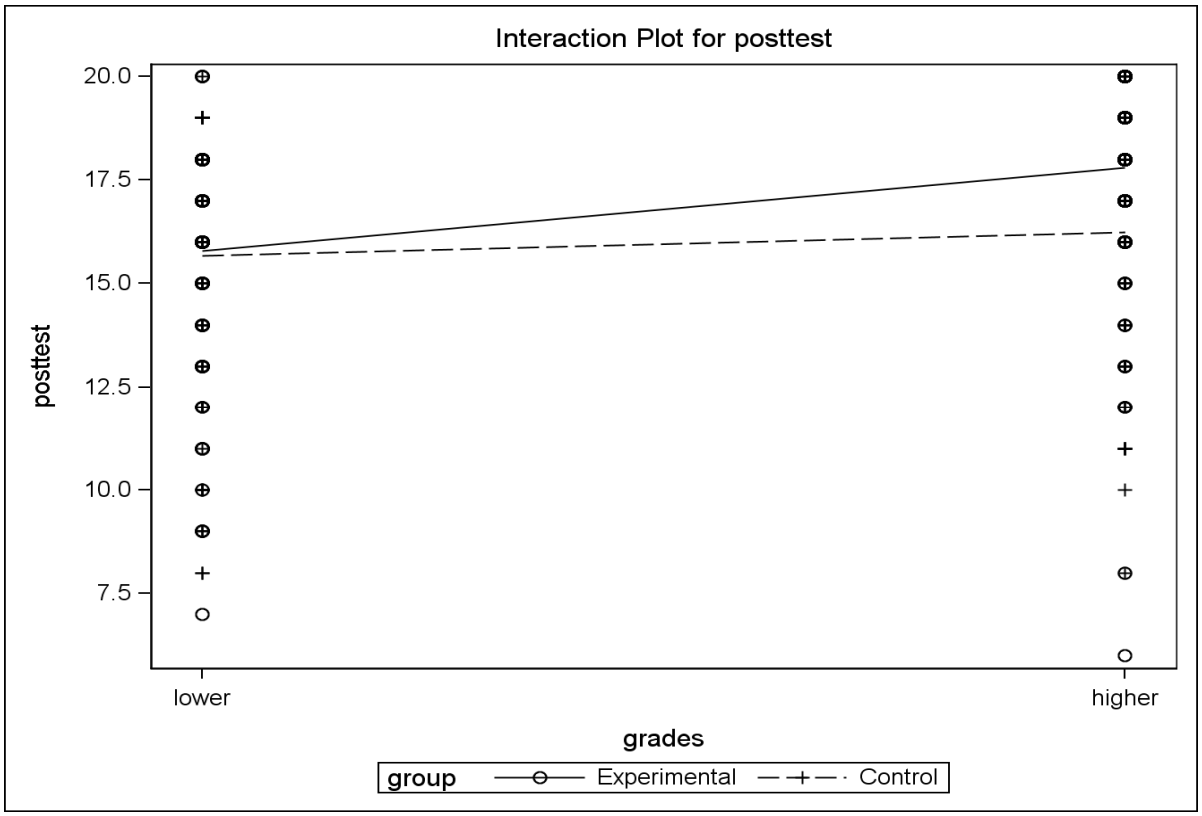


Figure 3. Interaction Plot for Posttest

Table 11

Grade Group Effect Sliced by Grades for Posttest

grades	DF	Sum of Squares	Mean Square	F Value	Pr > F
lower	1	1.141410	1.141410	0.20	0.6536
higher	1	174.523530	174.523530	30.83	<.0001

Instrument Playing Test Results

For the instrumental data set, as for the singing set, 626 students took part in the project initially. However, 18 were absent from the school on one or both of the test days, leaving a total N of 608: 300 in the experimental group and 308 in the control group. (Unlike the singing data set, no scores were removed from the instrumental data because of previous knowledge of the material.) Table 12 shows the number of students involved in the study by grade level and group.

Table 12

Number of Participants Involved in Playing Test by Grade Level and Group

Grade	Frequency	Percent
1	117	19.24
2	89	14.64
3	103	16.94
4	98	16.12
5	84	13.82
K	117	19.24

Group	Frequency	Percent
Control	308	50.66
Experimental	300	49.34

The same analysis was used for instrument test as for the singing test scores. For the first research question, the data showed a significant difference in posttest scores between the control and experimental groups. The mean score of control group students' posttest score in control group was $\overline{X}$=12.1526, and the mean posttest score for the experimental group was $\overline{X}$=12.8233, a difference of 0.6707 points in favor of the experimental group. (See Table 13.) The 95% confidence interval was -0.9790 and -0.3624. (See Table14.) Figure 4 presents graphs and box plots showing the distribution of posttest scores for the control group and for the experimental group. Based on these data, the experimental group made significantly higher posttest scores than the control group.

Table 13

Comparison of Two Independent Sample *t*-test Scores for Playing Test

Group	N	Posttest Mean	*t*-value	df	p-value
Control	308	12.1526	-4.27	606	0.0001
Experimental	300	12.8233			
Diff (1-2)		-0.6707			

Table 14

Two Independent Sample *t*-test

Group	*N*	Posttest Mean	StdDev	Std Error	Min	Max	95% Interval Lower Bound	Confidence for mean Upper Bound
Control	308	12.1526	2.0145	0.1148	5.0000	16.0000	11.9267	12.3785
Experimental	300	12.8233	1.8503	0.1068	4.0000	16.0000	12.6131	13.0336
Diff (1-2)		-0.6707	1.9352	0.1570			-0.9790	-0.3624

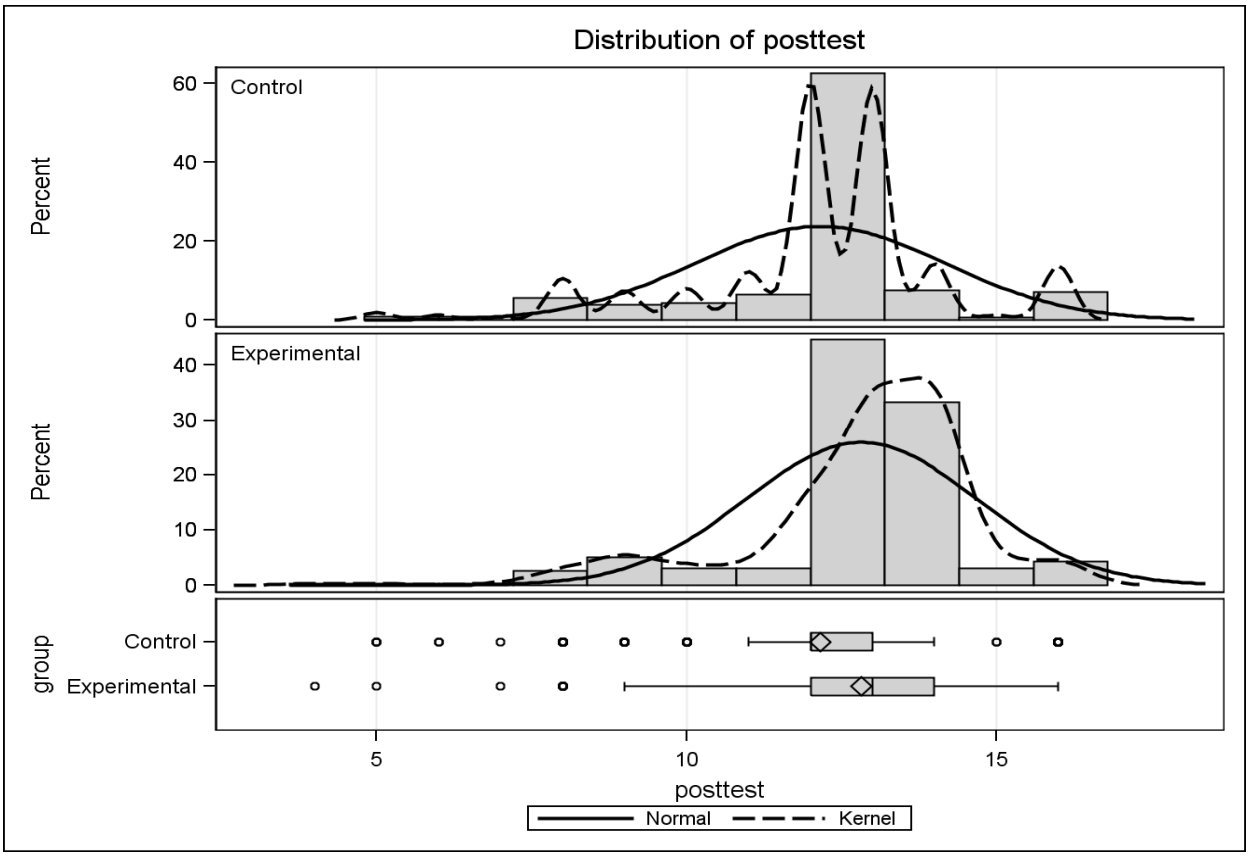


Figure 4. Distribution of Posttest

The second research question addressed retention of gains. The mean retention score for instrumental playing was $\overline{X}$ =9.9838 for the control group and $\overline{X}$=11.7667 for the experimental group (see Table 15). Figure 5 presents graphs and box plots showing the distribution of retention scores for the control group and for the experimental group on the item of the scale dealing with rubric uses.

The output also gives the estimated difference of –1.7829, which is to say, the control group was 1.7829 points lower than the experimental group in retention playing scores. The 95% confidence intervals were –2.0760 and –1.4898, showing that the experimental treatment enhanced students' achievement. According to this data, it can be concluded that there is

significant difference between the experimental and control groups in instrumental retention scores.

Table 15

T-test for Variable Retention for Instrumental Playing Test

Group	N	Retention Mean	t-value	df	*p*-value
Control	308	9.9838	-11.91	606	.0001
Experimental	300	11.7667			
Diff (1-2)		-1.7829			

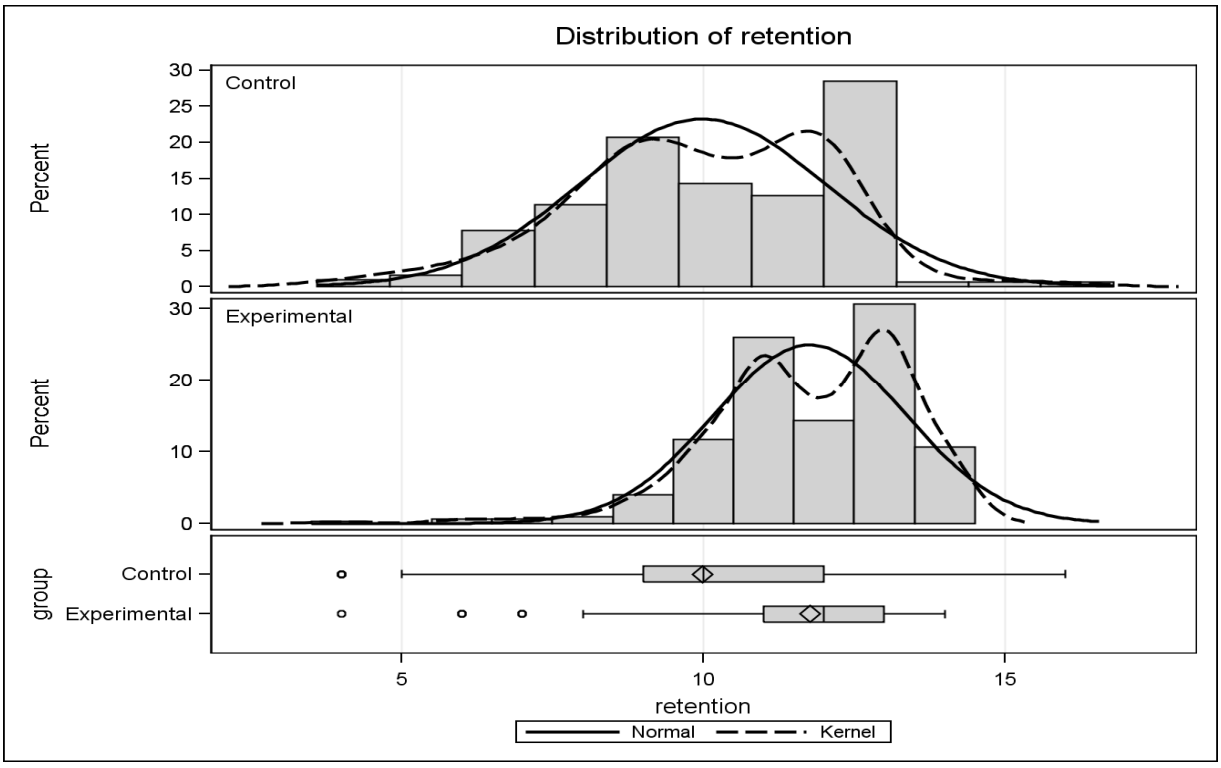


Figure 5. Distribution of Retention Scores for Instrumental Playing Test.

Table 16

Two Independent Sample *t*-test

Group	*N*	Retention Mean	StdDev	Std Error	Min	Max	95% Interval Lower Bound	Confidence for mean Upper Bound
Control	308	9.9838	2.0585	0.1173	4.0000	16.0000	9.7530	10.2146
Experimental	300	11.7667	1.5981	0.0923	4.0000	14.0000	11.5851	11.9482
Diff (1-2)		-1.7829	1.8458	0.1497			-2.0760	-1.4898

For the third research question, a linear model was used for the analysis of the playing test, as with the singing test. Table 17 shows that that there were 323 upper grade students (53.12% of the total population) and 285 lower grade students (46.88 %).

Table 17

Number of Students in Lower and Upper Grades

Grades	*N*	Percent
Upper	323	53.12
Lower	285	46.88

Along with the linear model for upper and lower grades, the combination of different levels of factor grade and group were analyzed. For the lower grades the means were 12.591195 for the experimental group and 11.432927 for the control group. For the upper grades the means were 13.085106 for the experimental group and 12.972222 for the control group (See Table 18). As the interaction plot graph in Figure 6 demonstrates, the use of a rubric during instrumental instruction has a more significant effect for lower grade students than for upper grade students. In addition, with *p*-value less than 0.0001, the findings revealed that the rubric method produces significantly higher posttest scores for experimental-group students in the lower grades. For upper grade students, the *p*-value of 0.6069 does not support the existence of a significant difference between control and experimental group. (See Table 19.)

Table 18

Model for Posttest Instrumental Playing by Grade and Group

Grades	Group	posttest LS MEAN	95% Confidence Limits	
Lower	Experimental	12.591195	12.302897	12.879493
Lower	Control	11.432927	11.149058	11.716796
Upper	Experimental	13.085106	12.778959	13.391253
Upper	Control	12.972222	12.669281	13.275163

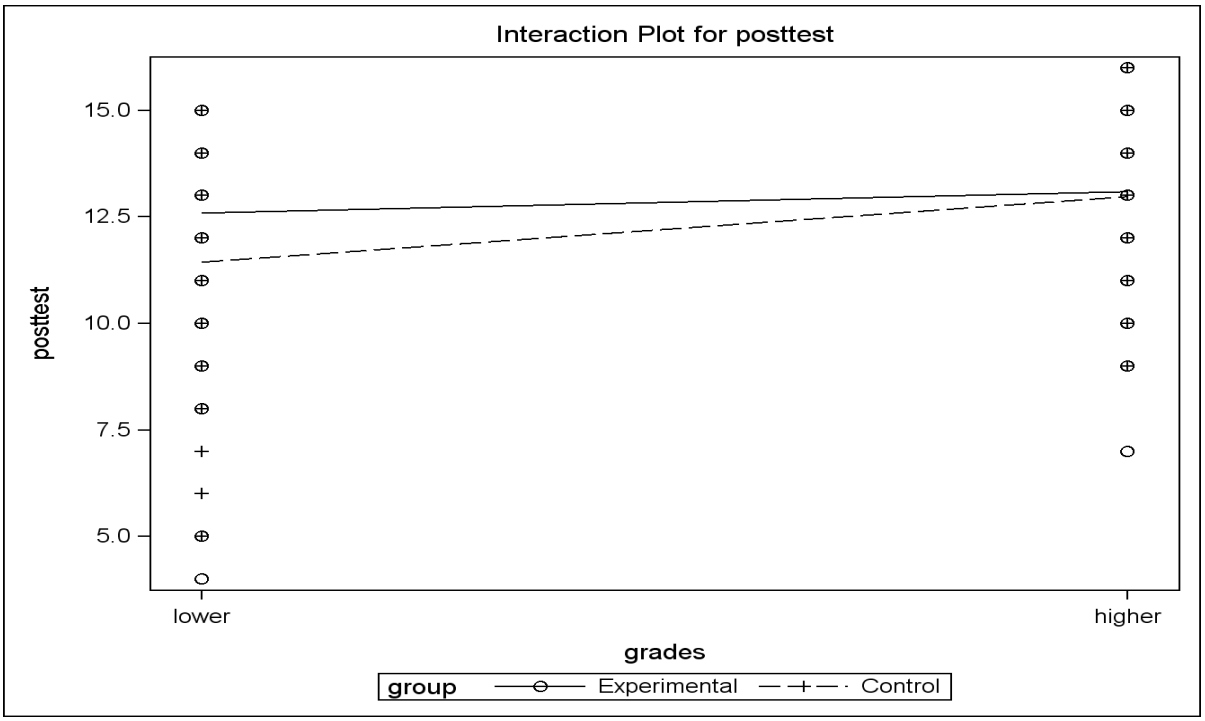


Figure 6. Interaction Plot for Posttest for Playing Instrument

Table 19

Grade Group Effect Sliced by Grades for Posttest

Grades	DF	Sum of Squares	Mean Square	F Value	Pr > F
Lower	1	108.307036	108.307036	31.61	<.0001
Upper	1	0.907826	0.907826	0.26	0.6069

For the fourth research question, 141 upper-grade students in the experimental group completed a seven-question survey about the use of rubrics in instruction (see Table 20). The survey employed a four-point Likert-type (strongly agree, agree, disagree, or strongly disagree). The results shows that overall, rubrics do benefit student learning in specific ways (Table 21). The majority of students responded “agree” or “strongly agree” to each of the seven questions. In

response to the final question—"Overall, did you like to use a rubric?"—93.61% of the students marked "agree" or "strongly agree."

Table 20

Number of Students Participating in the Survey

Grade	N
3	53
4	51
5	37
Total	141

Table 21

Survey Results

About rubrics:	Strongly Agree	Agree	Disagree	Strongly Disagree
1. Do rubrics help you better understand the standards and expectations of lessons?	36.88%	**56.74%**	5.67%	0.71%
2. Did you have fewer questions about the assignment**?**	31.21%	**62.41%**	5.67%	0.71%
3. Do you feel that you were more organized?	41.13%	**52.48%**	4.26%	2.13%
4. Did you work harder than usual to meet the highest grade?	43.26%	**50.35%**	5.67%	0.71%
5. Do you think that using rubrics helped you earn a better grade?	40.43%	**53.19%**	4.96%	1.42%
6. Did you feel that the teacher feedback was clear and consistent?	41.13%	**53.19%**	4.96%	0.71%
7. Overall, did you like to use a rubric?	43.26%	**50.35%**	4.96%	1.42%

Summary

This research study examined the impact of rubric use during instruction on students' learning performance. The present chapter included a review of the research problem, the purpose of the research study, restatement of the research questions, descriptive statistics results, and findings. The researcher was able to answer the research questions based on the analysis of data.

The research method of two sample independent *t*-test and the linear model allowed the researcher to determine if there was a change in scores of each group and allowed to statistically

compare the groups. By analyzing the data, the researcher was able to identify disparities between the control and the experimental group.

The findings revealed that the experimental group, which had the use of the teacher-made rubric throughout instruction, showed significantly higher test scores than the control group in both singing and playing posttests (see Table 22). Thus, the data support the finding that the use of rubrics in elementary general music instruction increases learning performance.

Table 22

Singing Test Comparison

	N	Posttest Mean	Follow-up Mean
Control	299	15.9398	13.8361
Experimental	291	16.7629	15.1340

Playing Test Comparison

	N	Posttest Mean	Follow-up Mean
Control	308	12.1526	9.9838
Experimental	300	12.8233	11.7667

CHAPTER 5

DISCUSSION AND CONCLUSIONS

This chapter includes an overview of the study: problem statement, purpose, methodology, and findings. Finally, conclusions are drawn and recommendations for future research are provided.

The problem addressed in this study was the need to investigate the effect of using an authentic assessment instrument in the elementary music classroom. The purpose of an authentic assessment is to follow students' improvement throughout the program, evaluate their needs, and support their ongoing progress, in contrast to a final summative evaluation.

As the literature affirms, there are several types of authentic assessment that may be used in the elementary classroom. Examples of these measurements are observations, interviews, essays, performance tasks, portfolios, journals, self-evaluations, and rubrics. Effective assessments give students feedback on how well they understand the information and on what they need to improve.

Published research also suggests that rubrics can enrich and facilitate the assessment process both for teachers and for students. The use of a teacher-made rubric can clarify teachers' expectations. Thus, rubrics eliminate or at least decrease ambiguity and uncertainty in the educational process (Ainsworth & Christinson, 1998; Burbridge, 1998; Levi, 2005; Rucker & Thomson, 2003).

The intent of this study was to determine whether a teacher-made rubric aids student comprehension and retention of materials related to singing and playing an instrument. A pretest-

posttest quasi-experimental design was utilized. The content of instruction and the conditions in which the instruction was delivered were the same for both experimental and control groups. The experimental group used a teacher-made rubric throughout instruction, while the control group did not. The descriptive statistics were based on the results of the singing and playing test scores collected at the end of the instructional period and one month later. The setting for this study was a large suburban elementary school in the southeastern United States.

Participants in both experimental and control groups received 45 minutes of similar instruction for four consecutive days, with the experimental group having access to a rubric. On the fifth day, the singing and playing posttests were administered. Four weeks after the posttest, a retention test was administered. The researcher collected the data from all pre-, post-, and retention tests. A two-sample independent *t*-test was used to determine if there was a significant difference between the posttest and retention scores of the two groups. The pretest scores were removed from the statistical measure, since most of students were unfamiliar with the song and the accompaniment pattern for the playing test before the instructional period began.

Conclusions

The research questions for this study were:

1. Is there a significant difference between the immediate pretest and posttest scores of students who use a rubric throughout the instructional period (experimental group) and students who do not (control group)?
2. Is there a significant difference in comprehension and retention test scores between the experimental and control groups one month after the completion of the treatment period?

3. Are there significant differences in score gains between upper- and lower-grade elementary students?
4. Do students react favorably to the use of a rubric?

On the basis of the findings, the first research question can be answered affirmatively: a significant difference in posttest scores was found in favor of the experimental group. Therefore it can be concluded that the use of a rubric has a positive effect on student achievement as measured immediately after instruction.

The second research question can also be answered in the affirmative: retention test mean scores for the experimental group were significantly higher than for the control group. This finding indicates that the use of a rubric can positively affect retention of skills and materials over a longer period of time.

Findings related to the third research question suggest that the use of the use of rubrics may have a greater positive effect on achievement in singing for elementary students in the upper grades than for students in the lower grades. However, on the instrumental playing task, the reverse was true: experimental group students in the lower grades showed more improvement in posttest scores than did upper grade students.

Experimental-group students in grades 3-5 completed the survey addressing the fourth research question. Respondents agreed that rubrics helped them better understand the standards and expectations of lessons. They had fewer questions about the assignment and felt that they were more organized. They reported that they worked harder than usual and agreed that using rubrics helped them earn a better grade. They felt that the teacher feedback was clear and

consistent. These findings indicate that the use of a rubric in the elementary general music classroom can make an important positive impact on student attitudes as well as achievement.

The findings of the present study support earlier research showing that the use of rubrics can improve students' academic achievement (Andrade, 2000, 2005; Bolton, 2006; Schafer, Swanson, Bene & Newberry, 2001). The mean scores of the experimental groups in this study increased significantly on the post- and retention tests. It is therefore concluded that the use teacher-made rubrics in elementary general music classes can promote students' comprehension and retention of material related to singing and playing an instrument.

Limitations

In considering the conclusions and implications of the present study, it is important to keep the following limitations in mind: (1) the research was conducted during a relatively short period of time rather than throughout an entire school year; and (2) the study was conducted in only one elementary school in the southeast. Because of these limitations, possibly affecting validity and reliability, the results of the study may not be generalized to the entire relevant population.

Recommendations for Future Research

This study on rubrics was conducted in a single school in the southeastern United States with elementary students in a general music classroom setting. To ensure generalizability, similar studies should be conducted in a wider variety of instructional and geographical settings. Further, it is recommended that similar research be conducted subjecting the rubric itself to more thorough review by a larger number of evaluators. In addition, increasing the number of qualified

scorers would ensure greater accuracy in evaluating student outcomes. Taken together, these two recommendations would strengthen both the validity and the reliability of future studies of this type.

The third recommendation is to explore involving students in the creation of the rubric. Such involvement could increase students' sense of ownership, not only of the rubric itself but of the learning that the rubric is intended to evaluate. In this study, the teacher wrote a very simple rubric to help students accomplish a given objective. Students taking an active role in developing the scoring criteria, self-evaluation, and goal setting more readily accept that the assessment is adequately measuring their learning (Eppink, 2002; Stauffer, 1999). Including students in the creation of the rubric may encourage them to feel more empowered, so that their learning becomes more focused and self-directed.

REFERENCES

Ainsworth, L., & Christinson, J. (1998). *Student generated rubrics: An assessment model to help all students succeed*. Orangeburg, NY: Dale Seymour Publications.

Andrade, H.G. (2000). Using rubrics to promote thinking and learning. *Educational Leadership, 57*(5), 13-18.

Arter, J., & Chappuis, J. (2007). *Creating and recognizing quality rubrics*. Upper Saddle River, NJ: Pearson Education.

Armstrong, C. (1994). *Designing assessment in art*. Reston, VA: National Art Education Association.

Asmus, E. (1999). Music assessment concepts. *Music Educators Journal*. *82*(2), 21.

Bolton, C. F. (2006). Rubrics and adult learners: Andragogy and assessment. *Assessment Update, 18*(3), 5-6.

Brinko, K. T. (1993). The practice of giving feedback. *Journal of Higher Education*, *64*(5), 575-5, 93.

Brookhart, S. (2013). *How to create and use rubrics for formative assessment and grading*. Alexandria, VA: ASCD.

Brophy, T. S. (2000). *Assessing the developing child musician: Guide for general music teachers.* Chicago: GIA Publications.

Burbridge, A. (1998). Assessment: Pencil, paper & performance tool! *Connection*, *13*(1), 9-11.

Dunbar, L. (2012). Performance assessment of the masses in 30 seconds or less. *General Music Today*, *25*(2), 5.

Epp, D. A. (2012). Teaching in Japan: Using assessment in elementary school music education. *Canadian Music Educator*, *54,* 29-36.

Eppink, J. (2002). Student-created rubrics. *Teaching Music*, *9*(4), 28.

Gordon, E. (2002). *Rating scales and their uses for evaluating achievement in music performance.* Chicago: GIA Publications.

Hale, C., & Green, S. (2009). Six key principles for music assessment. *Music Educators Journal*, *95*(4), 27-31.

Hall, L. (1997). *Strategies for teaching: Guide for music methods classes*. Reston, VA: MENC.

Hickey, M. (1998). Assessment rubrics for music composition. *Music Educators Journal, 85*(4), 26-32, 52.

Koops, L. (2008). Using a musicianship portfolio as assessment of music and education course objectives in a "Music for Elementary Teachers" course. *Journal of Music Teacher Education, 18*(1), 39-54.

Leonhart, A. (2005). Using rubrics as an assessment tool in your classroom. *General Music Today, 19*(1), 10-15.

Linn, R., & Gronlund, N. E. (2000). *Measurement and assessment in teaching.* (8th ed.) Upper Saddle River, NJ: Prentice Hall.

Nutter, K. (1999). Managing assessment in music education: Schools across the country are responding to the need for assessment management. *Teaching Music, 7*(1), 26-31, 59.

Odegaard, D. (2009). *Music curriculum writing 101: Assistance with standards-based music curriculum and assessment writing*. Chicago: GIA Publications.

Quinlan, A. (2006). *A complete guide to rubrics: Assessment made easy for teachers, K-college*. Lanham, MD: Rowman & Littlefield Education.

Scott, S. J. (2001). A task-centered approach to performance-based assessments. *General Music Today, 14,* 10-14.

Shafer, D. W., Swanson, G., Bene, N., & Newberry G. (2001). Effects of teacher knowledge of rubrics on student achievement in four content areas. *Applied Measurement in Education, 14(2),*151-170.

Smith, J., & Kaschub, M. (2009). *Minds on music: Composition for creative and critical thinking*. Lanham, MD: Rowman & Littlefield Education.

Stauffer, S. L. (1999). Beginning assessment in elementary general music. *Music Educators Journal, 86*(2), 25-30.

Stevens, D., & Levi, A. (2005). *Introduction to rubrics: An assessment tool save grading time, convey effective feedback, and promote student learning*. Sterling, VA: Stylus Publishing.

Stiggins, R. J. (2005). *Student-involved assessment for learning.* (4th ed.) Upper Saddle River, NJ: Merrill Prentice Hall.

Taggart, G., & Phifer, S. (1998). *Rubrics: A handbook for construction and use.* Lancaster, PA: Technomic Publishing AG.

Wesolowski, B. (2012). Understanding and developing rubrics for music performance assessment. *Music Educators Journal*, *98*(3), 36-42.

Weber, E. (1999). *Student assessment that works.* Needham Heights, MA: Allyn & Bacon.

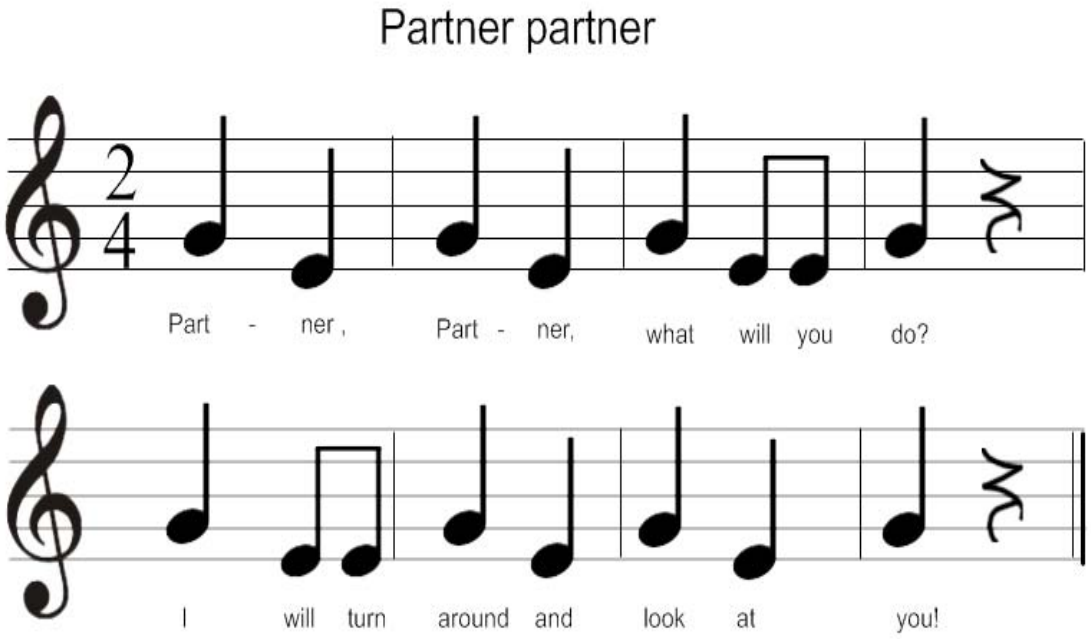
APPENDIX A

Melodic and Rhythmic Content

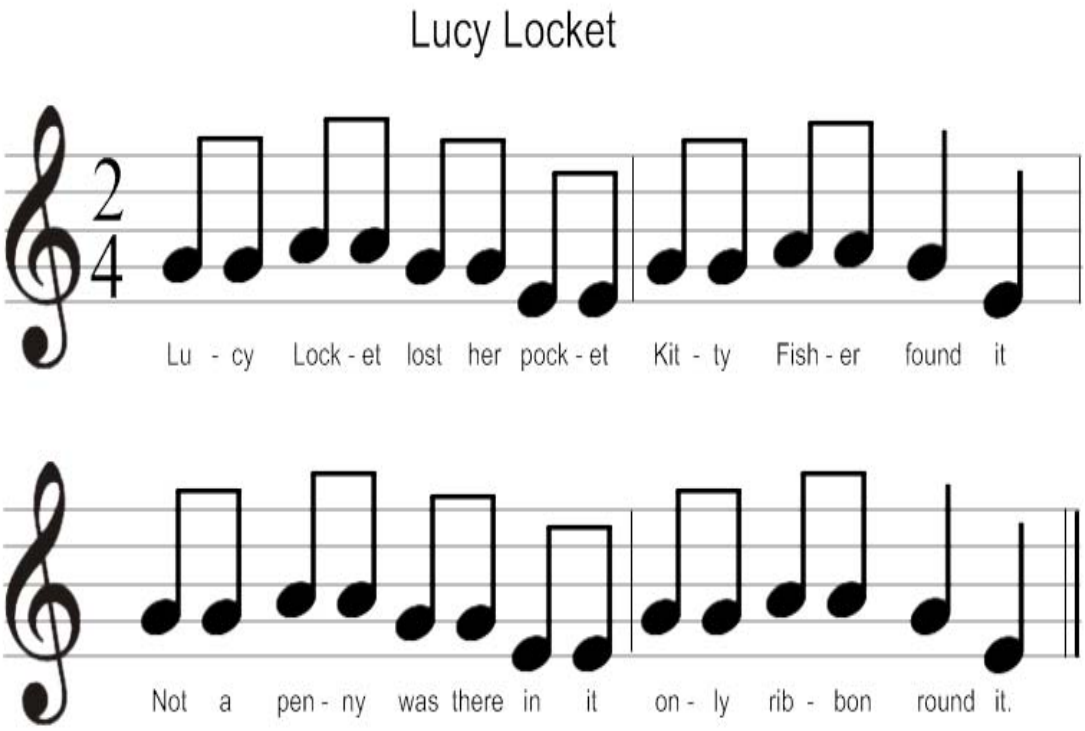
Grade Level	Repertoire	Melodic content	Rhythmic content
Kindergarten	Partner, Partner	*sol-mi*	Quarter note, two eighth-notes
1st grade	Lucy Locket	*sol-la-mi*	Quarter note, two eighth-notes
2nd grade	A-Tisket, A-Tasket	*do-re-mi-fa-sol-la*	Quarter note, two eighth-notes
3rd grade	Li'l Liza Jane	*do-re-mi-sol-la-high do*	Quarter note, two eighth-notes, half note, dotted half note
4th and 5th grade	Nine Hundred Miles	*la-ti-do-re-mi-fa-sol-la*	Quarter note, two eighth-notes, half note, triplet, dotted half note, dotted eighth-note followed by 16th note.

APPENDIX B

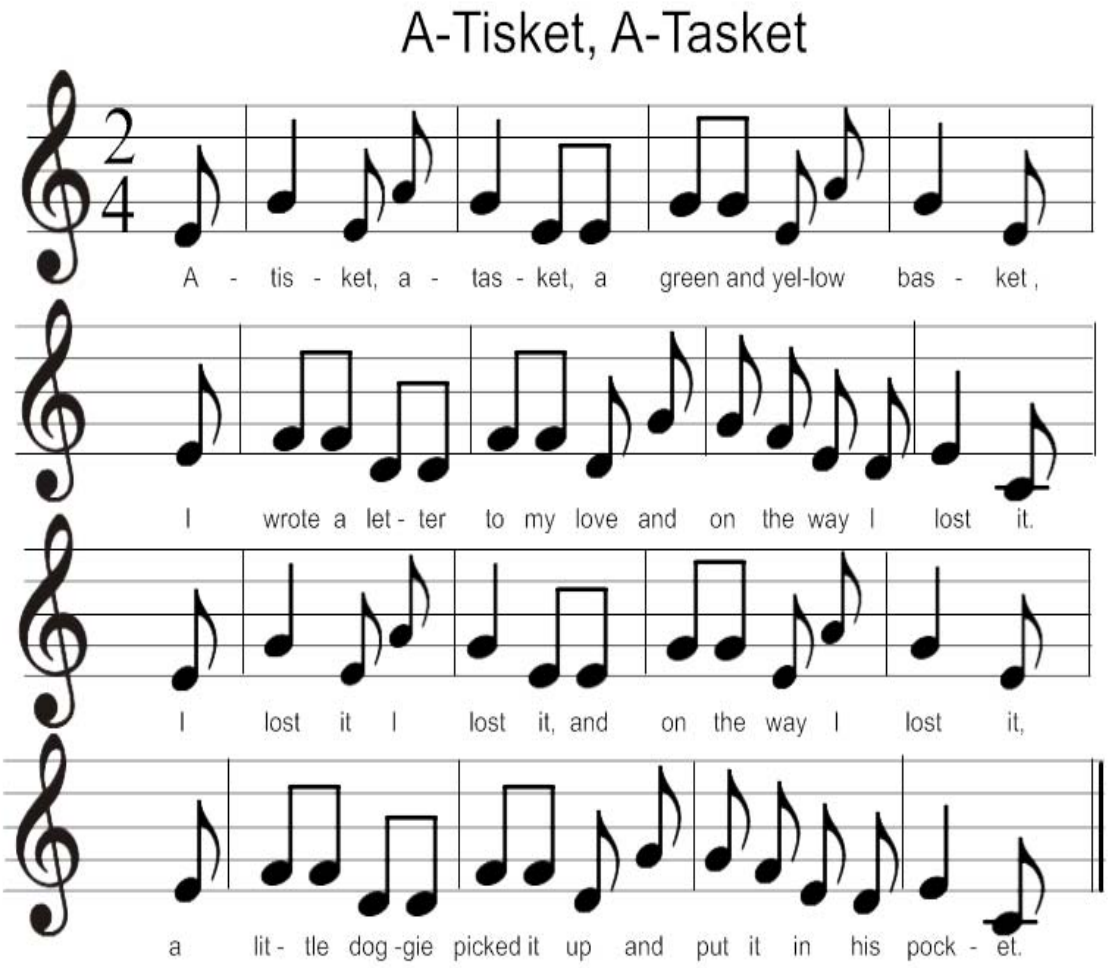
Musical Scores



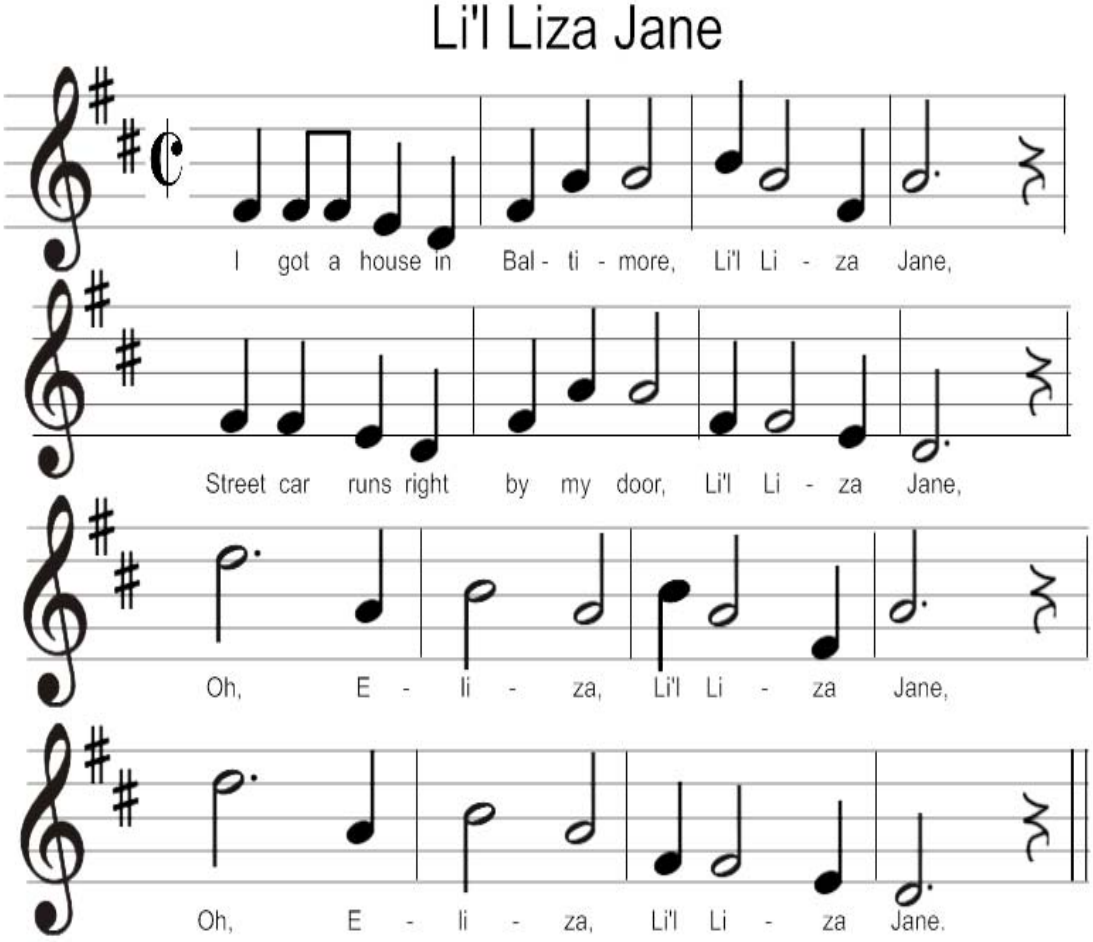
APPENDIX B (CONTINUED)



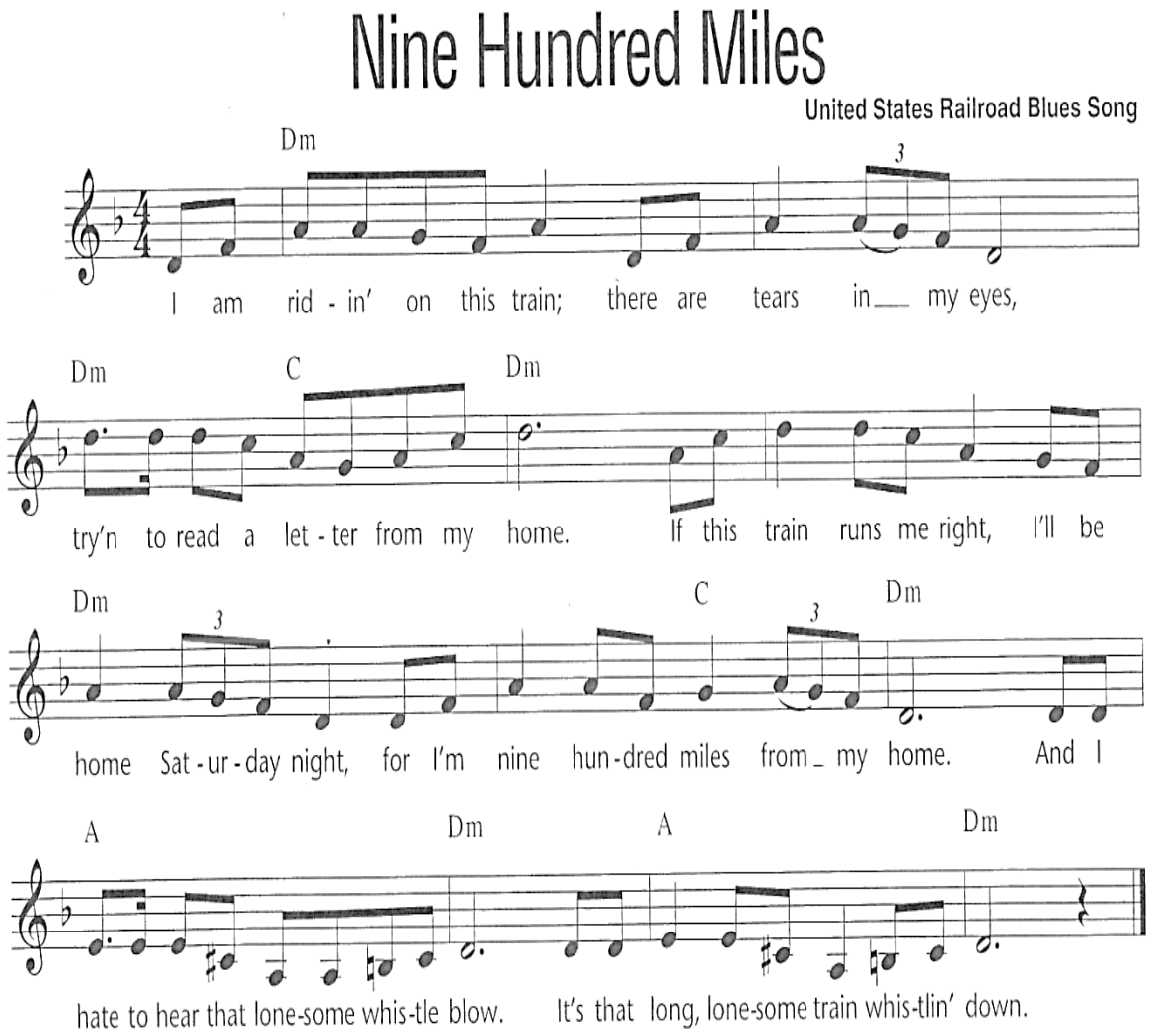
APPENDIX B (CONTINUED)



APPENDIX B (CONTINUED)

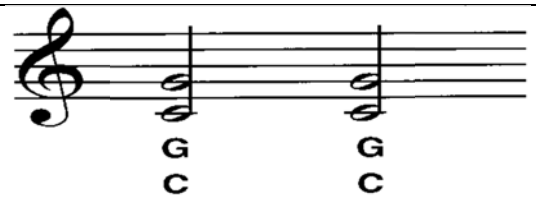
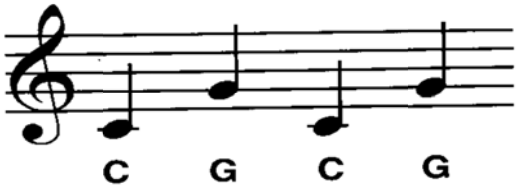
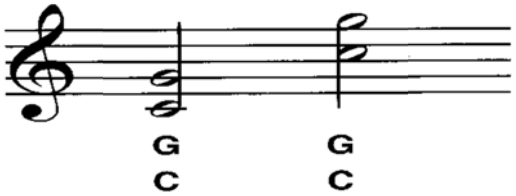
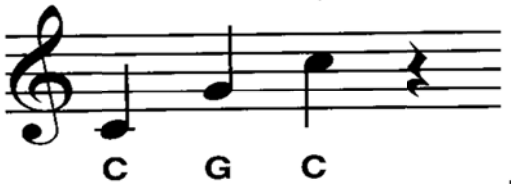
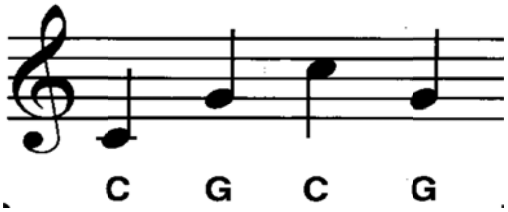


APPENDIX B (CONTINUED)



APPENDIX C

Accompaniment Patterns

Grade	Pattern Type	Score
Kindergarten	Chord Bourdon	
1st grade	Broken 5ths Bourdon	
2nd grade	Levels Bourdon	
3rd grade	Crossover Bourdon	
4th and 5th grade	Challenge crossover Bourdon	

APPENDIX D

Sample Rubrics

Rubrics for singing test (Grades 3-5)

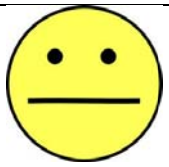
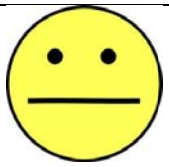
	CRITERIA	Excellent (4) 0 mistake	Very Good (3) 1-3 mistakes	Good (2) 4-6 mistakes	Developing (1) 6+ mistakes	LEVEL ATTAINED
1	**Melody Accuracy**	Sings all pitches in tune.	Sings most pitches in tune.	Sings some pitches in a singing voice and others with a speaking voice.	Is beginning to sing pitches correctly.	
2	**Rhythmic Accuracy**	Sings all rhythms correctly with a steady beat.	Sings most rhythms correctly with a steady beat.	Sings some rhythms correctly and changes speed of the beat when difficult.	Is beginning to sing rhythm correctly.	
3	**Text**	Sings with all words correctly.	Sings most words correctly.	Sing some words correctly.	Is beginning to sing with words.	
4	**Tone and Breath Support**	Sings with all clear tone and enough breath support.	Sings mostly with clear tone and enough breath support.	Sing some with clear tone and have less breath support.	Is beginning to sing with clear tone and not enough air support.	
5	**Posture**	Maintains a proper posture all throughout the song.	Mostly proper posture.	Posture acceptable through most of the song.	Posture is not consistent.	

Total Grade _____/20

Comments:

APPENDIX D (CONTINUED)

Rubric for singing for lower grade (Grade K-2)

	CRITERIA	Excellent (4) 0 mistake	Very Good (3) 1-3 mistakes	Good (2) 4-6 mistakes	Developing (1) 6+ mistakes	LEVEL ATTAINED
1	**Melody Accuracy**					
2	**Rhythmic Accuracy**					
3	**Text**					
4	**Tone and Breath Support**					
5	**Posture**					

Total Grade _____/20

Comments:

APPENDIX D (CONTINUED)

Rubrics for Playing Orff Instruments for upper grade (3-5)

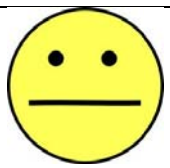
	CRITERIA	Excellent (4) 0 mistake	Very Good (3) 1-3 mistakes	Good (2) 4-6 mistakes	Developing (1) 6+ mistakes	LEVEL ATTAINED
1	**Melodic Accuracy**	Play all pitches accurately	Plays most pitches accurately	Plays some pitches accurately	In beginning to play pitches correctly	
2	**Rhythmic Accuracy**	Play all rhythms with precision and fluency in a proper steady tempo	Plays most rhythms accurately	Plays some rhythms accurately	In beginning to play rhythm correctly	
3	**Posture**	Maintains proper posture throughout the entire performance	Maintains proper posture through most of the performance	Maintains acceptable posture through most of the performance	Is beginning to demonstrate brief intervals of good posture	
4	**Ensemble Precision**	Play together with a group all the time with precision and fluency in a proper steady tempo	Play together with a group most of time	Play together with a group some of the time	In beginning to blend with others	

Total Grade _____/16

Comments:

APPENDIX D (CONTINUED)

Rubrics for Playing Orff Instruments for lower grade (K-2)

	CRITERIA	Excellent (4) 0 mistake	Very Good (3) 1-3 mistakes	Good (2) 4-6 mistakes	Developing (1) 6+ mistakes	LEVEL ATTAINED
1	**Melodic Accuracy**					
2	**Rhythmic Accuracy**					
3	**Posture**					
4	**Ensemble Precision**					

Total Grade _____/16

Comments:

APPENDIX E

Survey Questions

Survey questions on Student Perceptions of Rubric Effectiveness in Elementary General Music Class

About Rubric	Strongly Agree	Agree	Disagree	Strongly Disagree
1. Does rubrics help me better understand the standards and expectations of lessons?	☐	☐	☐	☐
2. Did you have fewer questions about the assignment?	☐	☐	☐	☐
3. Do you feel that you were more organized?	☐	☐	☐	☐
4. Did you work harder than usual to meet the highest grade?	☐	☐	☐	☐
5. Do you think that using rubrics helped you earn a better grade?	☐	☐	☐	☐
6. Did you feel that the teacher feedback was clear and consistent?	☐	☐	☐	☐
7. Overall, did you like to use a rubric?	☐	☐	☐	☐

APPENDIX F

Parental Permission Letter

Dear Parent or Guardian,

We are doing a research study to find out how children like you learn how to sing and play in a music class. We are asking you to be in the study because you are in a class that is learning a general elementary music. If you agree to be in the study, you don't have to do anything extra. We will make a copy of your grades and tests and use it for our study. We hope to learn something about singing and playing instrument that will help other children in the future.

You do not have to say "yes" if you don't want to. No one, including your parents, will be mad at you if you say "no" now or if you change your mind later. We have also asked your parent's permission to do this. Even if your parent says "yes," you can still say "no." Remember, you can ask us to stop at any time. Your grades in school will not be affected whether you say "yes" or "no." We will not use your name on any papers that we write about this project. We will only use a number so other people cannot tell who you are.

You can ask any questions that you have about this study. If you have a question later that you didn't think of now, you can [insun.ko@xxxxxxxx] contact me any time.

Name of Child: _____Parental Permission on File: □ Yes □ No

(For Written Assent) Signing here means that you have read this paper or had it read to you and that you are willing to be in this study. If you don't want to be in the study, don't sign.

Signature of Child: _____Date: _____

(For Verbal Assent) Indicate Child's Voluntary Response to Participation: □ Yes □ No

Signature of Researcher: _____Date: _____