THE USE OF NOTATED AND AURAL EXERCISES AS PEDAGOGICAL PROCEDURES INTENDED TO DEVELOP HARMONIC ACCURACY AMONG BEGINNING JAZZ IMPROVISERS

James Edwin Laughlin, B.A., B.M.E., M.M.E.

Dissertation Prepared for the Degree of

DOCTOR OF PHILOSOPHY

UNIVERSITY OF NORTH TEXAS

August 2001

APPROVED:

Darhyl Ramsey, Major Professor
Hildegard Froehlich, Minor Professor
Roger Warner, Committee Member
Michael Steinel, Committee Member
Graham Phipps, Director of Graduate Studies in Music
Warren Henry, Chair of the Department of Music Education
Thomas Clark, Associate Dean for Academic Affairs of the College of Music
C. Neal Tate, Dean of the Robert B. Toulouse School of Graduate Studies

Laughlin, James Edwin, The use of notated and aural exercises as pedagogical procedures intended to develop harmonic accuracy among beginning jazz improvisers. Doctor of Philosophy (Music Education), August 2001, 107 pp., 11 tables, 3 illustrations, references, 27 titles.

This study compared the effects from the use of aural and notated exercises as pedagogical procedures for teaching harmonic accuracy to beginning jazz improvisation students. The methods of pedagogy were identified from published pedagogic and historical sources, the results of oral and written traditions of jazz pedagogy. The performance objective was produced from a review of the related literature as a recognized and measurable characteristic of jazz improvisation.

The purpose of this study was to compare measurements of harmonic accuracy, following the use of notated and aural exercises as experimental procedures of jazz improvisation pedagogy.

A lesson plan, materials, curriculum and outline were developed followed by student recruitment and participation. A total of 20 student volunteers participated in the methods of pedagogy (aural or notation). Data collection consisted of a musical background questionnaire and pre and posttest performance recordings. Student recordings were evaluated by six judges using the "experimental performance evaluation measure." Statistical analyses were conducted, including comparisons of pre to posttest effects between, and among the methods of pedagogy.

Although all student participants performed mostly from notated music prior to this study, students who received the aural method of pedagogy produced greater improvement for all measurement items. While the aural method produced no differences between grade level, the notation method produced significantly lower scores for 9th graders compared to 12th graders; no other significant grade level differences were noted. Conclusions were that although many sources of pedagogy do not include aural exercises as the predominate activity, beginning improvisers who have more experience reading music than playing by ear, learn better from aural, than notated exercises.

ACKNOWLEDGMENTS

The conclusion of this project is a testament to a loving wife, family and community of friends who supported this endeavor. My foremost gratitude is to Jesus Christ, who revealed to me the truth about Love and Grace, so that I might fulfill His desires. To my wife Tammy, and children, Alex and Nick, without your love and understanding this task was not possible.

I would also like to thank my major professor, Dr. Darhyl Ramsey, for always encouraging me to complete the project. To the committee members – in particular Mike Steinel, I am appreciative of your wisdom and assistance; and to the volunteer evaluators, thank you for donating your time and talents to this project. I owe special recognition to David Fike, my bass player / statistician; your many hours of assistance were invaluable.

Thanks to my church and pastors, Dr. Kurt Oheim and Rob Leathers, who helped me realize a deeper faith in Christ. The many deeds from my friends at Pinnacle Community Church, who never failed to offer encouragement, prayers and friendship, will be remembered. Finally, I thank my parents, Kathy and Linard Laughlin for their support and love through 21 years of college.

TABLE OF CONTENTS

Page

ACKNOWLEDGMENTS iii

LIST OF TABLES v

LIST OF ILLUSTRATIONS vi

Chapter

1. INTRODUCTION..... 1

   Introduction of the Problem
   Background
   Rationale
   Statement of the Purpose
   Research Questions
   Delimitations

2. REVIEW OF RELATED LITERATURE 11
   Jazz Improvisation Research Studies
   Selected Pedagogical Publications
   Historical Documentation

3. METHODOLOGY..... 29

   Development of Experimental Procedures
   Pilot Testing Procedures
   Analysis
   Main Study

4. RESULTS 39

5. DISCUSSION AND RECOMMENDATIONS..... 56

APPENDICES 63

REFERENCE LIST..... 105

LIST OF TABLES

Table Page

1. Inter-rater reliability for combined pretest and posttest scores 43
2. Inter-rater reliability comparisons, Kruskal-Wallis ANOVA 45
3. Pretest equivalence, Mann-Whitney Comparisons 46
4. Comparisons of pretest and posttest scores by method 48
5. Comparisons of posttest scores by method 48
6. Posttest minus pretest scores between methods 49
7. ANCOVA comparisons of posttest scores between methods 50
8. ANCOVA comparisons of posttest scores between methods for Q2 52
9. Comparisons of absolute and adjusted posttest scores by method 53
10. 2 way ANOVA comparisons of student grade level by treatment methods 54
11. Mann-Whitney comparisons of posttest notation scores by 9^{th} and 12^{th} graders 55

LIST OF ILLUSTRATIONS

Figure Page

1. Interactions within the research model 41

2. Judge numeric value assignments 42

3. Comparison of absolute score values and posttest minus pretest values……45

CHAPTER 1

Introduction, Rationale and Purpose

Introduction

Aural and notated exercises are commonly used pedagogical procedures for jazz improvisation. Historical accounts portray the learning style of many early jazz musicians as an aural apprenticeship, through imitation of live and recorded performances. However, recently published materials, according to Brown (1990), present notated musical exercises rather than aural development as the predominate skill. While authors, such as Coker (1989) and Baker (1989), recommend that comprehensive jazz improvisation curricula should include the use of both aural and notated methods of pedagogy, there has been no investigation of their comparative effectiveness.

Notated and aural exercises are used to develop performance skills, although the method of pedagogy is not apparent to the listener. For example, a student who acquires harmonic concepts through the use of notated theory exercises may then demonstrate these concepts through an improvised solo. In comparison, harmonic skills might be developed directly on one's instrument "by ear," without the use of music notation reading skills. To the listener it becomes difficult to determine the method of pedagogy, and consequently, its effectiveness. The comparative outcomes from the use of aural or notated pedagogy have not been the focus of research investigation. Therefore, comparisons of aural and notated methods of pedagogy are the subject of this study.

Rationale

Historical documentation provides evidence that jazz performers learned to improvise primarily through aural apprenticeship. Recently published jazz teaching materials, on the other hand, emphasize the use of notated exercises rather than aural methods. An examination of historic documentation will describe and illustrate the use of aural-imitation while published pedagogical sources will demonstrate the use of notated musical exercises as a recognized tool of jazz pedagogy.

Historical Documentation

Historical accounts attribute the origination of jazz to African tribal music. Because this lineage was passed from one generation to the next by oral tradition, African slaves entering the United States carried with them an oral musical tradition. Southern (1983) describes the development of jazz through "aural traditions." She states:

> Jazz is primarily an aural kind of music, its written score represents but a skeleton of what actually takes place during a performance. Performances of the same work differ from player to player, for each recreates the music in his own individual way. Jazz is learned through oral tradition, as is folksong, and those who would learn to play it do so by listening to others playing jazz. (p. 363)

Jazz began as a type of folk music, performed and passed along predominately by ear (Southern, 1983). As jazz became more stylized, a written tradition emerged that involved the use of arrangements and composition. Charters & Kunstadt (1962) describe the use of notation in early jazz as a means to create stylized ensemble performances. Although the history of jazz includes a lineage of oral and written traditions, recognized soloists emerged from the oral tradition.

Authors have documented accounts of historically significant jazz performers who learned through oral tradition. Woideck (1999) provides a biographical account of Charlie Parker that includes information about Parker's musical training. He began by participating in his public school band and continued with self-guided efforts. Woideck states:

> Other than his school band experience, Parker had no formal music instruction per se, but in the tried-and-true "oral tradition," the many more-experienced older musicians he encountered were potential informal teachers. Parker took what he gleaned from them, plus ideas he heard in live and recorded music, and embarked on a period of self-study in the "woodshed." (p. 5)

Woideck continues by describing Parker's apprenticeship and the significance of hearing not only live, but recorded solos. He states that "being able to play repeatedly and to study methodically [Lester] Young's improvised solos on record made an important difference for Parker. This 1937 breakthrough marked the beginning of Parker's period of most rapid musical growth" (p. 12).

Clark Terry's account, documented by Dance (1970), provides further testimony from a recognized jazz performer. Terry described an apprenticeship that involved recordings of Louis Armstrong and Roy Eldridge, but he also stated the importance of hearing live trumpet players locally. Terry expressed influential moments with professional jazz trumpet players from his youth, portraying a system of pedagogy based on aural-imitation.

According to Southern (1983) early jazz musicians, particularly African Americans, learned their craft through aural-imitation by using recorded and live jazz

performances as their primary resources. Unlike live performances, however, recordings allow students the opportunity to replay a selected example until they can imitate various aspects of the “model” performer in their own playing.

The implications of sound recording to the jazz tradition have been described as having a profound and formulating influence on developing performers. Charters and Kunstadt (1962) state, “jazz has always been closely tied to the phonograph record” (p. 60). Lyons (1980) writes:

> Jazz has no better friend than records. As the music is characterized by spontaneous, emotional improvising, a jazz performance can never be recreated with exactly the same feeling, even by the same players. The music's foremost heroes have peak periods of creativity, freshness of imagination, and technical skill which they themselves cannot recapture once the delicate web of circumstances is altered. Nor can the best jazz, which always involves an element personal to the musicians involved, be written down in all its nuances for "duplication" by others. Thus sound recording is the only means of documenting jazz's development and its artistic triumphs. (p. 7)

Due to the nature of aural-learning sound recording, along with other types of recording media, has been crucial to jazz apprenticeship.

Selected Pedagogical Sources

Berg (1990), Coker (1989) and Lawn & Hellmer (1990) were reviewed with regards to their use of notated and aural exercises. Berg made use of a notated “practice-sheet method” in his pedagogical text. “The practice sheet is a notation of the chords and scales contained in the progression of a song” (p. 10). Notated exercises are presented in

all chapters of the book, while no chapter or major heading was devoted to the use of aural exercises.

Coker (1989) recommended a curriculum outline and how to teach from it. He made use of notated musical exercises within this curriculum, including written exercises contained within his jazz piano text. Although aural development is discussed as a necessary skill for jazz improvisation, no major heading within the recommended curriculum is devoted to this objective.

Lawn & Hellmer (1990) presented formal written musical theory knowledge, along with traditional ear training exercises, intended to be applied to a jazz solo. "This book provides all the necessary tools to understand not only the theories associated with jazz styles but also the relationship of information to arranging, composing, and improvising" (p. xiii). Ear training was described as important to any jazz theory text with exercises presented in seven of the fifteen chapters.

Issues pertinent to the scope of this study were: 1) identifying an appropriate sample group, and 2) identifying a significant and measurable performance variable. When comparing the use of aural and notated methods of pedagogy it was necessary to consider participants' musical background and experience level.

Sample Group

Observed changes from beginning level improvisers provided accountability for various types of research error. The pretest-posttest experimental design reduced regression type error, due to the ability to establish treatment group homogeneity prior to the experimental treatment sessions. Due to the difficulty to account for experienced

players musical development, beginning jazz improvisers provided for greater control of the observed effectiveness of the pedagogical methods.

Performance Variables

The development of a lesson plan intended for beginning level students began with the specification and definition of a learning objective. For the purposes of comparing the two pedagogic methods, a performance objective selected and defined from the research literature ensured the use of a recognized and measurable variable. Harmonic skills were described in many research studies as fundamental criteria of jazz improvisation performance. Although studies provided varying descriptions of this performance variable, measurement procedures consistently produced good reliability. The inclusion of this objective as a measurement variable in studies that involved beginning improvisers indicated the appropriateness of this variable as a learning objective for this level of student.

Studies including Briscuso (1972), Segress (1979), Zwick (1987), Tumlinson (1991), Pfenninger (1990), and Burnsed and Price (1984) included descriptions of harmonic accuracy as an observable and measurable performance skill of jazz improvisation. The recognition of harmonic aspects in early studies, Briscuso's (1972) and Segress' (1979) use of the term "harmonic awareness," led to terminology more specific to performance, Pfenninger's (1990) term "note accuracy" and Tumlinson's (1991) term, "harmonic appropriateness."

Briscuso (1972) compared jazz improvisation ability to scores obtained from the MAP (Gordon, 1965). His use of high school band students with no prior training or experience in jazz improvisation is comparable to the selection of student level from the

present study. An evaluation instrument was developed through the selection and use of five performance criteria (harmonic awareness, rhythmic development and interest, melodic expressiveness, ability to play with jazz style, individuality). No further defining information was provided for the performance variable, harmonic awareness. Reliability estimates for the overall measure were between r = .74 and r = .85.

Segress (1979) also used the description "harmonic awareness" among other evaluative variables, using a five-point Likert-scale rating system. He developed and tested a first semester college level jazz improvisation curriculum. Harmonic aspects, chord and scale exercises, were included as areas of instruction. The evaluation of these aspects is measured through the use of the term "harmonic awareness." Segress reported a reliability estimate of r = .99.

Pfenninger (1990) used the term "note accuracy" as a description of tonal evaluation. He investigated the measurement of jazz improvisation achievement by creating three experimental rating scales (tonal, rhythm, and expression). These rating scales were developed from a literature review and a university jazz educator questionnaire. Three jazz educators and two non-jazz music educators listened to and rated twenty taped performances using the three experimental rating scales. A higher correlation coefficient was reported between the tonal and rhythm scales than those involving the expression rating scales. Reliability estimates for the tonal rating scale were reported as r = .77.

Damron (1973) investigated self-instructional pedagogy for high school level beginning jazz improvisers. Although the method of pedagogy involved both written and aural exercises, his selection and use of dorian mode exercises provided an example of

appropriate material for this level of student. Student evaluation involved the statement "overall quality of the improvised solo," with no additional defining terminology.

Aitken (1975) investigated experimental pedagogy intended for high school trumpet players. His treatment method used audio-imitation and notated exercises as separate units of instruction. Aitken did not investigate comparisons from the use of aural and notated exercises; his inclusion of these exercises for high school level students, however, provided guidance to this study.

Burnsed and Price (1984) developed an adjudication form based on 33 construct areas identified from the literature, and three selected university jazz studies programs. An adjudication form was produced from categorical groupings of these areas along with one additional area provided by the authors, called "overall effect." "Tonal materials" was used as one of six construct areas to be measured, using the term "note choice" as the defining phrase. Interjudge reliability was reported as r = .735.

Tumlinson (1991) produced seven hypothesized constructs of jazz improvisation performance, including the term "harmonic appropriateness." His use of 33 descriptive items as the "descriptive improvisation measurement instrument" (DIMI) involved 4 individual items related to the measurement of harmonic accuracy. Terms within these descriptions included "demonstrates harmonic appropriate performance skills," "tone selection," "ability to solve harmonic problems," and "exhibits harmonic inventiveness." A Cronbach's Alpha coefficient was used to determine reliability estimates for each evaluation item, with a reported Alpha = .86 for item number 1, tonal fluency.

Harmonic skills have been identified in previous research studies as important to jazz improvisation pedagogy and evaluation. A review of the literature revealed that the

evaluation of how accurately harmonic aspects were performed was achieved through the use of descriptive terminology, usually rated by evaluators on a Likert type measurement scale. Therefore, the use of harmonic accuracy as the performance objective for this study provided a recognized and measurable criterion. The selection of high school beginning level jazz students allowed for greater accountability of research variables, including controlled measurement before and after experimental pedagogy. Because no study had investigated the outcome of aural and notated methods of jazz improvisation pedagogy in relation to harmonic accuracy, the purpose of this study was as follows.

Purpose of Study

The purpose of this study was to compare measurements of harmonic accuracy, following the use of notated and aural exercises as experimental procedures of jazz improvisation pedagogy.

Specific research questions addressed by this study were:

1. What were the effects on high school band members who had not participated in a jazz band or experienced performing a jazz solo, from the use of notated and aural exercises, involving the development of harmonic accuracy?

2. What were the comparative outcomes from the use of aural and notated methods of pedagogy by student grade level?

Delimitations of Study

This study was limited to high school band students who had not participated in a jazz band or performed a jazz improvised solo. A student questionnaire verified the selection of the sampled population, along with pretest observations. In order to compare similar performance variables from all participants, polyphonic instruments (such as

guitar or xylophone) were instructed to perform single line solo's. This allowed the evaluators to judge student performances with greater consistency. The performance variable, harmonic accuracy, was selected in order to define a lesson objective for beginning level students.

CHAPTER 2

Review Of Related Literature

The related literature is presented in the following three sections: jazz improvisation research studies, selected pedagogical sources, and historical documentation. The first section presents a review of jazz improvisation studies which identified performance characteristics and developed measurement devices while documenting validation procedures. The second section details three pedagogical sources which proposed the development of harmonic performance characteristics using notated musical exercises. The third section provides accounts from historical documentation that describes predominately aural apprenticeship among innovative jazz performers.

Jazz Improvisation Research Studies

Stating that "there have been no experimental research studies related to the teaching of jazz improvisation" (p. 5) prior to his study, Briscuso (1972) was the first to utilize empirical methodology in jazz improvisation research. He compared an ability measure of spontaneous and prepared jazz improvisation solos to scores obtained from the MAP (Gordon, 1965). His stated purpose "was to investigate the extent to which students who possess different levels of musical aptitude possibly differ in their ability in jazz improvisation" (p. 3). Jazz improvisation scores of high, average, and low were correlated with scores from the battery of tests which determine the MAP, and the degree of predictiveness from the MAP scores to spontaneous and prepared jazz improvisation was investigated.

An experimental research design involving 48 junior and senior high school students who had a minimum of three years of band but no training in jazz improvisation was used. All students were given the MAP as a pretest, followed by the experimental treatment sessions that involved thirty weeks of a non-required jazz improvisation course, which met as a two-hour session once per week. The author described the sessions in terms of students learning the jazz style and developing improvisation skills.

Evaluation of jazz improvisation abilities was measured through the selection of five performance criteria (harmonic awareness, rhythmic development and interest, melodic expressiveness, ability to play with jazz style, individuality). Each student performed a spontaneous and prepared solo over a thirty-two-bar, AABA progression, and a twelve-bar blues progression. Three judges rated each performance criterion using a five-point Likert scale; each judges' scores were summed for each performance and all judges scores were totaled for each. Inter-judge reliability coefficients were estimated between $r = .74$ and $r = .85$.

Although no correlational significance was reported between the jazz improvisation ability score and the MAP scores (with the exception of the "musical sensitivity" portion of the MAP), the results of this study concluded that students who score above the median percent on the MAP can benefit from jazz improvisation instruction.

Briscuso's contribution to this study includes the use of research procedures and the selection of harmonic awareness as a significant performance criterion of an improvised jazz solo.

Damron (1973) studied a self-instructional sequence in jazz improvisation for secondary instrumentalists. The dorian mode was chosen as the exclusive use of theory knowledge. Ear training was cited as an important element with a call and response method described. A review of existing jazz pedagogy materials revealed five types: (1) texts, (2) exercise books, (3) instruction book-record packets, (4) keyboard oriented books, and (5) transcription of recorded solos. The experimental design was described as "a combination graphic-aural method," using cassette tape recordings and an instrumental performance book (individual books transposed for various instruments). The format was divided between performance and theory units.

A posttest-only-control group design was chosen with a total of forty students, randomly selected from five secondary schools. Of the forty students, twenty were members of their school stage band and concert band while the other twenty were members of their concert band only. Two groups were randomly selected from both sets of students. Administration of the experimental pedagogy occurred at students homes where they were asked to keep a log of how much time they spent with the packet (parents were involved). Five weeks were chosen as the time allotment, concluding with a posttest recording from both groups. Three judges "who were music educators with experience in jazz education" (p. 28) scored each recording with the statement "overall quality of the improvised jazz solo" as the only criterion.

Statistical analyses were performed revealing significant inter-judge reliability and higher experimental group scores. A Kendall's coefficient of concordance between judges' scores revealed a significant correlation of $w = .70$. A two-way analysis of

variance was performed that revealed a significant difference between the experimental group and the control group at the .05 level.

Damron's research design provides insight about the investigation of beginning jazz improvisers. The use of the dorian mode for beginning level students, and the use of a call and response technique were appropriate to the development of a lesson plan for this study.

Aitken (1975) developed a self-instructional method for high school trumpet players. He revealed four elements of improvisation (scales, patterns, cliches, and nuances) as the basis for his experimental method, which "is substantiated by numerous professional jazz educators and musicians" (p. 17). A search of available methods for these four elements provided the review for the development of his experimental treatment procedures. Content validity was reported through the inclusion of three letters from jazz educators, who stated their approval of the experimental method.

Aitken's use of audio-imitation and the design of his self-instructional method provided assistance in the development of procedures for the present study. The method consisted of a text with written instructions and musical exercises, an introduction to the text, and an audio cassette of pre-recorded examples for aural-imitation exercises. The introduction was intended to explain to the student how to use the method as well as explain necessary procedures and terminology. Included were the following topics: (a) Responsibility of the student; (b) listening and imitating; (c) how to practice; (d) developing a schedule and systematic routine; (e) aural discrimination; (f) making use of the piano and voice in addition to the trumpet "to develop and improve hearing" (p. 29); (g) tone quality, vibrato, tonguing, time; and (f) the use of text.

Each of Aitken's one hundred twelve units included three sections: explanation, exercise, and imitation. "All lessons are brief so the student can master a certain scale, pattern, cliche or nuance in a short period of time" (p. 30). Audio-imitation consisted of recorded examples of student exercises and a call and response approach. A flow chart describing the procedure for each lesson was provided to the student in the introduction to the text. The procedures were direct quotations while the chart was presented in numerical sequence rather than as a flow chart. Following is the lesson sequence:

> (1) Student studies music visually and aurally; (2) student listens to tape while following his music; (3) student practices exercises slowly and evenly, until he plays notes and interpretation correctly; (4) student plays, imitating the tape, plays notes and interpretation correctly; (5) repeats step 3 until achieves step 4; (6) student memorizes scale, pattern, cliche or nuance; (7) with text closed, student plays the exercise from memory in all keys with current notes and nuances; (8) call and response; and (9) next lesson. (p.33)

Aitken's use of an audio imitation method is closely linked to the traditional approach used to learn jazz improvisation. While formal knowledge might provide better understanding and could reduce the amount of time to achieve more advanced levels, the aural aspect remains a fundamental skill. A review of Aitken's procedures assisted in the development of aural-imitation pedagogical procedures.

Segress (1979) proposed and tested a jazz improvisation curriculum that utilized an empirical research design. He stated that while a number of prior studies which developed jazz curricula existed, these "writers failed to evaluate their work or provide any empirical evidence of success or failure" (p. 30). The research purpose, "to develop

and evaluate a comprehensive first semester college jazz improvisation curriculum" (p. 13) was clearly stated and adhered to throughout the study. The systems approach model used by Segress has been widely implemented in curriculum studies before and after his study. Segress (1979) implemented the following steps:

> (1) Identification of the problems from documented needs, (2) identification and writing of the instructional objectives and performance standards, (3) identification of the entering competencies and construction of the entry test, (4) preparation of the evaluation instruments from the instructional objectives to be used as pretest and posttest, (5) identification of the possible curriculum strategies from alternatives, (6) selection and development of the curriculum, (7) implementation of the curriculum in small group field tests and pilot program, (8) revision of the curriculum as required, (9) implementation of the curriculum in a classroom situation, and (10) evaluation to determine the effectiveness of the curriculum. (p. 16)

The methodology included the development, implementation, and evaluation of the designed curriculum. Problems and solutions concerning evaluation were discussed. Pre-treatment evaluation consisted of a jazz theory, listening and improvisation performance test, and student questionnaire. Students who tested and participated in the study were grouped into two categories: "those enrolled in the semester of the designed curriculum" and "those participating in a jazz ensemble but not in the improvisation class" (p. 39). Three visiting evaluators graded the performance tests (both pretest and posttest).

An analysis of covariance comparing the pretest results to the posttest results revealed a significant change. The estimated reliability coefficients, based on a Kuder-Richardson formula, were r = .98 on the pretest, r = .99 on the posttest theory / listening portion, and r = .92 on the performance evaluation. Conclusions stated that the students exposed to this curriculum had benefited in the following specific areas: 1) improved music theory fundamentals; 2) the identification of jazz tunes, composers, musical forms and prominent jazz performers; 3) improvisation performance in a jazz style; and 4) the development of a positive attitude toward improvement.

Zwick (1987) developed an experimental sequential format for jazz improvisation instruction through a content analysis of selected jazz materials. Prior to the analysis, he surveyed nine jazz educators through a questionnaire on a set of criteria for the inclusion of text materials. Included in these criteria were the following: the recommended text must be published and available, levels from beginner to advanced must be included, the author must be a jazz educator, and the text must be for general instruction (not specific to any instrument or element). Any recordings used with the materials would be considered beyond the recommended sequential format.

A list of seventeen instructional areas, derived from the selected materials, were used to analyze the data. The following instructional areas were: (01) History of improvisation, (02) prerequisites for study of jazz improvisation, (03) jazz improvisation fundamentals, (04) ear training; (05) jazz style, (06) analysis, (07) form and structure of jazz music, (08) melodic improvisation, (09) patterns for improvisation, (10) chord progressions, (11) rhythm section, (12) substitutions, (13) transcription of jazz solos, (14)

improvising on jazz music, (15) scales for improvisation, (16) non-harmonic tones; and (17) the blues (p. 76).

Zwick constructed a coding system to weigh the emphasis of each instructional area according to chapter heading, significant part of a chapter, and listings in several chapters. A second coding table was then constructed to record the introduction and sequence of each instructional area. Through content analysis each instructional area was recorded above the median percent of emphasis, except for (02) Prerequisites for Study of Jazz Improvisation and (16) Non-Harmonic Tones. Based on his analysis, Zwick proposed a sequence in which each area above the median percent should be introduced.

Zwick's identification of seventeen instructional areas gathered from the jazz pedagogy literature provided a summarization of this literature. Because many of these instructional areas involve the development of harmonic performance skills, this summarization further affirms that this area is prominent in both the pedagogy and research literature.

Pfenninger (1990) studied the problem of objectively measuring jazz improvisation achievement. Through a review of the literature and a questionnaire administered to university jazz educators, instruments were developed to rate three performance areas, including the tonal rating scale, the rhythm rating scale, and the expression rating scale. A pilot study was conducted with the researcher and one other judge acting as evaluators who rated four performances. The actual study employed six judges, three jazz educators, two non-jazz music educators and the researcher. Using the three rating scales, each judge independently rated twenty taped performances, rating each three separate times over a three-week period. Each week the judges received a new

tape of the same performances in a new random order. Upon completion of this process, one judge repeated the entire process one week later in order to measure test-retest reliability. In addition to this measure, inter-judge reliability coefficients were analyzed along with intercorrelations among the rating scales.

Pfenninger found that the intercorrelation between the expression and rhythm dimensions was high (r = .71), moderately low between the tonal and rhythm dimensions (r = .40), and low between the tonal and expression dimensions (r = .18). Interjudge reliability coefficients yielded a high overall composite for the tonal rating scale (0.77), and rhythm rating scale (r = .78), but lower for the expression rating scale (r = .67). The composite of all three rating scales revealed a high overall reliability (r = .79). The researcher interpreted the lower correlation coefficient of the expression rating scale to the subjective nature of this area of jazz improvisation achievement. The low correlation between the tonal dimension and both the rhythm and expression dimension "indicate that those dimensions have relatively little in common and that the scales were being used independently of each other" (p. 41). Pfenninger's identification and use of a tonal rating scale among the three scales he developed and tested reveals harmonic performance as a significant aspect of jazz improvisation.

Burnsed and Price (1984) developed and tested an adjudication form for the evaluation of jazz improvisation. Jazz evaluation criteria were collected from three university jazz studies programs and a review of the jazz improvisation literature, producing 33 terms denoted as the constructs of jazz improvisation. These terms were grouped into five category headings: 1) Technical facility, 2) melodic and rhythmic

development, 3) style, 4) tonal materials, and 5) emotional effect. These authors produced a sixth category, 6) overall effect, to complete the adjudication form.

A panel of eight judges rated each area on a Likert type scale from 1 (low) to 5 (high); a total of eleven tape-recorded solos was rated by each judge. Four of the judges were described as having "extensive jazz performance and knowledge," while four were described as having "extensive experience in the non-jazz idiom" (p. 36). The taped solos were randomly ordered from "professional and student jazz musicians of varying degrees of skill and style" (p. 36).

The results of this study were obtained from total scores from the overall measure in addition to item analysis. Mean scores and rank order from the data were revealed for all judges and by judge classification (jazz versus non-jazz). Although disagreement between judge classification was determined, a Kendall's coefficient of concordance analysis yielded a significant correlation ($w = .73502$, $x2 = 58.8016$, $df = 10$, $p = .001$) indicating "very good" interjudge reliability. Individual rank orders by judge classification also produced significant interjudge reliability according to these authors. Finally a correlation matrix for category ratings revealed that each category heading is significantly related to the others. The category "emotional effect" produced a lower correlation, indicating that, "it may represent a distinct and less clearly defined construct of jazz improvisation" (p. 38). It was discussed that most of the judges "commented that the accompanying group for each solo had an effect on their evaluations. Future research should try to control this factor" (p. 39).

Through the development of their rating scale, Burnsed and Price (1984) categorized tonal materials as one of five significant areas. Among the defining

terminology within the tonal scale, “note choice” provided the clearest defining characteristic.

Tumlinson (1991) identified and grouped variables into seven hypothesized constructs of jazz improvisation performance. 266 variables identified from the literature were reduced to 33 after pilot testing and eliminating redundancies. Recorded excerpts of n = 60 student and n = 60 professional performers were rated by two jazz educators with the use of an experimental measurement instrument. The “descriptive improvisation measurement instrument” (DIMI) was created from the identified variables and used by the selected judges. Through factor analysis Tumlinson discovered that unique variables were identified between combined, student and professional performance samples, finding only two identified constructs consistent across samples. “This indicated that the populations of student and professional performers are so unlike that combining them appeared to confound the results” (p. 125). The identified constructs for each sampling group were discussed and compared.

Although similar, the differences in identified constructs between student and professional performers suggested variation in a listener’s attention. Tumlinson suggested that because the professional brings the listener to a relaxed state regarding such elements as style and harmonic / melodic congruity, attention may turn toward other areas of the performance. Student performances, on the other hand, may create a less comfortable state to the listener, who then concentrates on evaluation. He suggests that criteria could be developed for each level of performance ability, thus clarifying pertinent variables and creating more reliable and valid measurement. “It is possible that evaluative variables would be more effective in distinguishing constructs for the student level, while detailed,

descriptive variables would be more effective in distinguishing constructs for professional performers" (p. 134).

In his recommendations, Tumlinson states:

> Rating scales based on the constructs of jazz improvisation should be developed. The variables as developed by the present study could serve as material upon which to develop the items for such a rating scale. The body of literature surveyed could also serve as supplementary material upon which to build these rating scales. These scales should focus on the specific playing level to be rated, student or professional. The variables that were collected could also be utilized to offer diagnostic information to the student and teacher. Pedagogy and curriculum development could also utilize these variables as a check (list to insure that all important jazz improvisation concepts have been addressed). (p. 135)

The identified variables for student performers were: (1) Jazz style/time feel, (2) harmonic and melodic congruity, (3) melodic breadth, (4) rhythmic and melodic variety, and (5) fluency. Hypothesized constructs which did not emerge from factor analysis were "individuality" and "form."

Of the 266 variables Tumlinson (1991) collected from the literature, 33 were identified and grouped into seven hypothesized constructs: 1) harmonic appropriateness; 2) rhythmic usage; 3) melodic usage; 4) jazz style; 5) individuality; 6) expressiveness; and 7) form. This study produced and codified data regarding the identification of observable performance characteristics of jazz improvisation. It represents to some degree a culmination of the research literature involved in the identification of jazz performance variables. The identification and inclusion of harmonic appropriateness as

one of the primary constructs of jazz improvisation performance provides further evidence for the selection and use of harmonic accuracy as the dependent variable for the present study.

Selected Pedagogical Publications

Many published pedagogical sources for jazz improvisation utilize notated examples and exercises. As examples of this, three sources that provided a unique strategy for the development of harmonic performance skills were reviewed. The use of notation was then compared to documented historical accounts involving primarily aural development.

Berg's (1990) publication was based on the premise that successful improvisation results when patterns appropriate to the musical style surround consonant tones. This method contains information regarding stylistic characteristics and how to achieve them. Notated examples were used throughout the book, along with a CD recording of examples and accompaniment. Each chapter dealt with a basic concept, including recommended exercises. Although many of these exercises involved notation, Berg advised in the preface that students should perform each concept on their instrument.

Lawn and Hellmer (1990) provided a jazz theory text intended for those with a background in traditional music theory knowledge. They describe the purpose of this text as the application of formal theory knowledge to the jazz style.

> While a knowledge of traditional music theory principles certainly is helpful to the understanding of jazz, such courses and books rarely address the specific characteristics unique and indigenous to the jazz language. As a result, students of

> music are frequently left with no understanding of relevant materials necessary for the study of jazz improvisation and composition or arranging. (p. xiii)

Although the text consisted predominately of notated examples and exercises, ear training exercises were also included in many chapters with an accompanying audio recording. The authors stated that "because jazz has developed over the past nine decades as a result largely of the oral and aural traditions, it seems that ear training must be a central component of any jazz theory text designed to provide not only theoretical comprehension but also an understanding of common practice application" (p. xiv). These authors recognize the distinction between theoretical comprehension and common practice application. Exercises throughout this text direct students to play the proposed concept on their instrument and in a jazz improvised solo when applicable. The use of music notation and other pedagogical tools were recognized by these authors as secondary to the final activity of the improvised jazz solo.

Coker (1989) proposed a jazz pedagogy curriculum and pedagogic tips intended as a guide for developing a university jazz studies program. In his rationale he listed descriptive elements of jazz as an alternative to providing a definition. Following are Coker's descriptive elements of jazz:

> 1) jazz is a musical art form, recognized as such around the world; 2) the key element of jazz is the craft of improvisation; 3) improvisation is a musical skill, requiring considerable time to develop. The theory and musical tools of improvisation may be mentally understood in a relatively short space of time, but the honing of the performance skills requires years...; 4) jazz is a very creative, personal sort of music...; 5) jazz is highly communicative and extremely

> spontaneous in that communication...; 6) as a musical style, jazz is very elusive, often confounding the critics, entrepreneurs, and audiences when they try to label and categorize the music as jazz or non-jazz...; and 7) elements of the jazz style have permeated nearly all known musical styles and musical functions... (p.14-15)

In his list of proposed curriculum course offerings, jazz theory, analysis of jazz styles, jazz piano, jazz composition and jazz arranging each made use of music notation skills. This curriculum, similar to many university jazz studies curriculums, included the development of pedagogical skills (such as notation, jazz piano, theory and historical knowledge) intended to be applied to the improvised jazz solo. In his argument for the inclusion of jazz programs in the university setting Coker stated:

> But it would be unfair to ignore the alternatives altogether. One of those alternatives, which could be described as 'the school of hard knocks' or 'street learning,' suggests that the best way to learn to play jazz is by listening, observation, experience, trial and error, imitation, etc. These are all valid notions and should be incorporated into the jazz learning experience, certainly. After all, how did all the great players of pre-jazz education years learn their craft? Can we improve on that? (p. 16)

He further stated "that modern jazz education does not ignore the methods of ‘street learning’" (p. 16). This recognition by Coker of an "old school" system of learning in comparison to modern jazz education is one of the premises of this study.

Historical Documentation

Jazz historical documentation has described early jazz as a product of oral folk tradition, and as such carried similar learning practices and procedures. Nettl (1976)

writes: "jazz has its origins in folk music: clearly it is related to folklore..." (p. 102). Eileen Southern (1971) describes the musical traditions of Black Americans from West Africa (from which most of the slave trade originated) as well as the development of indigenous styles in American culture. She traces the lineage of this music as the product of oral tradition until the most recent traditions. The nature of jazz is a folk tradition, according to Southern, due to the improvised solo which is performed differently each time.

The written score, both in performance and during the learning process, becomes less significant because of the nature of this activity. Nettl (1976) describes the methods used to learn materials within a folk tradition.

> Another basis for defining folk music is the manner in which it is transmitted. People learn some things through reading and other things by being told or shown. News read in a newspaper and a skill learned from a textbook are elements of culture transmitted in written tradition. Information passed from one person to another through speech is transmitted in oral tradition... (p. 22)

Cultures of varying complexities, according to Nettl, may involve oral traditions. She writes that "members of urban cultures, living in the centers of literate societies, also learn much by oral tradition, directly from other people" (p. 22). She further states that "music transmitted through oral tradition has generally been accepted by scholars as folklore, and oral tradition is the most commonly accepted criterion of folk music today" (p. 22).

Jazz history can also be described in terms of a written tradition, involving complex arrangements and notated compositions. As improvised jazz became stylized,

certain characteristics were captured by arrangers and composers in notated scores. The written, or notated, aspects of jazz began in order to provide various ensemble participation as well as to produce solo sheet music. Recognized soloists of early jazz styles, such as Charlie Parker, were produced, however, from a predominately aural tradition of pedagogy.

As a folk tradition, early jazz styles were learned predominately through listening and imitating other jazz musicians. Gunther Schuller (1968) describes the inadequacies of verbal explication and notated musical examples to describe this music.

> ...jazz, a basically improvised music defying notation and in which recourse to the written score is both impossible and - if scores existed - irrelevant. Despite the limitations of musical notation, a score by Beethoven or Schoenberg is a definitive document, a blueprint from which various slightly differing interpretations can be derived. A jazz recording of an improvised performance on the other hand is a one-time thing, in many instances the only available and therefore "definitive" version of something that was never meant to be definitive...The jazz historian is forced to evaluate the only thing that is available to him: the recording. (p. x)

The significance of sound recording to the tradition of jazz soloists has had a formulating influence. Schuller states that "in an improvised art the recording is all we can go by" (p. x). He describes the significance of recorded jazz performances due to the importance placed on the improvised jazz solo.

> Whereas we are interested primarily in the Eroica and only secondarily in someone's performance of it, in jazz the relationship is reversed. We are only

> minimally interested in West End Blues as a tune or a composition, but primarily interested in Armstrong's rendition of it. (p. x)

The influence of hearing a selected jazz performance repeatedly has further promoted the oral tradition of aural learning.

The use of recorded exercises as the method of aural-imitation pedagogy follows the tradition described from influential performers. The development of a single lesson objective, presented as notated or recorded exercises, was produced from the descriptions from formal pedagogy and historical documentation.

CHAPTER 3

Methodology

The purpose of this study was to compare measurements of harmonic accuracy, following the use of notated and aural exercises as experimental procedures of jazz improvisation pedagogy. To fulfill this purpose, procedures and materials were developed for each treatment group. A pilot study was conducted, resulting in modifications to the measurement instrument. Students for the main study were recruited from local high school band programs. The pretest-posttest evaluations provided data for the comparisons of each method of pedagogy.

Development of Treatment Procedures

Lessons, curricula and materials were developed and used as treatment and evaluation procedures. Lesson plans for both the aural and notation methods of pedagogy used the same learning objective (demonstration of harmonic accuracy). Materials developed included lesson materials used during the treatment sessions, student take-home materials (recorded aural or notated exercises), and evaluation materials.

Development of Lesson Plan

A lesson plan was developed (Appendix A) that included the same lesson objective for each group. The lesson objective was created in a conceptual manner that could be applied to the aural and notated methods of pedagogy. The instructional objective was identified as the demonstration of harmonic accuracy. Damron (1973) measured experimental procedures involving beginning jazz high school students. His

use of the dorian mode for this level of student was common among jazz pedagogues. In order to measure harmonic accuracy, a dorian mode chord sequence involving two distinctive areas of tonalities was used. This was chosen in order to measure a student's success with harmonic accuracy involving tonal modulation. It was the opinion of this researcher that the demonstration of tonal modulation from beginning level improvisers was a reasonable performance trait for a single lesson plan.

The Miles Davis song "So What" was selected as a recognized song from the jazz tradition that used a modal harmonic structure. The instructional objective was administered through the development and implementation of an instructional curriculum (included as Appendix A). This curriculum was developed from the materials in order to fulfill the lesson objective.

A lesson outline was written, including six units of instruction, to be presented in a single session. Unit one consisted of presenting and practicing the primary scale and arpeggio material. Unit two provided knowledge and performance experience regarding the use of scale and arpeggio materials within the selected harmonic structure. Unit three provided single measure melodic patterns, similar to imitative exercises used by Segress (1979) and Aitken (1975). These exercises, intended to be performed by the instructor, were to be repeated during the subsequent measure by the students. Unit four consisted of a listening example and an analysis by Reeves (1995) of the Miles Davis solo. Unit five provided to the student an alternative solution for the instructional objective, a procedure used by Segress (1979). Unit six allowed each student to improvise alone with a pre-recorded rhythm section of professional jazz musicians (Aebersold, 1991).

Lesson Materials

Materials used during the treatment sessions included notated or recorded exercises consisting of scales and arpeggios from which the modal tonalities were based. Additional lesson materials included a pre-recorded accompaniment (Aebersold, 1991), a jazz solo example (Davis, 1959) and an analysis of the solo excerpt (Reeves, 1995).

According to Segress (1979) the presentation of an alternative solution is a recommended feature of the "Systems Approach Model" that he used in his study. In the analysis of the Davis solo, Reeves (1995) identified the dorian scale as the primary harmonic strategy. He also identified two other scale types, the harmonic minor and blues scales. The harmonic minor scale was selected as an alternative solution as demonstration of the instructional objective.

The accompaniment used during the treatment sessions, and for pre- and post-test recordings was selected from Jamey Aebersold's (1991) Vol. 54, "Maiden Voyage" (letter of permission as Appendix H). Track 2 from this volume was used as a "modal" progression involving two dorian tonalities.

A listening example of the Miles Davis solo on the tune "So What," (Columbia CL1355) was presented to Group A (aural) during each of the four sessions. Group B (notation) practiced the solo from a notated transcription (produced by this researcher) during each session (Appendix C). This material was presented to students in the classroom and not reproduced as take-home material.

A portion of the Reeves (1995) analysis (contained within Appendix A) was read to the students prior to listening to, or during the study of, the transcription. Within the

quote read to students, Reeves brings attention to melodic aspects and strategy; this provided the only instruction regarding melodic aspects of improvisation.

Production of Take-Home Materials

Materials were given to students for take-home practice. A computer-generated accompaniment, using the software "Band In A Box for Windows, version 8" of the selected chord progression was produced on CD and given to students from both groups as take-home material. Copies of the computer-generated accompaniment were given to students due to copyright protections of the Aebersold recording. Group A (notation) received notated exercise sheets transposed for their instrument (Appendix B), and a CD recording of the accompaniment. Group B (aural) received a CD recording of the exercises performed by the researcher, and a recording of the accompaniment.

Notated exercise sheets were produced using the computer software "Finale 2000." Separate sheets were produced for each instrument, transposed and within standard playing ranges (Appendix B). Exercises from the notated sheets were performed by the instructor and recorded to a CD. The accompaniment used for the exercises, as well as the separate track accompaniment, were produced by the computer software "Band In A Box for Windows," version 8. The style used within the program was "j~oleo.sty," with the "A" style selected for the "A" section, and the "B" style selected for the bridge.

Development and Modification of the Performance Measure

A performance evaluation instrument (Appendix J) was developed to measure the research questions specific to this study, and to provide related items which could be used to investigate internal consistencies. Evaluation devices that had been validated from

previous research studies were used as models, including Burnsed and Price (1984), Pfenninger (1990), Segress (1979), and Tumlinson (1991). Items from these studies that pertained to the research questions of the present study were considered or modified for the new measurement instrument.

A newly developed instrument was needed rather than using one of the existing instruments, due to the purpose of the present study. Rather than a general measurement of all performance aspects, the primary lesson objective, harmonic accuracy, became the specific performance area to be evaluated.

A preliminary evaluation instrument was developed and used for the pilot study data. As a result of analyses, modifications to the instrument included eliminating general evaluation items that did not measure one of the research questions of this study. Duplication of items that evaluate performance characteristics from the research questions were also added in order to assess internal consistencies. Although Tumlinson's (1991) “descriptive improvisation measurement instrument” (DIMI) was formulated to evaluate all performance aspects of a jazz solo, four items that measure the characteristics pertinent to this study were used. These items were included within the new evaluation instrument for the purposes of this study.

A total of 6 Likert-scale (range = 1-5) items were developed for the evaluation instrument, with the addition of item number Q7 to allow further written comments from the evaluators. Item numbers Q2 – Q5 were drawn directly from Tumlinson’s (1991) DIMI. Item numbers Q1 and Q6 were created to assess instrument internal consistencies.

Questions intended to measure the performance criterion, harmonic accuracy, were correlated to determine internal consistencies from the evaluation instrument.

Results were generated using a Cronbach's Alpha test to compare combined pretest and posttest scores among the related question items. The results of this comparison produce Alpha = .97, indicating that this group of related questions consistently produced similar scores (good internal consistencies). The Cronbach's Alpha statistic indicates good correlations among questions measuring the performance objective, harmonic accuracy. Thus, a high degree of internal consistency was demonstrated, therefore, reliability of the evaluation instrument was inferred.

Pilot Testing Procedures

Local high school students were recruited from the same school who had not participated in a jazz band or had not experienced performing a jazz solo. A letter (Appendix D) explaining the extent of their participation was sent home with volunteer students, along with the "human subjects consent form" (Appendix F), to be signed by the students and their parents. Two groups of 8 students (n = 16) were asked to attend two sessions during a one week period in January 1999.

The first session consisted of collecting the consent forms, directly followed by recording each student playing an improvised solo with the CD recorded accompaniment. No instruction was given prior to the pre-treatment recording of each student. Following the pre-treatment recording session, materials were presented to each group with instructions given on how to practice (one hour minimum) prior to the second session two days later. During the second session the same lesson plan was presented as a review of materials given during the first session. A post-treatment recording of each student was made and used as post-test observations.

Analysis of Pilot Data

The pre- and post-treatment recordings were recorded on digital audio tape (DAT) and transferred to CD. Students from both groups and sessions were randomly assigned to separate tracks on the CD recording. A preliminary version of the evaluation instrument was developed using a 5-point Likert type scale. Two judges evaluated the 32 separate track performances using the preliminary version of the measurement instrument (finalized version provided as Appendix J). Inter-rater reliability coefficients were calculated for each item from the measurement instrument, producing values ranging from $r = .12$ to $r = .78$. Descriptive statistics were calculated from pretest and posttest scores for each treatment method. The results produced higher mean values for the aural method than for the notation method. Comparisons of pretest to posttest scores between treatment methods using a Mann-Whitney test (for ordinal type data) indicated that statistical significance was not achieved for all evaluation instrument items.

As a result of the pilot testing, procedural changes that occurred prior to the main study included a modification of the evaluation instrument. The lesson plan and materials were not modified based on the results of the pilot study.

Student Recruitment

A letter describing the need for student volunteers to be involved in a research study was sent to each of four Amarillo, TX high school band directors. The directors were then contacted regarding student involvement and time requirements, two schools agreed to participate in the study. Due to the band schedules and extra rehearsals during the time period, directors requested that their students not participate beyond the month of September. Directors consented to one session per week for a total of four weeks. This

was sufficient in the opinion of the researcher to gain significant data results from a single repeated lesson plan.

An enrollment sheet was taken to each school with two available times and the dates of the four-week course. Students were not informed regarding the presence of varying pedagogy. Students and directors were informed that volunteers were needed for a jazz improvisation research study at no cost to students. Randomization between schools and groups occurred by allowing students to "blindly" sign up for their most convenient time. Students were instructed not to volunteer unless all sessions could be attended. Letters (Appendix E) were made available as students signed up at their local schools.

The enrollment sheets were later collected and students were contacted; information was collected and a human subjects consent form was mailed to each student's parents. Although two local high schools consented and approved the treatment times and period of involvement, students from one of the schools were unable to participate with the earlier time period due to a rehearsal conflict. As a result the earlier group (notation) was comprised entirely of students from the same school, while the second group (aural) included students from both schools. Because the conflict arose after the schools' participants and researcher were under way with the study, a decision was made to continue.

The earlier group (notation) had a sample size of 12 while the latter group (aural) had a sample size of 8. A power calculation was figured prior to student participation in order to determine the probability of significance based, in part, on the sample size. Considering that statistically, 99.7 percent of a normally distributed population falls

within ± 3 standard deviations, scores from a 5-point Likert scale could assume a S. D. of 0.7. A power calculation with a power of 0.8 determined that group sample sizes of n = 8 and n = 12 would be sufficient to detect a deviation of ± 1, with an estimated S. D. of 0.7.

Student Participation

Student participation began with the collection of consent forms that had been signed by the students and their parents. Questionnaires also were completed prior to the pre-treatment recording session. A single lesson plan was presented during each of the four sessions, with a post-recording session concluding student involvement. Student performances were recorded to Digital Audio Tape (DAT) and cassette tape as backup. A total of 20 students who participated in the main study each performed a pre- and a post-treatment solo. A total of 40 student solos were randomly assigned to separate CD recording tracks, which were mailed to each of six evaluators.

Use of Evaluators

Evaluators were contacted during November 2000 who agreed to participate with this study. A total of 6 evaluators representing three levels of professional assignments were selected. Two high school jazz band teachers, two graduate teaching assistants from a jazz degree program, and two university faculty who direct a jazz ensemble were used as evaluators.

Evaluator materials, including a cover letter with evaluator instructions (Appendix I), the evaluation instrument (one per CD track), and the CD of randomized student performances were mailed in early November 2000. As the evaluator data arrived to this researcher, a computer spreadsheet was formatted using Microsoft Excel 2000 and the

data were entered. By February 2001, the evaluator results were received and entered into the computer spreadsheet.

CHAPTER 4

Results

Although a degree of randomness was predicted through participants' blind choice of treatment groups, local band schedules beyond the control of this study prevented the participants in one school from choosing between the treatment sessions. This resulted in an unbalanced sampling from the two participating schools, creating conditions for a pretest – posttest quasi-experimental design. Mitchell & Jolley (1988) describe the process of quasi-experimental design.

> …the challenge in quasi-experiments is inferring ceteris paribus without the aid of random assignment or control of relevant variables. The first step to meeting this challenge is to be aware of all the variables that might account for a relationship between the treatment and the effect. Once you've identified these variables, you'll try to demonstrate that these variables didn't account for the relationship. (p. 242)

These authors state that "all possible threats to ceteris paribus fall under the eight threats to internal validity," categorized as the spurious eight (p. 242).

Eight causes of type 1 error that must be accounted for, according to these authors, include history, maturation, mortality, testing, instrumentation, selection, statistical regression and selection–maturation interaction. Campbell & Stanley (1966) explain that one method to account for selection-maturation is to reveal equivalent pre-treatment sample groups. A comparison of pretest scores between method groups would

satisfy this assumption. Because no student dropped out of the study, mortality error can be ruled out. Instrumentation error was accounted for by 1) randomly assigning both pretest and posttest recordings to one CD recording, 2) using the same performance measure for pretest and posttest evaluation, 3) observing interjudge reliability, and 4) observing internal consistencies of the performance measure.

Regression, according to Campbell & Stanley (1966), can be compensated for by the use of an analysis of covariance. The assumptions of ANCOVA include homogeneity of regression, which must be met prior to assuming the results. A comparison of pretest scores also will account for selection, as both groups were observed as statistically equivalent prior to the treatment period. Maturation was considered during the scheduling of the treatment period that lasted for four weeks. A condition of this study required the use of high school students who had not, previously or currently, participated in a jazz band. Because both groups were observed as statistically equivalent on pretest scores, and students were not participating in jazz activity elsewhere during the course of the study, a comparison between posttest scores by method would account for variation between treatment procedures. Likewise, testing may also be accounted for by comparisons between groups of pretest scores, posttest scores, and from pretest to posttest by method.

Mitchell & Jolley (1988) summarize that causality can be inferred from a quasi-experimental design when temporal precedence is established, covariation is observed, and the spurious eight are ruled out. The pretest-posttest design by nature establishes temporal precedence due to the nature of the design, as treatment follows pretesting. Covariation will be discussed in terms of the research model and its resulting relationships.

The Research Model

The research model in Figure 1 illustrates some of the observations and analyses undertaken in this study. The relationships of each aspect of this model will address inter-rater reliability, measurement consistencies, and the probable effect from each treatment method. A comparison of pretest scores between groups will provide an estimate of equivalent samples from the population. Pretest to posttest scores by method will also be investigated to determine treatment effects as well as a comparison of posttest scores between methods to reveal treatment differences.

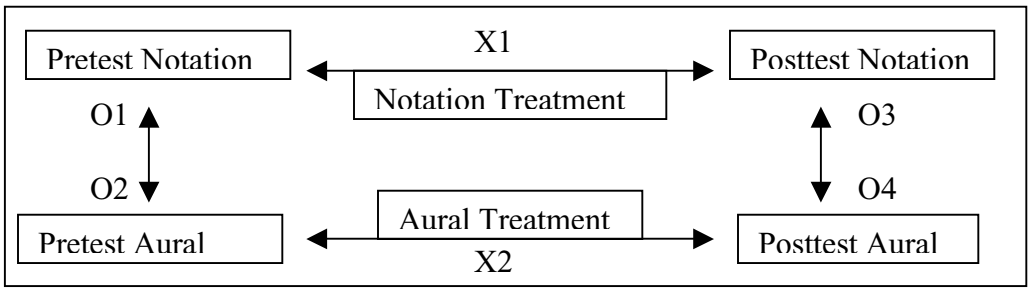


Figure 1. Interactions within the Research Model.

Procedures were chosen based upon the interactions of observations and treatment methods illustrated in the research model; the research questions will be discussed following the results of these analyses. Inter-rater reliability and internal consistency estimates of the measurement instrument also were investigated. Because specific items for the measurement instrument were developed to evaluate the research questions from this study, data from individual evaluation items were generated and discussed in relation to the research questions. A comparison of pretest scores between method groups provided pre-treatment equivalencies. Observations from pretest to posttest scores by method provide data from which to describe treatment effects. Finally, a comparison of posttest scores between methods produced descriptive and inferential data of post-

treatment differences between the two methods. An analysis of covariance was calculated for each question item, from pretest to posttest scores, by treatment effects between student groups, while accounting for regression type errors.

Inter-Rater Reliability

Six professional jazz pedagogues evaluated students' pretest and posttest performances. Reliability estimates were calculated from the score values assigned by the six judges (2 high school teachers, 2 graduate teaching fellows, 2 university teachers). For pretest and posttest scores combined, reliability estimates using a Spearman's rho coefficient were calculated. All estimates were statistically significant at the .05 level. In general, the highest correlations occurred between like evaluators (judge 1 & judge 3 = university teachers; judge 2 & judge 4 = high school teachers; judge 5 & judge 6 = graduate teaching fellows) while the least agreement occurred across evaluator types. Numeric values were assigned to each judge in the order in which the data were received and entered into the data spreadsheet. Judge assignment values are illustrated in Figure2.

Judge 1	University Teacher 1
Judge 2	High School Teacher 1
Judge 3	University Teacher 2
Judge 4	High School Teacher 2
Judge 5	Graduate Teaching Fellow 1
Judge 6	Graduate Teaching Fellow 2

Figure 2. Judge numeric value assignments.

Table 1 lists the highest and lowest coefficient values between judges for each evaluation item. The highest value from all questions occurred with item Q3, producing r = 0.845 between J2 and J3 (J2 = high school teacher, J3 = university teacher).

Table 1.

Inter-Rater Reliability Estimates for Combined Pretest and Posttest Scores.

Q1	Highest Correlation	r = 0.778	J3 to J2
	Lowest Correlation	r = 0.465	J4 to J6
Q2	Highest Correlation	r = 0.830	J3 to J5
	Lowest Correlation	r = 0.502	J3 to J6
Q3	Highest Correlation	r = 0.845	J2 to J3
	Lowest Correlation	r = 0.563	J3 to J6
Q4	Highest Correlation	r = 0.747	J3 to J4
	Lowest Correlation	r = 0.523	J4 to J5
Q5	Highest Correlation	r = 0.780	J1 to J2
	Lowest Correlation	r = 0.508	J3 to J5
Q6	Highest Correlation	r = 0.876	J2 to J3
	Lowest Correlation	r = 0.532	J2 to J4

Conrad (1997) discussed inter-rater reliability estimates among musical performance evaluators. Reliability estimates are discussed from various types of music research and performance measures, indicating that estimates > .50 are described as having a high consistency. All inter-rater correlations produced r values above 0.5, which, according to Conrad, may be interpreted as good reliability. This would indicate that for most of the correlation estimates between evaluators from this study, the consistency is high when compared to other musical performance measures from the literature.

Because each judge appeared consistent from the amount of change from pretest to posttest scores, this researcher was prompted to investigate posttest scores minus

pretest scores. This would, in essence, compare the amount of change from pretest to posttest scores. When manipulating the raw data for posttest minus pretest scores, a missing score for either session resulted in a blank assignment rather than a score of 0.0 (which would indicate no change). Although not all judges completed all of the question items, missing fields were allowed by the computer software, SPSS version 10.0, without conflicting with the analysis results.

Comparisons among judges' scores using a Kruskal-Wallis ANOVA were calculated using both pretest and posttest scores combined, and with posttest minus pretest scores (Table 2). The Kruskal-Wallis procedure was used because it is the non-parametric equivalent to the ANOVA comparison of more than 2 levels of the independent variable (judge n = 6). The null hypothesis that judges' scores were the same is rejected when the p value is significant ($p < 0.05$). The data displayed in Table 2 reveal that for pretest and posttest scores combined, the null hypothesis is rejected for all question items. This suggests that for raw score data, student scores by judge were different. When comparing posttest minus pretest scores by judge, most revealed p values greater than 0.05. Therefore, the null hypothesis is not rejected, indicating that the judges' score differences (variation) were not statistically significant. Thus, there is not sufficient evidence to reject the claim that posttest minus pretest scores differed by judge. When comparing the significant questions listed in Table 2, fewer items reached statistical significance for posttest minus pretest scores, indicating that the judges' scores were statistically equivalent for these evaluation items.

Table 2.

Between Judge Comparisons, Kruskal-Wallis ANOVA

Question	pretest and posttest	posttest minus pretest
Q1	H = 20.83, p = 0.0009 *	H = 9.86, p = 0.079
Q2	H = 24.95, p = 0.0001 *	H = 6.83, p = 0.234
Q3	H = 23.35, p = 0.0003 *	H = 4.89, p = 0.429
Q4	H = 30.11, p = 0.0001 *	H = 3.84, p = 0.573
Q5	H = 37.24, p = 0.0001 *	H = 2.58, p = 0.765
Q6	H = 18.98, p = 0.0019 *	H = 12.76, p = 0.026 *

(* $p \leq 0.01$, ** $p \leq 0.05$)

These data suggest that while judges' scores varied with respect to absolute values (both pretest and posttest), they did not differ with respect to relative change in scores from pretest to posttest. For example, Figure 3 presents the scores for evaluation item Q1 from each judge for one particular student. While the absolute score values are different (J6 = 1 while J4 = 5) the differences from posttest to pretest scores are more equal.

Q1 for student 6A	Pretest	Posttest	Post minus Pre
J1	2	2	0
J2	2	4	2
J3	3	3	0
J4	5	5	0
J5	3	2	1
J6	1	1	0

Figure 3. Comparison of absolute score values and posttest minus pretest values.

Pre-Treatment Equivalence

The pretest / posttest quasi-experimental design must establish a level of statistical equivalence between pre-treatment groups, according to Campbell & Stanley (1966), prior to a discussion of treatment effect. A comparison of pretest scores from the

two sample groups was conducted using a Mann-Whitney test due to the comparison of the two variables represented by ordinal type data. The data (Table 3) suggest that the two groups did not differ in a statistically significant manner.

Table 3.

Pretest Equivalence, Mann-Whitney Comparisons.

	Notation Method n = 78	Aural Method n = 42
Q1	U = 1326.5, p = 0.063	
Q2	U = 1331, p = 0.089	
Q3	U = 1319.5, p = 0.075	
Q4	U = 1174, p = 0.128	
Q5	U = 1780, p = 0.403	
Q6	U = 1310.5, p = 0.084	

When applying the Bonferroni's adjustment for 17 comparisons ($p<0.05/6$ which is equivalent to $p<0.008$) to detect if there were statistical differences between the pretest aural and notation groups by student scores, no statistical differences were found. Conclusions (when applying a Bonferonni's adjustment to correct for type 1 error) indicate that there was not sufficient evidence to conclude that the treatment groups differed.

Comparisons by Method

Figure 1 illustrates the appropriate comparisons for the research model, including comparisons by method using pretest scores (pre-treatment equivalence), from pretest to posttest by method (treatment influence) and between posttest scores by method (treatment differences). Because the two sample groups were statistically equivalent prior

to the treatment sessions, the degree of change from pretest to posttest scores would indicate a probable effect caused by either the experimental notation or aural treatment methods.

Table 4 lists the results from a Mann-Whitney comparison (ordinal type data) of pretest to posttest scores for each treatment method (notation and aural). The mean and standard deviation values for each observation (pretest and posttest scores) are also provided, although they were generated and are reported as descriptive data. The Mann-Whitney probability values for each method as compared from pretest to posttest scores were all significant at the $p<0.05$ level. When the Bonferroni's adjustment for 24 comparisons ($p<0.0021$) was revealed, all items continued to reach statistical significance. Descriptive analyses of these significant items reveal that the mean score values for the aural method were greater than for the notation group.

When comparing pretest to posttest scores by method, the aural method produced greater increases in the mean score values. For example, evaluation item Q2 produced a pretest aural mean value of 1.62 ± 0.69 (mean value ± standard deviation), while the posttest aural mean value produced 3.36 ± 1.10. Evaluation item Q2 produced a pretest notation mean value of 1.96 ± 0.98, and a posttest notation mean value of 2.78 ± 1.09. The descriptive data (mean value and standard deviation) for evaluation items that reached statistical significance indicate that the aural method produced greater increases for student scores from pretest to posttest in comparison to the notation method.

Table 4.

Comparisons of Pretest and Posttest Scores by Method

	Pretest	Posttest	Mann-Whitney
Q1 Notation	1.86 ± 0.89	2.64 ± 1.03	U = 4303, p = 0.0001 **
Q1 Aural	1.52 ± 0.63	3.19 ± 0.99	U = 1592.5, p = 0.0001 **
Q2 Notation	1.96 ± 0.98	2.78 ± 1.09	U = 4128.5, p = 0.0001 **
Q2 Aural	1.62 ± 0.69	3.36 ± 1.10	U = 1575, p = 0.0001 **
Q3 Notation	1.95 ± 1.03	2.82 ± 1.13	U = 4163, p = 0.0001 **
Q3 Aural	1.57 ± 0.70	3.36 ± 1.01	U = 1605, p = 0.0001 **
Q4 Notation	2.00 ± 1.05	2.72 ± 1.13	U = 3818.5, p = 0.0001 **
Q4 Aural	1.66 ± 0.78	3.19 ± 1.02	U = 1217, p = 0.0001 **
Q5 Notation	1.79 ± 0.92	2.45 ± 1.05	U = 4136, p = 0.0001 **
Q5 Aural	1.95 ± 0.99	2.95 ± 1.10	U = 1307, p = 0.0001 **
Q6 Notation	1.95 ± 0.94	2.69 ± 1.06	U = 4233.5, p = 0.0001 **
Q6 Aural	1.63 ± 0.77	3.22 ± 1.01	U = 1475.5, p = 0.0001 **

A Mann-Whitney comparison of posttest scores between the method groups indicated that all evaluation items were statistically equivalent. When comparing the descriptive data in Table 5, for those items which reached statistical significance, the aural method produced higher mean score values than the notation method. Evaluation item Q2, for example, produced a posttest notation mean value of 2.78 ±1.09, as compared to the posttest aural mean value of 3.36 ± 1.10.

Table 5.

Comparisons of Posttest Scores by Method

	Notation	Aural	Mann-Whitney
Q1 Post	2.64 ± 1.03	3.19 ± 0.99	U = 2116.5, p = 0.0062 *
Q2 Post	2.78 ± 1.09	3.36 ± 1.10	U = 2045, p = 0.0089 *
Q3 Post	2.82 ± 1.13	3.36 ± 1.01	U = 2016.5, p = 0.0145 *
Q4 Post	2.72 ± 1.13	3.19 ± 1.02	U = 1741, p = 0.0341 *
Q5 Post	2.45 ± 1.05	2.95 ± 1.10	U = 2039, p = 0.0221 *
Q6 Post	2.69 ± 1.06	3.22 ± 1.01	U = 2025, p = 0.0135 *

Because the between judge comparisons (Table 2) produced higher correlations when using posttest minus pretest score values, these values should be investigated when comparing treatment effects. Table 6 provides descriptive data (means and standard deviations) and Mann-Whitney values from comparisons of posttest minus pretest scores compared between methods. Although all question items reached statistical significance, information from these analyses produced results unique to the results produced from absolute score values (Table 5). The mean values generated from posttest minus pretest scores represent the amount of change from pretest to posttest for each question. An examination of these values indicates that the aural method produced a greater increase for all items.

Table 6.

Posttest minus Pretest scores between Methods.

Questions	Notation Method	Aural Method	Mann-Whitney
Q1	0.782 ± 0.962	1.667 ± 1.119	U = 2403, p = 0.0001 *
Q2	0.769 ± 1.248	1.738 ± 1.191	U = 2332.5, p = 0.0001 *
Q3	0.821 ± 1.225	1.786 ± 1.116	U = 2333, p = 0.0001 *
Q4	0.756 ± 1.311	1.310 ± 1.179	U = 2049.5, p = 0.0198 *
Q5	0.654 ± 1.115	1 ± 0.988	U = 1910.5, p = 0.1174
Q6	0.744 ± 1.122	1.548 ± 1.435	U = 2263, p = 0.0004 *

An analysis of covariance is appropriate, according to Campbell & Stanley (1966), in order to account for regression error. Because the ANCOVA is a test of adjusted mean values controlling for variation in pretest scores (covariate), it compares posttest scores as if the pretest scores were statistically equal (adjusted). These authors state: “simple gain scores are also applicable but usually less desirable than analysis of

covariance" (p. 49). Although Table 6 presents the simple gain scores, (in terms of posttest minus pretest) an analysis of covariance is also warranted for each question item. Hinkle, Wiersma and Jurs (1988) explain that a test of the homogeneity-of-regression assumption must be met prior to interpreting the ANCOVA results. A linear regression model from pretest to posttest scores was tested for all questions (data presented as Appendix K). A comparison of the confidence interval for slopes from the regression models for aural and notation methods was done for each question as a test for homogeneity of regression. In all cases the slopes did not differ in a statistically significant manner, indicating that the homogeneity-of-regression assumption has been satisfied for all evaluation items.

An analysis of covariance was tested, defining posttest scores as the dependent variable, treatment method as the independent variable, and pretest scores as the covariate. The results from these analyses, summarized as Table 7, verify the results from posttest minus pretest comparisons presented in Table 6. The evaluation items that reached statistical significance for the simple gain score analyses were also significant from the ANCOVA results.

Table 7.

ANCOVA comparisons of posttest scores between methods.

Q1	F (1,117) = 8.4, p = .0*
Q2	F (1,114) = 5.6, p = .0*
Q3	F (1,114) = 3.9, p = .0*
Q4	F (1,104) = 6.4, p = .0*
Q5	F (1,117) = 2.6, p = .1
Q6	F (1,115) = 3.8, p = .0*

The results from the ANCOVA for evaluation item Q2 are presented as Table 8. The "tests of between-subjects effects" reveals that for this item (Q2), the level of significance was 0.019 for the main effect from method. This may be interpreted as an indication that Q2 produced significantly different scores by method for posttest observations when controlling for pretest score variation. The "descriptive statistics" reveal a higher posttest mean value for method 1 (aural, 3.21 ± 1.09, $n = 48$) than for method 2 (notation, 2.81 ± 1.13, $n = 69$). For all significant items revealed from the ANCOVA results (data presented as Appendix K), the aural method produced higher posttest mean values than the notation method.

Table 9 provides the mean values and standard deviations produced from absolute score values (presented in Table 5) and the adjusted values from the ANCOVA results. The statistic and significance also are presented in order to compare the influence from the ANCOVA adjustment. For all question items, the mean values for each method group had a tendency to move toward the combined score mean value of both methods. Because all notation group mean values were below this average, the ANCOVA adjustment values became higher, while the aural group values produced lower adjusted values because the aural means were higher than the mean score averages.

Comparison by Method and Grade Level

A two-way ANOVA comparing students' grade level to treatment method revealed that while grades 10, 11 and 12 were statistically homogenous, grades 9 and 12 differed. Table 8 lists the results of the two-way ANOVA for evaluation item Q2 (data for all items presented as Appendix M).

Table 8.

ANCOVA comparisons of posttest scores between methods for Q2.

Univariate Analysis of Variance

Between-Subjects Factors

		n
Method	1	48
	2	69

Descriptive Statistics

Dependent Variable: Q2POST

		Mean	Std. Deviation	n
Method	1	3.21	1.09	48
	2	2.81	1.13	69
	Total	2.97	1.13	117

Levene's Test of Equality of Error Variances

Dependent Variable: Q2POST

F	df1	df2	Sig.
.006	1	115	.936

Test the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept+Q2+Method

Tests of Between-Subjects Effects

Dependent Variable: Q2POST

		Type III Sum of Squares	df	Mean Square	F	Sig.
Source	Corrected Model	12.007	2	6.004	5.073	.008
	Intercept	138.734	1	138.734	117.227	.000
	Q2	7.552	1	7.552	6.381	.013
	METHOD	6.673	1	6.673	5.639	.019
	Error	134.916	114	1.183		
	Total	1182.000	117			
	Corrected Total	146.923	116			

Table 9.

Comparisons of Absolute and Adjusted Posttest Scores by Method

	Notation	Aural	
Q1 Post [Absolute] ANCOVA	[2.64 ± 1.03] 2.60 ± 1.0	[3.19 ± 0.99] 3.0 ± 0.9	[U = 2116.5, p = 0.0062] * F = 8.401, p = .004 *
Q2 Post [Absolute] ANCOVA	[2.78 ± 1.09] 2.81 ± 1.13	[3.36 ± 1.10] 3.21 ± 1.09	[U = 2045, p = 0.0089] * F = 5.639, p = .019 *
Q3 Post [Absolute] ANCOVA	[2.82 ± 1.13] 2.8 ± 1.1	[3.36 ± 1.01] 3.1 ± 1.1	[U = 2016.5, p = 0.0145] * F = 3.972, p = .049 *
Q4 Post [Absolute] ANCOVA	[2.72 ± 1.13] 2.76 ± 1.14	[3.19 ± 1.02] 3.22 ± 0.99	[U = 1741, p = 0.0341] * F = 6.417, p = .013 *
Q5 Post [Absolute] ANCOVA	[2.45 ± 1.05] 2.50 ± 1.06	[2.95 ± 1.10] 2.81 ± 1.12	[U = 2039, p = 0.0221] * F = 2.641, p = .107
Q6 Post [Absolute] ANCOVA	[2.69 ± 1.06] 2.76 ± 1.07	[3.22 ± 1.01] 3.07 ± 1.06	[U = 2025, p = 0.0135] * F = 3.821, p = .053

Results from the posthoc tests from the two-way ANOVA presented in Table 10 list the mean difference, standard error and level of significance between each grade level. Although comparisons for item Q2 were not statistically significant (p = 0.354) when comparing student grade level and treatment method, the results presented as Table 10 reveal 9th grade scores differed significantly from 12th grade scores.

Table 11 provides a comparison of 9th and 12th graders' posttest scores by method (notation and aural) for evaluation item Q1. The results of a Mann-Whitney test revealed that the aural group was equivalent (p = 0.1797) while the notation group differed (p = 0.0002). A comparison of the mean rank values for the notation method indicates that 9th graders scored lower than 12th graders. These results revealed that although posttest student scores from the aural treatment group remained equivalent between 9th and 12th graders, student scores from the notation treatment group were different.

Table 10.

Two-Way ANOVA Comparisons of Grade Level by Treatment Method.

Tests of Between-Subjects Effects
Dependent Variable: Q2

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	38.735[a]	8	4.842	4.832	0.000
Intercept	839.754	1	839.754	837.978	0.000
Grade Level	19.606	4	4.901	4.891	0.001
Treatment Method	7.671	1	7.671	7.654	0.007
Grade Level * Treatment Method	3.293	3	1.098	1.095	0.354
Error	109.231	109	1.002		
Total	1198.000	118			
Corrected Total	147.966	117			

[a.] R Squared = 0.262 Adjusted R Squared = 0.208)

Post Hoc Tests (Absolute Values)

		Multiple Comparisons			95 % Confidence Interval	
(I)	(J)	Mean Difference	Std. Error	Sig.	Lower Bound	Upper Bound
10	11	.43	.27	.621	.40	1.27
	12	.41	.25	.620	1.19	.38
	9	.68	.27	.169	.15	1.52
11	12	.84	.27	.058	1.70	1.74E-02
	9	.25	.29	.945	.66	1.16
12	9	1.09*	.27	.005 *	1.95	.23

Table 11.

Mann-Whitney Comparisons of Posttest Scores by 9th and 12th Grade Level.

Q1 Post Aural	n	Rank Sum	Mean Rank	U
9th Graders	12	99.0	8.25	51.0
12th Graders	6	72.0	12.00	21.0
2-tailed p	0.1797			

Q1 Post Notation	n	Rank Sum	Mean Rank	U
9th Graders	12	118.0	9.83	248.0
12th Graders	24	548.0	22.83	40.0
2-tailed p	0.0002			

CHAPTER 5

Summary, Discussion and Recommendations for Further Research

Summary

Aural and notated methods of pedagogy are common to jazz improvisation. Their comparative effectiveness, however, had not been investigated. The use of beginning level improvisers provided for a controlled experimental study. Harmonic accuracy was described as an important performance variable to jazz improvisation pedagogy and evaluation, and an appropriate learning objective for beginning level students.

The measured effects to harmonic accuracy, rather than observing the effects to all performance characteristics, was chosen in order to produce a single lesson plan developed for beginning level jazz improvisers. Evaluation procedures developed and used by researchers who had addressed harmonic performance issues, were also reviewed in order to develop an appropriate evaluation procedure for the purpose of this study.

The research questions for this study were developed as a logical progression of the research literature. Because observable characteristics of jazz improvisation had been investigated and measured, this study observed the effects of varied pedagogy on a single performance aspect selected from the literature. Because the research questions were developed in order to clarify issues documented in the research literature, the results from this study will be addressed as implications to the literature.

A lesson objective was identified from the primary research question followed by the development of lesson materials, curriculum and an outline of instruction for the two

methods of pedagogy. Each experimental treatment method used the same lesson objective and curriculum. Whereas lesson materials were presented to the notation group with written musical examples, they were presented to the aural group as recorded examples on a CD. Students were recruited from local high schools, producing 12 volunteers for the notation group and 8 volunteers for the aural group. A schedule was approved between the researcher and high school band directors, resulting in a treatment period of four weeks.

Data collection consisted of recording each student improvising a jazz solo with a pre-recorded accompaniment both prior to and following the treatment sessions. These solo performances were randomly assigned to CD tracks and mailed to six professional jazz pedagogues acting as evaluators. A performance evaluation instrument was developed in order to observe the specific research questions of this study. Each evaluator received instructions, a separate evaluation sheet for each CD track, and the randomized CD of students' pretest and posttest performances. As this researcher received the completed evaluations, the data were entered into a computer spreadsheet followed by statistical procedures using the computer software SPSS, version 10.0.

The research design and model were identified and discussed in relation to the research questions and the resulting appropriate analysis procedures. As a result of the use of a five-point Likert scale, non-parametric procedures (with the exception of the two-way ANOVA, linear regression, and ANCOVA) were used for simple comparisons of the ordinal type data. Inter-rater reliability was estimated between all evaluators, revealing higher correlations between similar evaluator types. Student score values represented as the difference from posttest to pretest scores produced more uniform

values among the judges than pretest and posttest raw score values. This indicated that although judges assigned different absolute score values to a particular student performance, their values increased or decreased similarly from pretest to posttest observations. The evaluation instrument was tested for internal consistencies in order to assess an estimate of correlation between similar questions. Items were produced in order to assess internal consistencies, which produced Alpha = .98.

Due to conditions beyond the control of this study, an unbalanced sampling from the two participating high schools occurred. In accordance with procedures identified for the "quasi-experimental design," it was necessary to determine that each treatment group produced equivalent scores prior to the treatment sessions. Pre-treatment equivalencies were compared for each question item using a Kruskal-Wallis ANOVA test. When accounting for the combined estimate for error, all question items were statistically equivalent.

Comparisons of student scores produced between the aural and notation treatment methods were conducted using the Mann-Whitney test. Descriptive data were also generated, including the mean value and standard deviation for each pretest and posttest evaluation item. When comparing pretest and posttest scores by method, all items produced greater mean values for posttest than for pretest observations. When comparing the mean values produced between the two methods, the aural treatment group revealed greater increases from pretest to posttest values.

Data from posttest scores were compared between the method groups, indicating that for evaluation items that reached statistical significance, the aural method produced greater mean score values. Further analyses of treatment effect were conducted using

posttest minus pretest score values in order to compare the amount of change between sessions.

Campbell & Stanley (1966) explain that an analysis of covariance is a recommended analysis for the quasi-experimental study in order to account for regression error. Prior to this procedure, however, a test for homogeneity of regression was met by using a test of linear regression between pretest and posttest scores. The results of the ANCOVA verified the significant evaluation items from the posttest minus pretest comparisons by method.

A comparison between the two treatment methods and student grade level was fulfilled using a two-way ANOVA procedure. These analyses revealed that although students from grade levels 10, 11 and 12 produced similar evaluation results, 9^{th} graders differed from 12^{th} graders. The F tables for all question items are presented as Appendix L. A post hoc analysis and a Mann-Whitney comparison for evaluation item Q1 revealed that while 9^{th} graders involved in the aural treatment method remained equivalent to 12^{th} graders, those involved in the notation group produced lower scores. Appendix N presents the Mann-Whitney comparisons for all question items.

Discussion

Research Question 1

What were the effects on high school band members who had not participated in a jazz band or experienced performing a jazz solo, from the use of notated and aural exercises, involving the development of harmonic accuracy?

The data produced from this study indicated that each group was statistically equivalent prior to the treatment sessions. Comparisons from pretest to posttest scores

indicated that both the aural and notation treatment groups produced significant improvement regarding the measurement of harmonic accuracy.

Comparisons of posttest scores indicated that the aural group produced significantly higher scores than the notation group. Because the two groups were equivalent prior to the experimental procedures, and because the students did not receive jazz improvisation instruction beyond the control of this study, it was concluded that the experimental procedures produced greater improvement for students who received the aural pedagogy.

Research Question 2

What were the comparative outcomes from the use of aural and notated methods of pedagogy by student grade level?

Results from the two-way ANOVA revealed grades 10, 11 and 12 as statistically equivalent, while comparisons between 9^{th} and 12^{th} graders were statistically different. A Mann-Whitney comparison indicate that although students from all grade levels from the aural group were equivalent, the notation group produced differences between 9^{th} and 12^{th} graders. The data from these comparisons indicate that for students who participated in the notation treatment method, 9^{th} graders received significantly lower posttest scores than 12^{th} graders.

Implications to the Literature

The conclusions reached from the results of this study contribute useful information to the research literature. All students indicated on the musical background questionnaire that they played from notated music rather than by ear. The conclusions from this study, however, indicated that although the predominately aural method of

pedagogy represented the least developed skill (in comparison to music notation skills), it produced greater improvement for the learning objective (harmonic accuracy). This researcher was surprised by these results, expecting the more developed skill for students (notation) to produce the greatest improvement.

Implications from these results to the pedagogy literature imply that the use of a predominately aural method of jazz pedagogy produced greater harmonic skills for high school beginning jazz improvisers. Few published pedagogical sources, according to Brown (1990), are primarily based on this type of pedagogy. Although Berg (1990), Coker (1989), and Lawn & Hellmer (1990) indicated the importance of aural development, these authors do not present a predominately aural pedagogical approach. The implications of this study to high school jazz programs that teach beginning jazz improvisers infer that published methods of pedagogy might not represent the most effective tradition of pedagogy.

While the historical and pedagogical literature provided evidence that oral and written traditions coexisted, many historically significant performers claim to have learned predominately by ear. Authors described the practice of band leaders who hired "session" players who could read notation while also hiring advanced soloists who sometimes could not. Because jazz has involved both oral and written traditions, methods of pedagogy have evolved from each system, although many significant soloists learned by ear. The results of this study verify these historical facts; a predominately aural method of pedagogy produced better jazz soloists.

Recommendations for Further Study

Following are recommendations for further study:

1) Continued investigation of the effects from the use of notated exercises and aural imitation as methods of jazz pedagogy is recommended. Because the present study primarily investigated these effects to the selected learning objective, harmonic accuracy, further studies should investigate these effects to other characteristics of jazz improvisation.

2) Because jazz history involves both oral and written traditions, the investigation of primarily aural or primarily notated methods for learning would have implications for the interpretation of this history. Further studies should investigate current issues which involve the oral and written traditions of jazz, and the predominate use of, and implications from, the selection of pedagogy.

3) Further investigation involving a comparison of student musical background to the selection of pedagogical method is recommended. The results of this study concluded that although all student participants indicated notated music skills as their primary musical background, the primarily aural method produced greater achievement for the lesson objective.

APPENDIX A

LESSON PLAN

Lesson Objective

All students will participate in the instructional curriculum and activities. The student will repeat a single lesson unit for a total of n = 4 sessions. The student will be instructed to practice the material and objectives for a minimum of one hour after the first, and prior to each of the remaining sessions. The student will be assessed both prior to, and after receiving the four lesson sessions. The student will improve upon the instructional objective.

Instructional Objective

The instructional objective was identified as the demonstration of "harmonic performance skills" through student jazz improvisation. The performance objective, "harmonic performance skills" will be measured both prior to and following the lesson presentation. The performance objective will be explained in terms of harmonic tonal areas, including major tonalities, scales, arpeggios, subsequent dorian modes and an alternative solution. Improvement from pre to post treatment assessment by the student of the performance objective will demonstrate successful instruction.

Instructional Curriculum

Materials provided to each student to practice prior to each session will be presented by the instructor. Appendix D contains the student notated materials for Group A (Notation); students will also receive a CD recording of the accompaniment and a notated transcription of a Miles Davis solo. Group B (aural) will receive a CD recording of the same exercises performed by the instructor along with the recorded accompaniment and a Miles Davis solo on the tune "So What" from Kind Of Blue, Columbia, 1959.

The following lesson exercises will be taught to each group (appendix H).

1) Concert “C” major scales, half notes, one octave.
2) Concert “C” major scale, eight notes, followed by triad arpeggio.
3) Concert “D” dorian scale, half notes, one octave.
4) Concert “D” dorian scale, eight notes, followed by triad arpeggio.
5) Concert “D” dorian mode, eighth notes, to the 9^{th}, arpeggio to the 9^{th}.
6) Concert “Db” major scales, half notes, one octave.
7) Concert “Db” major scale, eight notes, followed by triad arpeggio.
8) Concert “Eb” dorian scale, half notes, one octave.
9) Concert “Eb” dorian scale, eight notes, followed by triad arpeggio.
10) Concert “Eb” dorian mode, eighth notes, to the 9^{th}, arpeggio to the 9^{th}.
11) Melody to “So What,” to be practiced with the recorded accompaniment.
12) Single measure melodic patterns followed by a measure of rest, to be practiced with recorded accompaniment.
13) Transcription of Miles Davis solo for Group A (notation); recording of miles Davis solo for Group B (aural).
14) Alternative solution, minor pentatonic scales.

Lesson Outline

Following the collection of consent forms, questionnaire and pre treatment recording, the lesson plan will be delivered. Lesson materials will be the same for each group, however the method of delivery will change (notation or aural).

Unit One

The following instructions will be given (the instructor may provide examples, answer questions, etc.)

Stylistic improvisation consists of pre-defined parameters within an existing harmonic framework (as opposed to free improvisation which means you can play anything). While the stylistic patterns which define a style are important, the harmonic framework (playing the correct notes) will be the focus of this lesson plan.

Students will be asked to play a concert "C" major scale in half notes as a group. Group A will play from notation, Group B will repeat the scale after the instructor has demonstrated. The remaining scale and arpeggio material, as outlined in the instructional curriculum, will be delivered in this manner.

Unit Two

Students will be given a brief definition of song structure as follows.

From the time a song begins until it ends, certain aspects will repeat and at other times be different. As an example, the song "So What," written by Miles Davis and recorded on his album "Kind of Blue" (Columbia CL 1355) in 1959, will be explained in terms of its structure. In his introduction presented on the album cover, Bill Evans (1959), the pianist on this recording, describes the structure like this. "So What is a simple figure based on 16 measures of one scale, 8 of another and 8 more of the first…" That totals 32 measures which is one of the most common song structures. Four 8 measure phrases make up this form, the first 8 measures are repeated followed by a different 8 measure phrase called the bridge.

The fourth 8 measure phrase returns to the original 8 measures. This structure is sometimes referred to as AABA. As you improvise over this structure you must develop of sense (either by counting or feeling the phrases) of when the bridge arrives. At this point you must change scales or tonalities, then return to the first scale or tonality after 8 measures. The entire 32 measure structure repeats over and over again. Each one of you will be asked to improvise through one 32 bar structure, then the next person will be instructed to begin.

The Aebersold (1991) Volume 54, track 2 recording will provide accompaniment while the instructor demonstrates the melody to the song, "So What." Students will then play the melody with accompaniment, group A from the notated exercise sheet, group B by ear. The instructor will demonstrate the dorian scales and arpeggios in 8^{th} notes through the harmonic structure with accompaniment. This will be repeated by the students, group A with the use of the notated exercises while group B will perform the material by ear or from memory.

Unit Three

Single measure melodic patterns will be performed by the instructor and repeated on the following measure by the students. Group A will play from the notated exercise sheet while group B will repeat the patterns by ear.

Unit Four

The following quote regarding the Davis recording from Reeves (1995) will be read to students.

> It is interesting to contrast the approaches of the different soloists, Adderley and Coltrane take a harmonic approach, whereas Miles Davis and Bill Evans are lyrical and use more space. Miles' sense of time is very subtle, laying behind the beat at times and marking the center of the beat at other times. His melodies are singable and each note is significant, exemplifying his "less is more" approach. Most of the notes are derived from the dorian mode, but we also find usage of the harmonic minor scale in bars 12, 24, and 63. Miles anticipates the key changes in bars 24, 48, and 56. His nuances include varied articulations, ghosted notes (bars 17, 19, 20 and 51), and half-valve effects (bar 40). At times he explores the upper extensions of the chord, using the 7^{th}, 9^{th}, and 11^{th} of the minor chord; this creates the bitonal effect of a C major triad over a D minor 7^{th} chord in bars 34, 49, 58, and 59. The blues scale is hinted at in bar 41, where he uses the ♭5th. Overall, this solo is a good example of meaningful simplicity and lyricism. (p. 33)

The recorded Davis solo will be presented to students; group A will receive a notated transcription while group B will listen. Structural areas will be indicated by the instructor during the listening example or within the transcription.

Unit Five

One alternative solution will be supplied to students regarding harmonic strategy. As depicted in the Reeves analysis, Davis' use of the corresponding harmonic minor scale demonstrates an alternative harmonic solution. The concert "D" and "Eb" harmonic minor scales and arpeggios will be presented to students in the same fashion as previous scales (half notes, 8^{th} notes, to the 9^{th}).

Unit Six

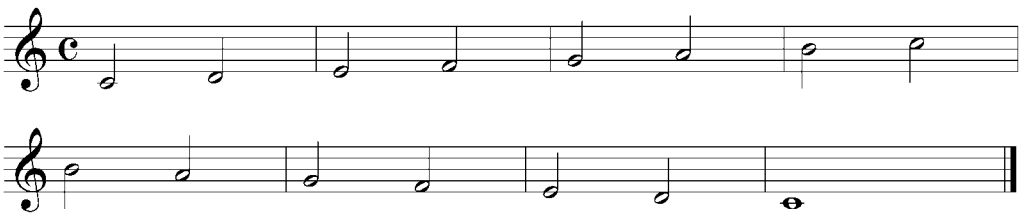
Each student will be instructed to improvise with the accompaniment; the instructor will point out harmonic areas during the first two sessions, but not during the third or fourth sessions.

APPENDIX B

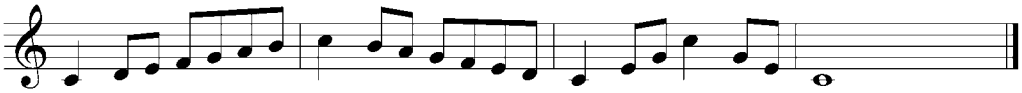
NOTATED EXERCISES

Concert Instruments
Unit I, page 1

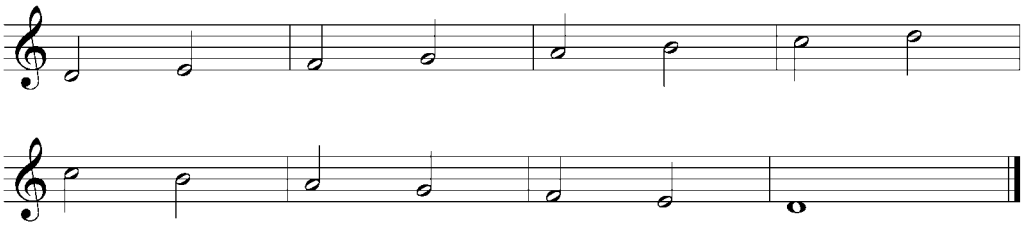
C Major Scale



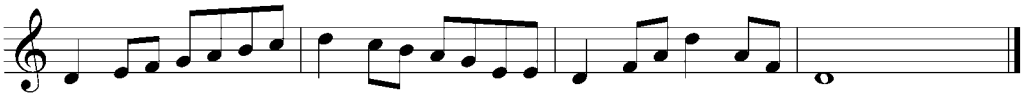
C Major Scale and Arpeggio



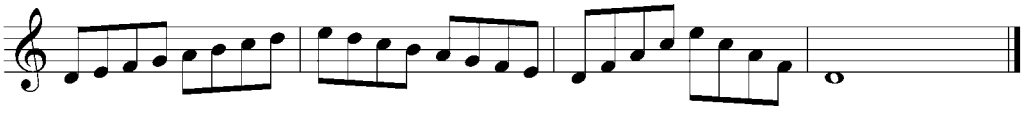
D Dorian Mode



D Dorian Mode and Arrpeggio

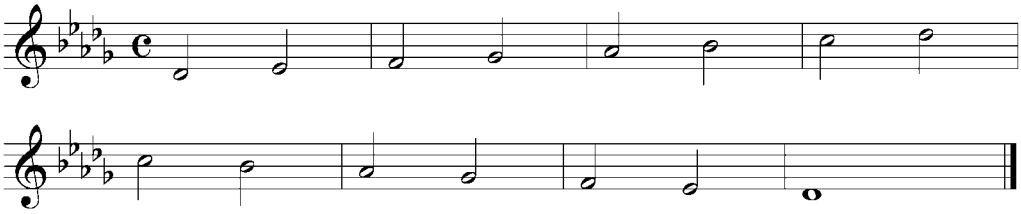


D Dorian Mode and Arrpeggio to the 9th

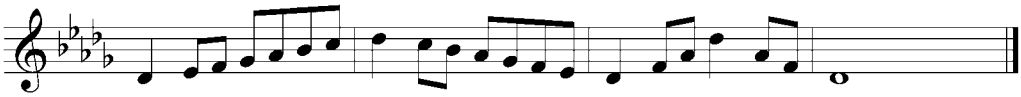


Concert Instruments
Unit I, page 2

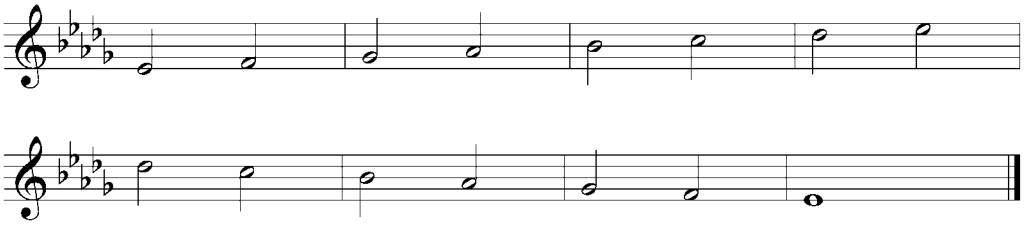
Db Major Scale



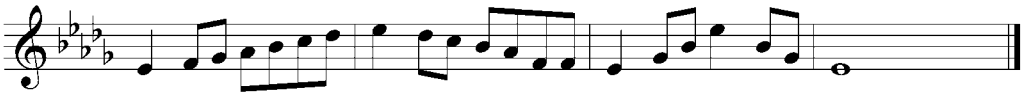
Db Major Scale and Arpeggio



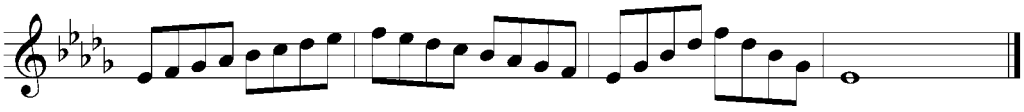
Eb Dorian Mode



Eb Dorian Mode and Arrpeggio

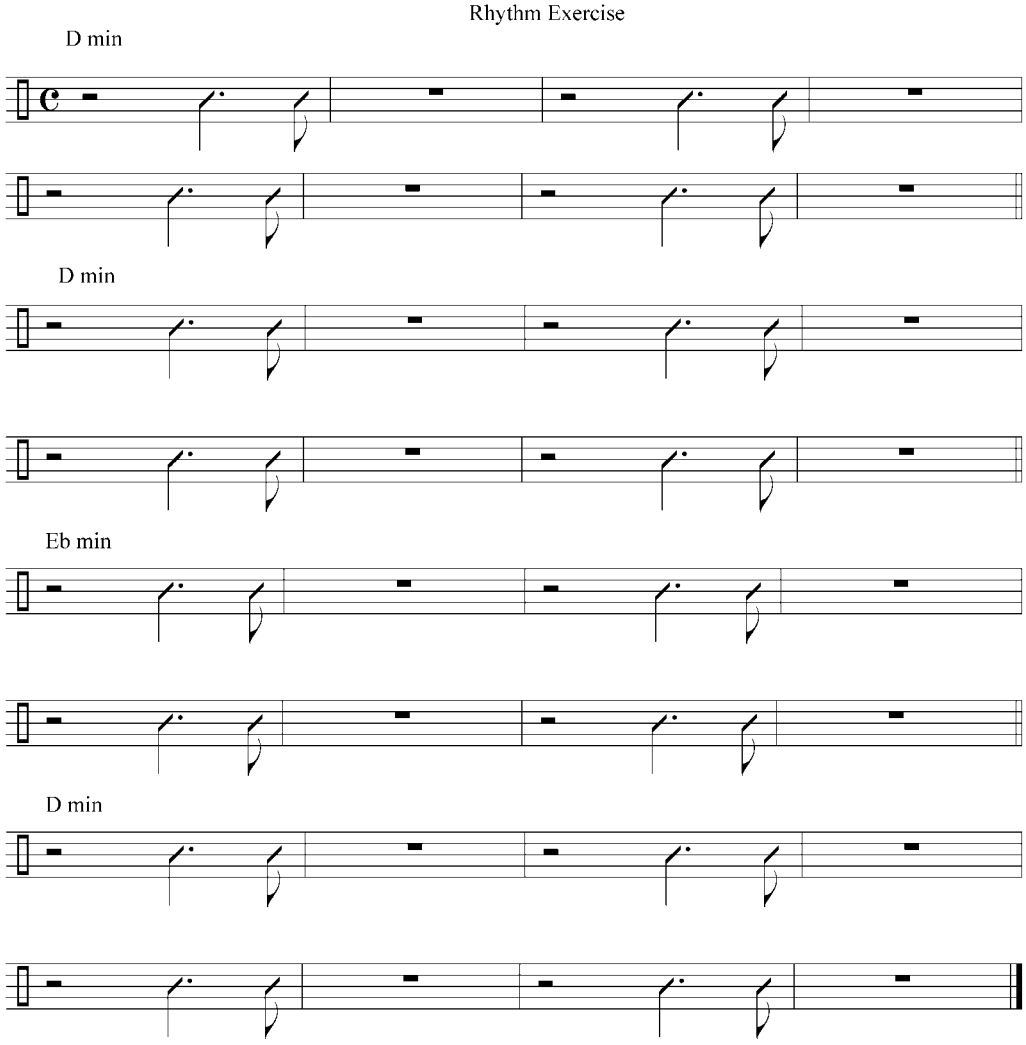


Eb Dorian Mode and Arrpeggio to the 9th

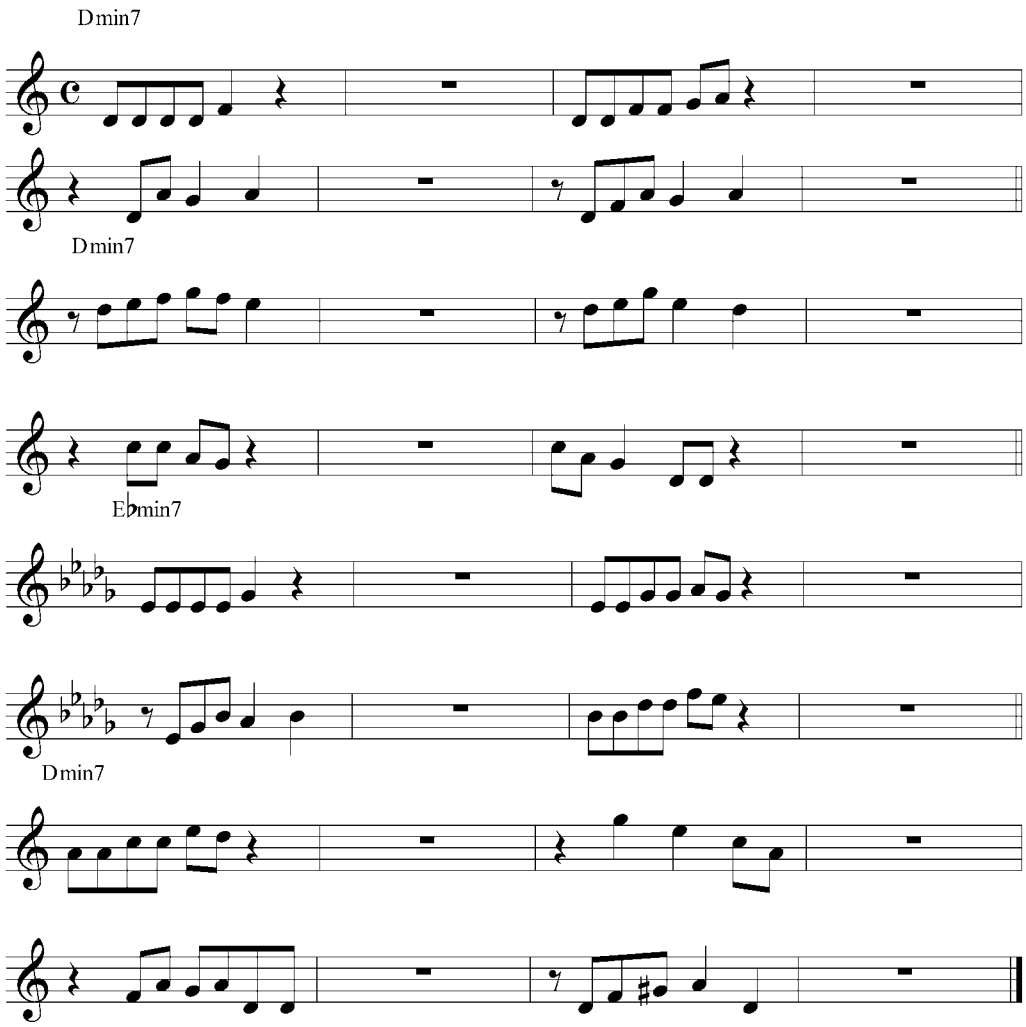


MILES

Rhythm Exercise

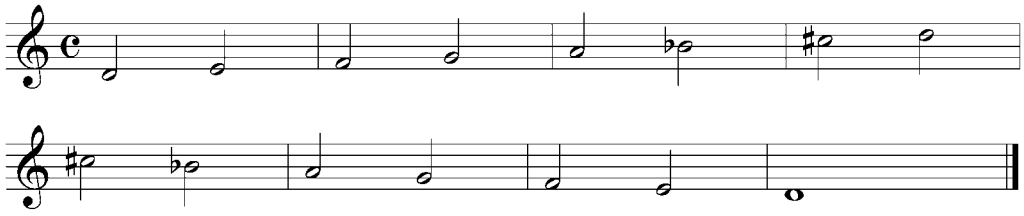


Unit III

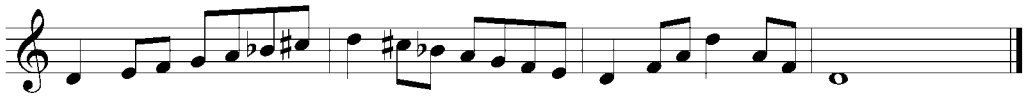


Concert Instruments
Unit V

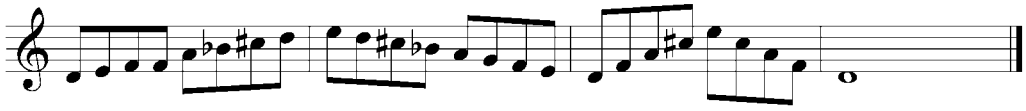
D Harmonic Minor Scale



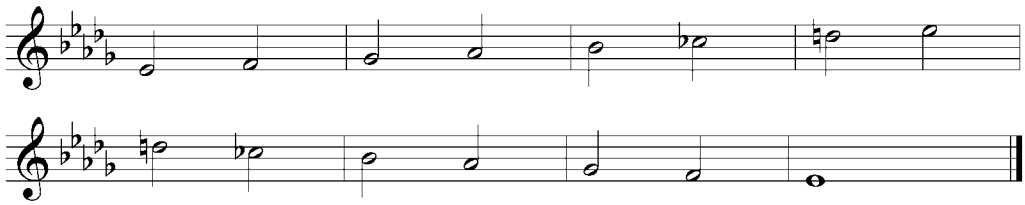
D Harmonic Minor Scale and Arpeggio



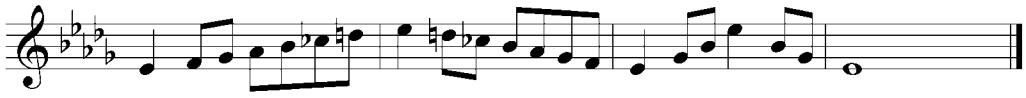
D Harmonic Minor Scale and Arpeggio to the 9th



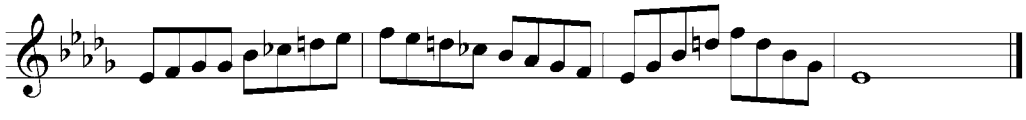
Eb Harmonic Minor Scale



Eb Harmonic Minor Scale and Arpeggio



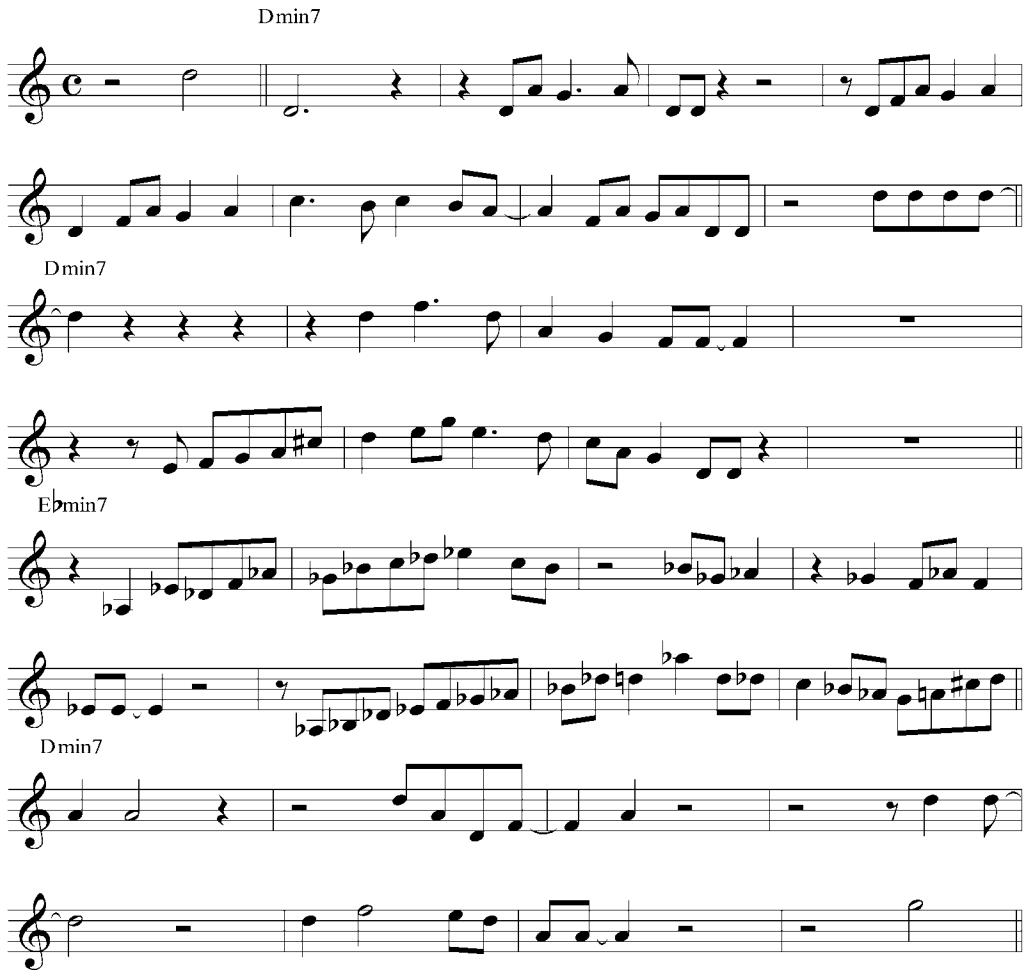
Eb Harmonic Minor Scale and Arpeggio to the 9th

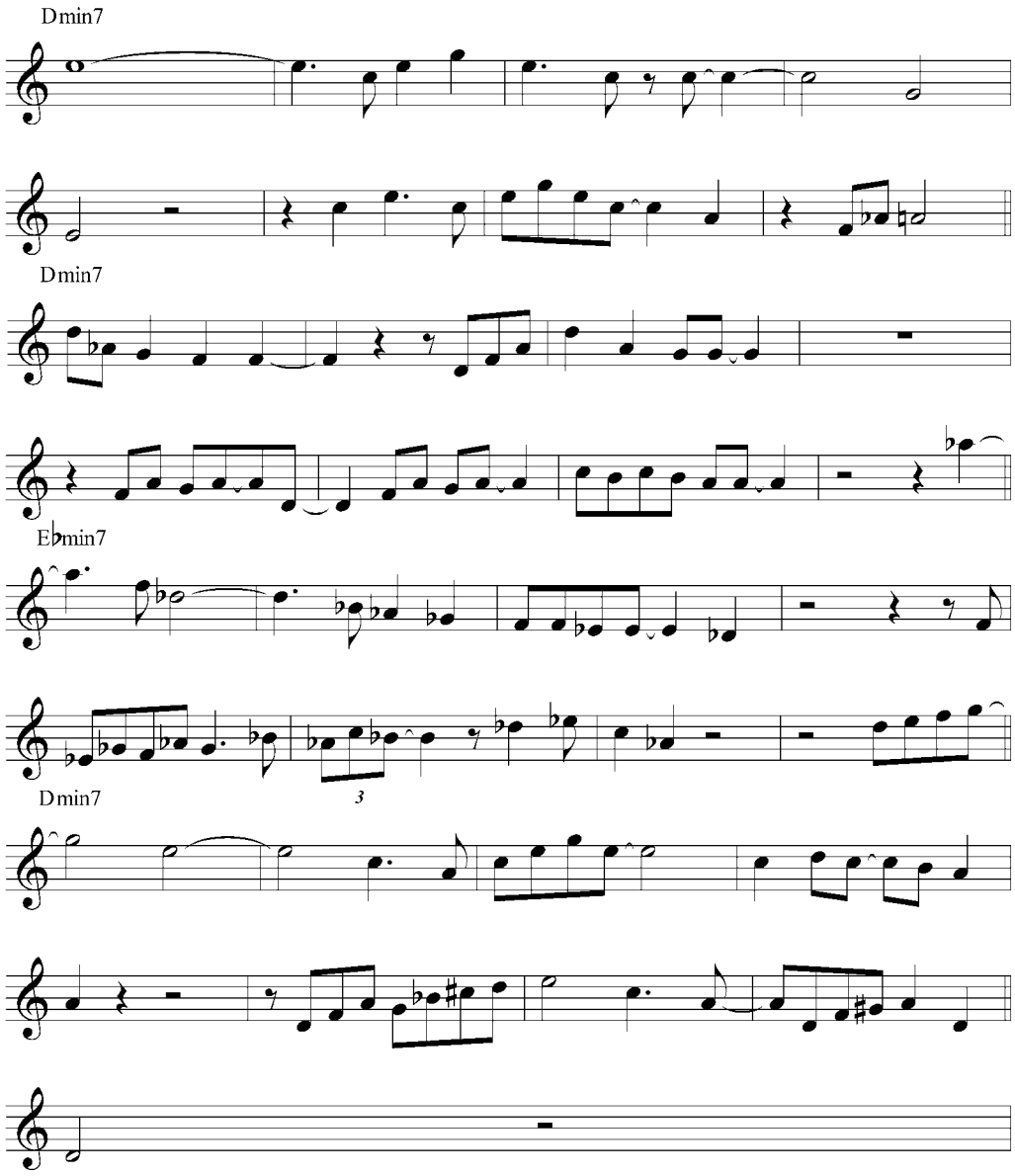


APPENDIX C

MILES DAVIS SOLO TRANSCRIPTION

Solo




Dmin7
Dmin7
Ebmin7
Dmin7
3

APPENDIX D

PILOT STUDY STUDENT LETTER

January 12, 1999

Thank you for agreeing to participate in this research study. You will be asked to read and sign a "Human Subjects Consent Form" which is required by The University of North Texas Graduate School. This letter will further explain what is required from your participation.

You have been asked to be present at two sessions this week, including today, Tuesday, and again on Thursday of this week. Because this research study involves teaching the same material with two different approaches you must only come at the time of your session. I must also ask that you not share your materials or discuss the first session with those involved in the other session until after Thursday. This will ensure that no one receives materials or guidance from both sessions. Because the purpose of this study is to investigate the two various teaching approaches, it is important that you not share your materials.

A recording session will occur during both sessions which will involve you performing a jazz improvised solo with a pre recorded accompaniment. The first session on Tuesday, prior to the lesson material, is simply to compare to the second session which will occur after the lesson material on Thursday. An adjudicator will rate various aspects of both sessions but will not know your identity. The results will be averaged with all participants and published as research findings. You name will not be included in this publication, assuring your anonymity.

The first session on Tuesday will involve a single jazz improvisation lesson plan. Materials will be provided for you to practice at home before the Thursday session. You are asked to spend at least one hour with the material on your own. You will be asked to record the amount of time spent with the take home material and provide that information on Thursday.

Sincerely,

Jim Laughlin
Associate Professor of Music
Amarillo College

APPENDIX E

STUDENT LETTER

August 22, 2000

Dear Area Band Student;

Thank you for your interest in a series of free jazz improvisation clinics. As I explained during my brief announcement to the band I am seeking volunteers to participate in a jazz improvisation class which will become a part of my dissertation study to complete my Ph.D. at the University of North Texas. I am seeking two separate groups of students (times and dates listed below) to meet a total of 4 times. You will be recorded before and after receiving jazz improvisation lesson materials and instruction. Any information collected will remain confidential and used only as a part of this dissertation study.

The times and dates of each group are as follows:

September 5	Group I	6:30 - 7:15 p.m.
	Group II	7:30 - 8:15 p.m.
September 12	Group I	6:30 - 7:15 p.m.
	Group II	7:30 - 8:15 p.m.
September 19	Group I	6:30 - 7:15 p.m.
	Group II	7:30 - 8:15 p.m.
September 26	Group I	6:30 - 7:15 p.m.
	Group II	7:30 - 8:15 p.m.

All sessions will meet in the Amarillo College Band Room [MB 104]. You must be willing to remain in one group and participate with each of the 4 sessions. You will not be charged for any materials or instruction related to this research study. Please call me for any questions prior to or during the course of the study.

Sincerely,

Jim Laughlin
Professor
Amarillo College
371-5345

APPENDIX F

HUMAN SUBJECTS CONSENT FORM

I, , agree to participate in an experimental research study involving a jazz improvisation curriculum for individual application. I understand that this study will involve comprehensive evaluation in the form of questionnaire, written exam and taped or live musical performance. The purpose of this study is to produce and evaluate a jazz improvisation curriculum for individual application which allows for unique experiences, prior knowledge and performance abilities. As a participant of the pilot study, I understand that data will be produced from my participation which will be recorded and manipulated in the effort to modify and improve this study.

As a participant, I understand that comprehensive evaluation will occur in multiple settings. Portions which I may complete at home include questionnaire, written testing, and musical ear training testing with accompanying cassette tape. Portions which I must complete with an evaluator include musical performance which may be taped or evaluated by a live adjudicator. I understand that the take home portion may take up to three hours to complete and the evaluator portion may take up to two hours to complete. In addition to the evaluation process, I understand my results will be made known to me and that recommendations will be made toward appropriate materials and pedagogy for improvement. One area of improvement will be selected and I will keep records of time which I practice the recommended assignments for the duration of two weeks, at which time I will participate in a post-test involving musical performance and questionnaire. The benefits of my participation in this study may result in improved jazz improvisation performance.

I understand that any information regarding my participation will remain confidential through the use of a code number assigned to my data. At the conclusion of this study the key which relates my name to the assigned code number will be destroyed. Under this condition, I agree that any information obtained from this research may be used in any way thought best for publication or education.

I understand that there is no personal risk or discomfort directly involved with this research and that I am free to withdraw my consent and discontinue participation in this study at any time. A decision to withdraw will not affect any services available to me.

If I have any questions or problems that arise in connection with my participation in this study, I should contact Dr. Jim Rauscher, Chairmen for the Music Department at Amarillo College.

(Date) (Signature of Participant)

(Date) (Investigator)

(Date) (Witness)*

* Witness signatures are required whenever the capacity of the subject to understand the description of the project and its associated risks is in question or when required by the IRB.
THIS PROJECT HAS BEEN REVIEWED BY UNIVERSITY OF NORTH TEXAS COMMITTEE FOR THE PROTECTION OF HUMAN SUBJECTS (Phone: 565-3940)

APPENDIX G

JAZZ IMPROVISATION QUESTIONNAIRE

Name:
Instrument:
School:
Grade:

How long have you played this instrument:

Do you mostly play by ear? ..... Yes No

Do you mostly play from music notation?..... Yes No

Have you ever played in jazz band or stage band?..... Yes No

Have you ever improvised on your instrument? ..... Yes No

Have you played major scales? ..... Yes No

Have you played dorian scales? ..... Yes No

Do you listen to jazz frequently (weekly)? ..... Yes No

Do you own any jazz recordings? ..... Yes No

Have you ever seen a professional jazz group? ..... Yes No

Rate your interest in learning jazz improvisation:
(highest) 5 4 3 2 1 (lowest)

Are you willing to practice between sessions? ..... Yes No

Have you ever made up songs on your instrument?..... Yes No

Have you ever played a song that you know without using music?..... Yes No

Have you ever played extra notes or "embellished" from your written music? Yes No

How would you rate your musical performance skills in general?
(highest) 5 4 3 2 1 (lowest)?

APPENDIX H

LETTER OF PERMISSION

JAMEY AEBERSOLD JAZZ
1211 Aebersold Dr.
New Albany, IN 4710

Jim Laughlin
1216 S. Bonham
Amarillo, TX 79102

Dear Jim,

We have no problem with your using our play-a-long track as outlined by the conditions in your letter of Jan.8, 2001. We do not represent ourselves as spokesmen for the copyright holder(s).

Laughlin's letter of Jan. 8, 2001 states:

The listed excerpt would be used to accompany volunteer student performances to be evaluated as a part of the research study. No more than one chorus (32 measures) will be used at any one time.

Six copies would be produced to a CD and sent to six evaluators to judge the student performances. All CD recordings would be returned and destroyed upon completion of the evaluation process.

Excerpts to be used are from Aebersold's (1991) A New Approach To Jazz Improvisation, Volume 54, track 2, "So What." The original CD will be played on a stereo system while high school student participants perform single chorus solos. A recording microphone will record both the student and the accompaniment. This recorded material would be reproduced to a CD and evaluated as a representation of pre and post treatment student abilities.

Good luck with your dissertation.

Sincerely,

Matt Eve, GM
Jamey Aebersold Jazz

APPENDIX I

EVALUATOR COVER LETTER

November 2, 2000

Dr. Terry Segress
Department of Music
SWOSU
100 Campus Dr.
Weatherford, OK 73096

Dear Dr. Segress:

Thank you for agreeing to evaluate my dissertation student data. Enclosed is a CD recording with 52 tracks representing individual student performances. Also enclosed is an evaluator instruction sheet which will explain the procedures. In general, the recordings represent high school students who have participated in concert band but NEVER in a jazz band. They represent beginners to jazz and jazz improvisation. They were taught a single lesson plan on more than one occasion; a recording was made of each student playing along with Jamey Aebersold's Vol 54, track 2 accompaniment to "Impressions" as pre and post treatment data. These were then randomized and recorded to the CD which is provided to you. Evaluate each track as a separate and equal performance.

Please return your evaluation sheets quickly in order for me to begin analysis procedures. Use the mailing labels provided, place the label with my name and address over the "send to" label that currently has your name and address. You may leave my name and address as the "from." Place the pre-paid postal label over the existing label and place in regular mail.

You may call me regarding any questions or comments at 806) 371-5345. Each track lasts about one minute and you do not have to complete all tracks in one sitting. Feel free to complete a few at one time during short breaks within your schedule. Please look over the evaluation sheet prior to the first track to make sure you understand the procedures. Thank you for providing this service to me with you busy schedule, I will contact you following receipt of these materials to verify your completion.

Sincerely,

Jim Laughlin
Professor
Amarillo College
P. O. Box 447
Amarillo, TX 79178

APPENDIX J

PERFORMANCE EVALUATION MEASURE

1	2	3	4	5
Strongly Disagree / Lowest Quality	Disagree	Neutral	Agree	Strongly Agree / Highest Quality

_____1. Harmonic Proficiency.

To what degree does the student demonstrate control of harmonic materials?

_____ 2. Demonstrates harmonic appropriate performance skills.

Does the soloist perform appropriate or correct notes for the harmonic environment? Indicate a "5" always, "4" mostly, "3" usually, "2" sometimes, "1" never.

_____ 3. Demonstrate control in selecting tones that correspond with the sounding chord.

You should simply document the manifestation of playing notes that work with the sounding chord. A "5" should be given when the player demonstrates command over the selection of notes. If they display mere capability, a "4" should be given. A "1" should be given when the player rarely plays notes that work with the chord.

_____ 4. Demonstrates an ability to solve problems and make adjustments.

Does the soloist show an ability to respond and recover from problems such as personal mistakes. The best players can make mistakes into brilliant moments by incorporating and developing the mistake ("5"). If players merely recover and makes an adjustment, they should receive a "4" response. If they constantly stumble with no recovery – "1." Because the progression involves two simple harmonic areas (modal), judge the players ability to change or adjust to the harmonic changes.

_____ 5. Exhibits harmonic / melodic inventiveness.

"5" = A high degree of using original material or original reworking of derived material.
"1" = Extremely low inventiveness.

_____ 6. Correct Notes

7. State further comments regarding any positive or negative features prominent to this student's performance.

APPENDIX K

LINEAR REGRESSION MODEL FOR ANCOVA ASSUMPTION

Q1 Aural	Coefficient	SE	p	95 % CI of Coefficient
Intercept	3.4063	0.6218	<0.0001	2.1167 to 4.6958
Slope	-0.1875	0.3869	0.6327	-0.9899 to 0.6149
Q1 Notation				
Intercept	2.7907	0.2722	<0.0001	2.2486 to 3.3329
Slope	-0.0805	0.1321	0.5441	-0.3437 to 0.1827

Q2 Aural	Coefficient	SE	p	95 % CI of Coefficient
Intercept	3.1979	0.6516	<0.0001	1.8466 to 4.5493
Slope	0.0625	0.4054	0.8789	-0.7783 to 0.9033
Q2 Notation				
Intercept	2.7166	0.2841	<0.0001	2.1504 to 3.2827
Slope	0.0292	0.1304	0.8237	-0.2307 to 0.2890

Q3 Aural	Coefficient	SE	p	95 % CI of Coefficient
Intercept	2.8000	0.5990	<0.0001	1.5578 to 4.0422
Slope	0.3000	0.3668	0.4222	-0.4607 to 1.0607
Q3 Notation				
Intercept	2.9337	0.2811	<0.0001	2.4236 to 3.5439
Slope	-0.0881	0.1285	0.4949	-0.3442 to 0.1679

Q4 Aural	Coefficient	SE	p	95 % CI of Coefficient
Intercept	3.0811	0.6609	<0.0001	1.6867 to 4.4755
Slope	0.1351	0.3468	0.7016	-0.5966 to 0.8669
Q4 Notation				
Intercept	2.5760	0.2843	<0.0001	2.0092 to 3.1428
Slope	0.0750	0.1259	0.5533	-0.1761 to 0.3261

Q5 Aural	Coefficient	SE	p	95 % CI of Coefficient
Intercept	3.3013	0.5422	<0.0001	2.1769 to 4.4257
Slope	-0.0830	0.2269	0.7181	-0.5536 to 0.3876
Q5 Notation				
Intercept	2.3328	0.2646	<0.0001	1.8057 to 2.8599
Slope	0.0646	0.1315	0.6247	-0.1973 to 0.3264

Q6 Aural	Coefficient	SE	p	95 % CI of Coefficient
Intercept	3.6270	0.5903	<0.0001	2.3957 to 4.8582
Slope	-0.3175	0.3461	0.3699	-1.0393 to 0.4044
Q6 Notation				
Intercept	3.0439	0.2765	<0.0001	2.4932 to 3.5946
Slope	-0.1804	0.1280	0.1628	-0.4353 to 0.0745

APPENDIX L

ANCOVA RESULTS OF POSTTEST SCORES BETWEEN METHODS

Q1 Post Between Methods

Source	SS	df	MS	F	Sig.
Corrected Model	16.441	2	8.221	8.420	.000
Intercept	106.670	1	106.670	109.261	.000
Q1	11.441	1	11.441	11.719	.001
Method	8.201	1	8.201	8.401	.004
Error	114.225	117	.976		
Total	1094.000	120			
Corrected Total	130.667	119			

Q2 Post Between Methods

Source	SS	df	MS	F	Sig.
Corrected Model	12.007	2	6.004	5.073	.008
Intercept	138.734	1	138.734	117.227	.000
Q6	7.552	1	7.552	6.381	.013
Method	6.673	1	6.673	5.639	.019
Error	134.916	114	1.183		
Total	1182.000	117			
Corrected Total	146.923	116			

Q3 Post Between Methods

Source	SS	df	MS	F	Sig.
Corrected Model	12.852	2	6.426	5.544	.005
Intercept	148.454	1	148.454	128.074	.000
Q3	10.245	1	10.245	8.839	.004
Method	4.604	1	4.604	3.972	.049
Error	132.140	114	1.159		
Total	1204.000	117			
Corrected Total	144.991	116			

Q4 Post Between Methods

Source	SS	df	MS	F	Sig.
Corrected Model	21.991	2	10.995	10.732	.000
Intercept	110.100	1	110.100	107.464	.000
Q4	16.594	1	16.594	16.197	.000
Method	6.574	1	6.574	6.417	.013
Error	106.552	104	1.025		
Total	1050.000	107			
Corrected Total	128.542	106			

Q5 Post Between Methods

Source	SS	df	MS	F	Sig.
Corrected Model	19.715	2	9.857	9.422	.000
Intercept	88.973	1	88.973	85.040	.000
Q5	16.902	1	16.902	16.155	.000
Method	2.763	1	2.763	2.641	.107
Error	122.410	117	1.046		
Total	969.000	120			
Corrected Total	142.125	119			

Q6 Post Between Methods

Source	SS	df	MS	F	Sig.
Corrected Model	8.346	2	4.173	3.809	.025
Intercept	134.289	1	134.289	122.572	.000
Q6	5.797	1	5.797	5.291	.023
Method	4.186	1	4.186	3.821	.053
Error	125.993	115	1.096		
Total	1114.000	118			
Corrected Total	134.339	117			

APPENDIX M
2 WAY ANOVA RESULTS, STUDENT GRADE LEVEL
BY TREATMENT METHOD

Q1 Student Grade Level by Treatment Method

Source	SS	df	MS	F	Sig.
Corrected Model	30.806	8	3.851	4.280	.000
Intercept	760.158	1	760.158	844.949	.000
Grade Level	14.699	4	3.675	4.085	.004
Treatment Method	7.719	1	7.719	8.580	.004
Grade Level by Treatment Method	2.176	3	.725	.806	.493
Error	99.861	111	.900		
Total	1094.000	120			
Corrected Total	130.667	119			

Q2 Student Grade Level by Treatment Method

Source	SS	df	MS	F	Sig.
Corrected Model	38.735	8	4.842	4.832	.000
Intercept	839.754	1	839.754	837.978	.000
Grade Level	19.606	4	4.901	4.891	.001
Treatment Method	7.671	1	7.671	7.654	.007
Grade Level by Treatment Method	3.293	3	1.098	1.095	.354
Error	109.231	109	1.002		
Total	1198.000	118			
Corrected Total	147.966	117			

Q3 Student Grade Level by Treatment Method

Source	SS	df	MS	F	Sig.
Corrected Model	42.427	8	5.303	5.636	.000
Intercept	853.592	1	853.592	907.152	.000
Grade Level	23.046	4	5.762	6.123	.000
Treatment Method	8.222	1	8.222	8.738	.004
Grade Level by Treatment Method	3.361	3	1.120	1.191	.317
Error	102.564	109	.941		
Total	1213.000	118			
Corrected Total	144.992	117			

Q4 Student Grade Level by Treatment Method

Source	SS	df	MS	F	Sig.
Corrected Model	34.615	8	4.327	4.341	.000
Intercept	699.839	1	699.839	702.199	.000
Grade Level	17.265	4	4.316	4.331	.003
Treatment Method	4.576	1	4.576	4.591	.034
Grade Level by Treatment Method	4.578	3	1.526	1.531	.211
Error	103.650	104	.997		
Total	1073.000	113			
Corrected Total	138.265	112			

Q5 Student Grade Level by Treatment Method

Source	SS	df	MS	F	Sig.
Corrected Model	33.139	8	4.142	4.219	.000
Intercept	639.212	1	639.212	651.024	.000
Grade Level	13.693	4	3.423	3.487	.010
Treatment Method	5.013	1	5.013	5.106	.026
Grade Level by Treatment Method	10.100	3	3.367	3.429	.020
Error	108.986	111	.982		
Total	969.000	120			
Corrected Total	142.125	119			

Q6 Student Grade Level by Treatment Method

Source	SS	df	MS	F	Sig.
Corrected Model	34.007	8	4.251	4.625	.000
Intercept	784.930	1	784.930	854.010	.000
Grade Level	18.974	4	4.744	5.161	.001
Treatment Method	7.484	1	7.484	8.143	.005
Grade Level by Treatment Method	2.221	3	.740	.805	.494
Error	101.102	110	.919		
Total	1118.000	119			
Corrected Total	135.109	118			

APPENDIX N
MANN-WHITNEY COMPARISONS OF 9TH TO 12TH GRADERS

Q1 Post Aural	n	Rank Sum	Mean Rank	U
9th Graders	12	99.0	8.25	51.0
12th Graders	6	72.0	12.00	21.0
2-tailed p = 0.1797				

Q1 Post Notation	n	Rank Sum	Mean Rank	U
9th Graders	12	118.0	9.83	248.0
12th Graders	24	548.0	22.83	40.0
2-tailed p = 0.0002				

Q2 Post Aural	n	Rank Sum	Mean Rank	U
9th Graders	12	96.0	8.00	54.0
12th Graders	6	75.0	12.50	18.0
2-tailed p = 0.1025				

Q2 Post Notation	n	Rank Sum	Mean Rank	U
9th Graders	12	116.0	9.67	250.0
12th Graders	24	550.0	22.92	38.0
2-tailed p = 0.0002				

Q3 Post Aural	n	Rank Sum	Mean Rank	U
9th Graders	12	93.0	7.75	57.0
12th Graders	6	78.0	13.00	15.0
2-tailed p = 0.0529				

Q3 Post Notation	n	Rank Sum	Mean Rank	U
9th Graders	12	103.0	8.58	263.0
12th Graders	24	563.0	23.46	25.0
2-tailed p = 0.0001				

Q4 Post Aural	n	Rank Sum	Mean Rank	U
9th Graders	9	58.0	6.44	41.0
12th Graders	6	62.0	10.33	13.0
2-tailed p = 0.1135				

Q4 Post Notation	n	Rank Sum	Mean Rank	U
9th Graders	12	129.0	10.75	237.0
12th Graders	24	537.0	22.38	51.0
2-tailed p = 0.0013				

Q5 Post Aural	n	Rank Sum	Mean Rank	U
9th Graders	12	103.0	8.58	47.0
12th Graders	6	68.0	11.33	25.0
2-tailed p = 0.3355				

Q5 Post Notation	n	Rank Sum	Mean Rank	U
9th Graders	12	121.5	10.13	244.5
12th Graders	24	544.5	22.69	43.5
2-tailed p = 0.0004				

Q6 Post Aural	n	Rank Sum	Mean Rank	U
9th Graders	12	96.5	8.04	53.5
12th Graders	6	74.5	12.42	18.5
2-tailed p = 0.1246				

Q6 Post Notation	n	Rank Sum	Mean Rank	U
9th Graders	12	114.5	9.54	251.5
12th Graders	24	551.5	22.98	36.5
2-tailed p = 0.0002				

REFERENCES

Aitken, A. E. (1975/1976). A Self - Instructional Audio - Imitation Method Designed to Teach Trumpet Students Jazz Improvisation in the Major Mode (Doctoral Dissertation, University of Oregon, 1975). Dissertation Abstracts International, 37/01, 18.

Baker, David (1989). A Comprehensive Method of Jazz Education for Teacher and Student. Van Nuys, CA: Alfred Publishing Co., Inc.

Berg, S. (1990). Jazz Improvisation: The Goal Note Method. Evergreen, Co: Lou Fischer Music Publishing Co.

Briscuso, J. (1972). A study of Ability in Spontaneous and Prepared Jazz Improvisation Among Students Who Possess Different Levels of Musical Aptitude (Doctoral Dissertation, The University of Iowa, 1972). Dissertation Abstracts International, 33/04, 1767.

Brown, T. W. (1990/1991). An Investigation of the Effectiveness of a Piano Course in Playing by Ear and Aural Skills Development for College Students (Doctoral Dissertation, University of Illinois at Urbana-Champaign, 1990). Dissertation Abstracts International, 51/12, 4052.

Burnsed, V. and Price, H. (1984). Improvisation & Evaluation. Jazz Research Papers, NAJE: Manhattan, KS, 1984.

Campbell, D. T. and Stanely, J. C. (1963). Experimental and Quasi-Experimental Designs for Research. Chicago: Rand McNally & Company.

Charters, S. B. & Kunstadt, L. (1962). Jazz; A History of the New York Scene. Garden City, NY: Doubleday.

Coker, J. (1990). The Teaching of Jazz. Rottenburg N., West Germany: Advance Music.

Conrad, D. (1997). Inter-rater Reliability: Assessing Musical Performance. Retrieved December 30, 2000, from the World Wide Web: http://members.accexxus.net/~dconrad/inter-rater-reliability.html

Dance, S. (1970). The World of Duke Ellington. New York: Charles Scribner's Sons.

Davis, M. (1959). So What. On Kind of Blue [CD] Columbia CK40579

Gridley, M. (1988). Jazz Styles: History and Analysis. 3rd edition. Englewood Cliffs, NJ.: Prentice Hall.

Hinkle, D. E., Wiersma, W., Jurs, S. G. (1988). Applied Statistics for the Behavioral Sciences. 2nd edition. Boston: Houghton Mifflin Company.

Lawn, R. J., & Hellmer, J. L. (1990). Jazz: Theory and Practice. Belmont, CA: Wadsworth Publishing Co.

Lyons, L. (1980). The 101 Best Jazz Albums: A History of Jazz On Records. New York: William Morrow and Co.

Mitchel, M. and Jolley, J. (1988). Research Design Explained. New York, NY: Holt, Rinehart and Winston, Inc.

Nettl, B. (1976). Folk Music In The United States: An Introduction (3rd ed., revised by Helen Myers). Detroit, MI: Wayne State University Press.

Pfenninger, R. C. (1990/1991). The Development and Validation of Three Rating Scales for the Objective Measurement of Jazz Improvisation Achievement (Doctoral Dissertation, Temple University, 1990). Dissertation Abstracts International, 51/01, 2674.

Schuller, G. (1968). Early Jazz: Its Roots and Musical Development. New York: Oxford University Press.

Segress, T. (1979/1980). The Development and Evaluation of a Comprehensive First Semester College Jazz Improvisation Curriculum (Doctoral Dissertation, North Texas State University, 1979). Dissertation Abstracts International, 40/12, 6065.

Southern, E. (1983). The Music of Black Americans: A History. New York: W.W. Norton & Company, Inc.

Tumlinson, C. D. (1991/1992). Theoretical Constructs of Jazz Improvisation Performance. (Doctoral Dissertation, University of North Texas, 1991). Dissertation Abstracts International, 52/12, 4258.

Wardlow, G. D. (1998). Chasin' That Devil Music: Searching for The Blues (Ed. Edward Komara). San Francisco, CA: Miller Freeman Books.

Wheaton, J. (1994). All That Jazz. New York: Ardsley House, Publishers, Inc.

Woideck, C. (1999). Charlie Parker: His Music and Life. Ann Arbor: The University of Michigan Press.

Zwick, R. A. (1987). Jazz Improvisation: A recommended Sequential Format of Instruction (Doctoral Dissertation, North Texas State University, 1987). Doctoral Dissertation Abstracts, 48/03, 5