INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

UMI®

THE EFFECTS OF SELF-EVALUATION, SELF-LISTENING, AND MODELING ON JUNIOR HIGH INSTRUMENTALISTS' MUSIC PERFORMANCE AND PRACTICE ATTITUDE

by

Michael Peter Hewitt

A Dissertation Submitted to the Faculty of the

SCHOOL OF MUSIC AND DANCE

In Partial Fulfillment of the Requirements
For the Degree of

DOCTOR OF PHILOSOPHY
WITH A MAJOR IN MUSIC

In the Graduate College

THE UNIVERSITY OF ARIZONA

2000

THE UNIVERSITY OF ARIZONA ®
GRADUATE COLLEGE

As members of the Final Examination Committee, we certify that we have read the dissertation prepared by Michael P. Hewitt

entitled The Effects of Self-Evaluation, Self-Listening, and Modeling on Junior High Instrumentalists' Music Performance and Practice Attitude

and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy

	6/30/00 Date
Dwayne E. Duss	6/30/00 Date
Donald G. [illegible]	6/30/00 Date
Stephen J. Paul	6/30/00 Date
	Date

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copy of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

Dissertation Director — 8/30/00 Date

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at The University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

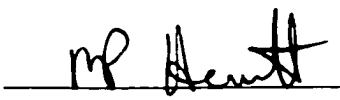
SIGNED: 

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Abstract

The purpose of this study was to examine the effects that self-evaluation, self-listening, and modeling have on junior high school instrumentalists' music performance and practice attitude. Eighty-two seventh, eighth, and ninth-grade students were randomly assigned to one of eight groups in the 2 x 2 x 2 pretest/posttest factorial design. Participants prepared three brief musical excerpts during 15-minute in-school weekly practice sessions and at home. Three adjudicators independently evaluated both pretest and posttest performances using the Woodwind Brass Solo Evaluation Form. Practice attitude was assessed by having students complete the Practice Attitude Questionnaire after each practice session. A general linear model repeated measures analysis with multiple dependent variables was performed to determine relationships among the variables. Results indicated that students who listened to a model recording improved their performance more so than students who did not listen to a model in the areas of tone, technique/articulation, rhythmic accuracy, tempo, interpretation, and overall performance, but not in the areas of intonation or melodic accuracy. A significant interaction between modeling, self-evaluation, and test was also discovered. Students in the Model/Self-Evaluation treatment group improved their performances more than the No Model/Self-Evaluation group in the areas of tone, melodic accuracy, rhythmic accuracy, interpretation, and overall performance. Also, there were no differences in scores between the Model/No Self-Evaluation and No Model/No Self-Evaluation for any performance subarea. In regard to practice attitude, no statistically significant interactions or main effects were found in the data. This indicates that mean practice

attitude scores remained constant throughout the duration of the study for all of the groups and also that the treatment groups did not differ from each other in terms of their attitude about the procedure at any time during the treatment period. Furthermore, each group exhibited a favorable attitude toward their particular practice strategy.

Chapter 1

Introduction

Introduction and Background

Musicians practice. Unquestionably, an essential part of being a quality musician in any genre is learning to practice effectively. While all musicians have developed their performance ability through some system of practice, there may be certain methods that are perhaps more efficient and effective. The familiar maxim, *Practice Makes Perfect*, may be true to some extent, but additional refinement may more honestly reflect the true purpose of practice. Thus, more precisely, *Efficient and Effective Practice Makes Perfect!*

As generally envisioned, music practice tends to evoke the image of a musician sitting alone in a room repeatedly performing warm-ups, exercises, scales, etudes, and musical works. For years, professional journals have reinforced this image through articles such as *Buzzing Practice for Beginners* (Reely, 1990), *Doing Double Duty on the Lip and Waistline* (Van Develde, 1981), *How's Your Technique* (O'Connell, 1967), and *Warm-up Procedures for the Cornetist* (Neilson, 1956). There can be no argument that essays like these are necessary and helpful to young musicians; however, they have also implicitly suggested the behaviors that constitute music practice.

In recent years, alternative definitions of practice have arisen which more fully bring together the diverse tasks essential for improving music performance. Ericsson, Krampe, and Tesch-Romer (1993) conceived the term "deliberate practice", which has also been adopted by Sloboda, Davidson, Howe, and Moore (1996) as "formal practice".

Deliberate formal practice involves wide ranging parameters that embody any and all activities that work to improve performance outside of "playing for fun". They have concluded that "level of expertise is a direct function of the amount of effortful formal practice of that skill undertaken by an individual" (Ericsson, 1996, p. 111). The authors state that to be most productive, music practice requires a well-defined task with an appropriate difficulty level for the musician, informative feedback, and opportunities for repetition and correction of errors (Ericsson, 1997).

In an even more encompassing definition Jørgensen (1995) views practicing as "self-teaching", which embodies a three-phase strategy: planning, the conduct of practice, and evaluation of practice. The planning stage incorporates strategies relating to the preparation of emotional, motivational, physical and musical elements. The conduct of practice stage deals with effective learning, monitoring and adjusting, and preparing for performance. The evaluating stage involves assessing products (performances), self-teaching, and the learning process. Furthermore, a meta-strategy classification is presented which formulates and controls the execution of the other three strategies.

Hallam (1997) offers a model that broadens the scope under which the notion of practice might be conceptualized. She states that effective music practice is "that which achieves the desired end-product, in as short a time as possible, without interfering negatively with the longer-term goals" (p. 181). Effective practice is "what works" in learning. In her model, derived from an example provided by Biggs and Moore (1993), the physical act of practice and repetition are consolidated with a number of other

elements she views as necessary for effective and efficient practice. This model suggests three stages of music practicing: presage, process, and product.

Presage factors include those aspects of the learning environment and characteristics of the learner that exist prior to practicing. The process stage involves what Hallam refers to as task-oriented and person-oriented strategies, the latter of which includes metacognition and thus self-evaluation. The concluding stage, product, deals with learning outcomes, which include both performance and affective results. Midway between the presage and process stages Hallam places an additional stage (unnamed) that includes the elements unique to the task being investigated, the repertoire and characteristics of the instrument being performed, and the specific performance requirements.

It is interesting that each of the recent concepts or models of practice presented incorporates some method of self-evaluation. Assuredly then, it can be seen that musicians' ability to effectively evaluate their own performances while utilizing various practice strategies is an essential skill to acquire if independent musicianship is desired. It would seem then that the attainment of independent musicianship would allow the student musician to become more involved in decision-making during ensemble rehearsals. Rehearsals in turn would then become more efficient as students are able to correct mistakes and make improvements individually rather than invariably relying on the leader of the ensemble to provide directives.

A fundamental element of self-evaluation is judging or comparing self-monitored information against a given standard, model, or goal (Davidson & Scripp, 1992; Slavin,

1991; Zimmerman, 1998). Thus in music performance, musicians must have a firm grasp of the aural concept they wish to evaluate prior to making a particular assessment. This concept can be generated either internally (via audiation) or externally from a live or recorded performance. External models appear to be superior to those that are internal as they tend to be more reliable and accurate (Bundy, 1987; Kepner, 1986). A recorded model is often preferable in a practice situation as it is more consistent and more readily accessible than a live performer.

Purpose and Research Questions

Given that self-evaluation is noted as an important element of practice by recently developed models and theories of practice and that a performance should be evaluated against a given external model to be effective, then the amalgamation of these methods should provide an optimal learning strategy for effective and efficient music practice. To date, there are no studies in the music education literature that examine the relationships of these strategies. Therefore, the purpose of this study was to examine the effects that modeling, listening to self on audiotape (self-listening), and self-evaluation have on junior high school instrumentalists' music performance and attitude about practice. Specifically, the following questions were addressed in the study:

Music Performance

1a. Is there a relationship between modeling, considered alone, and music performance?

1b. Is there a relationship between self-listening, considered alone, and music performance?

1c. Is there a relationship between self-evaluation, considered alone, and music performance?

1d. Is there a relationship between modeling and self-listening, considered *in combination*, and music performance, independent of the effects of modeling and self-listening alone?

1e. Is there a relationship between modeling and self-evaluation considered *in combination*, and music performance, independent of the effects of modeling and self-evaluation alone?

1f. Is there a relationship between self-listening and self-evaluation considered *in combination*, and music performance, independent of the effects of modeling and self-evaluation alone?

1g. Is there a relationship between modeling, self-listening, and self-evaluation considered *in combination*, and music performance, independent of the effects of modeling, self-listening, and self-evaluation alone?

Practice Attitude

2a. Is there a relationship between modeling, considered alone, and practice attitude?

2b. Is there a relationship between self-listening, considered alone, and practice attitude?

2c. Is there a relationship between self-evaluation, considered alone, and practice attitude?

2d. Is there a relationship between modeling and self-listening, considered *in combination*, and practice attitude, independent of the effects of modeling and self-listening alone?

2e. Is there a relationship between modeling and self-evaluation considered *in combination*, and practice attitude, independent of the effects of modeling and self-evaluation alone?

2f. Is there a relationship between self-listening and self-evaluation considered *in combination*, and practice attitude, independent of the effects of modeling and self-evaluation alone?

2g. Is there a relationship between modeling, self-listening, and self-evaluation considered *in combination*, and practice attitude, independent of the effects of modeling, self-listening, and self-evaluation alone?

Hypotheses

The following null hypotheses were investigated:

Music Performance

1a. There is no significant relationship between modeling, considered alone, and music performance.

1b. There is no significant relationship between self-listening, considered alone, and music performance.

1c. There is no significant relationship between self-evaluation, considered alone, and music performance.

1d. There is no significant relationship between modeling and self-listening, considered *in combination*, and music performance, independent of the effects of modeling and self-listening alone.

1e. There is no significant relationship between modeling and self-evaluation considered *in combination,* and music performance, independent of the effects of modeling and self-evaluation alone.

1f. There is no significant relationship between self-listening and self-evaluation considered *in combination,* and music performance, independent of the effects of modeling and self-evaluation alone.

1g. There is no significant relationship between modeling, self-listening, and self-evaluation considered *in combination,* and music performance, independent of the effects of modeling, self-listening, and self-evaluation alone.

Practice Attitude

2a. There is no significant relationship between modeling, considered alone, and practice attitude.

2b. There is no significant relationship between self-listening , considered alone, and practice attitude.

2c. There is no significant relationship between self-evaluation, considered alone, and practice attitude.

2d. There is no significant relationship between modeling and self-listening, considered *in combination,* and practice attitude, independent of the effects of modeling and self-listening alone.

2e. There is no significant relationship between modeling and self-evaluation considered *in combination,* and practice attitude, independent of the effects of modeling and self-evaluation alone.

2f. There is no significant relationship between self-listening and self-evaluation considered *in combination*, and practice attitude, independent of the effects of modeling and self-evaluation alone.

2g. There is no significant relationship between modeling, self-listening, and self-evaluation considered *in combination*, and practice attitude, independent of the effects of modeling, self-listening, and self-evaluation alone.

Definitions

Modeling

For the present study, modeling was defined as listening to an "ideal" audiotaped version of the music being performed for the purpose of emulating the exemplar performance. The recorded version was considered by the participants in the study to be in perfect condition, and one that they truly wished to simulate.

Self-Evaluation

The term self-evaluation referred to the weekly completion of an evaluation form by the practicing musician. As part of weekly treatment, selected groups of students were asked to evaluate the quality of their own performances. The areas of tone, intonation, melodic accuracy, rhythmic accuracy, technique/articulation, and tempo were appraised. The form was to serve as a diagnostic tool, providing students with information they could then use to identify performance areas in need of improvement.

Self-Listening

Self-listening referred to listening attentively to one's own performance of the music practiced via audiotape. Granted, students may indeed listen to themselves while

they are practicing, but studies (Bundy, 1987; Kepner, 1986) have shown that they may not be able to listen accurately to their performances, perhaps due to "sensory blocking". Thus, placing the performance on audiotape allows the musician to focus on the performance rather than performing.

Student Attitude

The definition for attitude used in the present study was posited by Cutietta (1992) as an adaptation and combination from Henerson et al. (1987) and Allport (1967). It states that "an attitude is a firmly held mental network of beliefs, feelings, and values that is organized through an individual's experience, and that exerts a directive and dynamic influence on the individual's perception and response to all objects and situations with which it is related" (Cutietta, 1992, p. 296). He states that "this definition acknowledges the interrelationship inherent in the multifaceted construct of attitude, and thus emphasizes similarities to all cognitive organization" (p. 296). Cutietta also states that attitude should be measured during or immediately following exposure to the stimuli being investigated. It should be assessed verbally with statements that show the interconnectedness of attitude components (feelings, beliefs, values) and methods of response (cognitive, behavioral, affective).

Limitations

In addition to the usual limitations of experimental studies, it must be mentioned that the process of self-evaluation used in this study is also quite limiting. In fact, it is particularly focused on the completion of one specific form once each week throughout the duration of the study. There was no attempt made to guide students through the

process or to have them formally select areas for improvement or strategies for bettering their performance. Also, students were not asked to provide information regarding how or why they selected areas for improvement. This particular self-evaluation process was chosen because it seemed highly practical in nature, and was a process students could and would be able to implement during their independent practice time. Two factors, the relatively short amount of time it took to complete the procedure, and the independence with which junior high students could implement it, were the most important factors in making the decision.

Overview of Remaining Chapters

This chapter served as an introduction to the current study, which was concerned with the effects that selected practice strategies of modeling, self-listening, and self-evaluation have on music performance and practice attitude. Chapter two is a review of related literature, specifically the areas of modeling, self-evaluation and self-listening. Chapter three explains the methodology used in the study as it specifically examines the design, procedures, and statistical procedures used, as well as presenting the instruments used for data collection. Results are shown in chapter four, while discussion, conclusions, and implications for music teaching and research are explored in the final chapter.

Chapter 2

Review of Related Literature

Introduction

The search for the literature reviewed for this study commenced with computer searches of general, psychological, educational, and music related databases. This revealed a variety of books, journal articles, and chapters in edited books, and doctoral dissertations that contained material relevant to the present study. The literature reviewed focused on the areas of modeling, self-evaluation, and self-listening. For each area, the general background of the topic is explored followed by a detailed review of each source that was closely associated to the present study. Each section ends with a summary of the literature examined in that section.

Modeling

Background.

Social learning theorists believe that most learning takes place through observing the behavior of other individuals, which is generally referred to as modeling (Bandura, 1977; Ormrod, 1999). Three types of models have been identified by Ormrod (1999): live models, symbolic models, and verbal instruction models. Live modeling occurs when an actual person demonstrates a particular behavior. Symbolic models are used when a person or character who is portrayed in a film, on television, in a book, or some other medium serves as the model. Verbal instructions are models that *describe* how to enact a certain behavior. No person, either live or symbolic, need be present in this situation.

Bandura (1977) lists four phases which must take place before an individual can effectively model a given behavior: attentional, retention, reproduction, and motivational. In the attentional phase, the learner needs to focus their attention on the behavior being modeled. Once the learner is attentive, the second phase is demonstrated when the learner recalls the behavior and is provided an opportunity to practice, or rehearse, the behavior. During the reproduction stage, the learner attempts to match the modeled behavior as closely as possible. It is important that the learner actually be physically capable of performing the behavior or this stage cannot be successfully actualized. Finally, students need to feel that what they are doing is important. They must *want* to demonstrate the behavior.

In general, modeling has been considered to be an effective tool for use in music education. Madsen, Greer, and Madsen (1975) have defined this term as the presentation, either live or recorded, of anything that may be later imitated by an observer. Tait (1992) presents three major areas of modeling that transpire in music teaching. The first, musical modeling, happens when the teacher provides a *total image* of the desired behavior either vocally or instrumentally. Aural modeling, is the use of phonetic vocalization in order to convey particular meanings or points of emphasis in the music. This vocalization can be demonstrated through humming, using syllables (du-ne-du-ne), or other vocal means. Physical modeling includes facial expressions, physical gestures, and conducting. Often, physical and aural modeling occur simultaneously. Tait states that although these three different types of modeling exist, they are all generally referred to in research literature with the generic term "modeling".

Related studies.

In an extensive review of the research on modeling in music education, Dickey (1992) examined studies a) of music teaching and conducting, b) using appropriate and inappropriate models, c) utilizing taped models. He derived the following conclusions:

1. Teacher demonstration-student imitation contributes significantly to the development of musical skills. Positive relationships exist between teacher modeling and student performance.

2. Students learn to make increasingly complex musical discriminations through modeling, via both musically appropriate and inappropriate demonstrations and imitations. The models used to teach music affect the way students think music should be performed. Students' performance preference, sense of correctness, group performance, and individual performance are all positively influenced by musical models.

3. Modeling is an effective strategy throughout a wide age distribution.

4. The use of a prepared tape as a model appears to be an effective teaching strategy for both elementary and college students.

5. Modeling is a more effective strategy than verbal description for teaching musical performance. Students prefer teaching that consists of more musical content and less verbal explanation.

6. Teachers who possess musical modeling skills, spend more time using them than teachers who do not have those skills.

7. To be effective, musical models must be accurate.

Rosenthal (1984) and Rosenthal, Wilson, Evans, and Greenwalt (1988) found that listening to an audiotaped model was an effective method among college music majors for learning a new piece of music. In the first study, Rosenthal (1984) compared the effectiveness of musical models and guided verbal models, alone and in combination with each other, and found that "direct modeling, without any added verbiage may be most effective in helping a student perform accurately" (p. 272). The group receiving only modeling treatment performed better than the guided model, guide only, and practice only groups for correct notes, rhythms, dynamics, and tempo, but not phrasing.

In the second study, Rosenthal et al. (1988) examined the effects of listening to an ideal recorded model with free practice, silent analysis, singing, and a control group on performance accuracy of notes, rhythm, phrasing or dynamics, articulation, and tempo. The model group received significantly higher scores than the singing group and the no practice group for rhythm and phrasing or dynamics. They received significantly lower scores than the singing group for tempo. Scores between modeling and all other groups for all additional performance areas were not different. From this study it appears that modeling alone is more effective than singing or not practicing for improving rhythm and phrasing, and equally as effective for improving other performance areas as practice and silent analysis. It must be noted that in each of the Rosenthal studies, the treatment was relatively short in duration, lasting approximately three minutes in the earlier study and an average of 15 minutes in the latter.

Anderson (1981) used a pretest-posttest design to examine the effects of using tape-recorded aural models for home practice on selected sight-reading and performance

skills. Subjects for the study were 80 sixth-grade woodwind students who were randomly placed in either an experimental or control group. All students were provided a set of exercises to practice both in class and at home during the 8-week treatment period. Students in the experimental group were additionally provided a cassette tape of solo clarinet performances of the musical exercises. Students were required to complete as many of the exercises as possible during the study. At the conclusion of the 8-weeks, students' tempo, pitch, rhythm, and articulation were tested on both a previously assigned piece (not the exercises) and in sight-reading a new piece. Results indicated that though both groups improved significantly from pretest to posttest, there were no significant differences between control and experimental groups with regard to any of the skills studied. This indicates that listening to a model did not affect students' performances any more than not listening to a model.

The comparison of verbal instruction and nonverbal teacher-student modeling was the focus of a study by Dickey (1991), which utilized a "pretest-posttest one-factor control-treatment group design with concurrent replication" (p. 133). Subjects for his study were 128 middle school students in four intact band classes. Students in the nonverbal group solved musical problems through a combination of teacher, student, or teacher/student modeling and discrimination of correct and incorrectly performed patterns. Students assigned to the verbal group received only verbal responses from the teachers and had questions asked of them that helped them to identify problems. Employing a number of different test instruments, Dickey concluded that instrumental ear-to-hand coordination and kinesthetic response skills could be effectively developed

and improved through the use of modeling strategies. Further, musical discriminations are not effectively taught through verbal descriptions.

The effects that audiotape and videotape models have on performance achievement were the focus of a study by Linklater (1997). Fifth- and sixth-grade (N = 146) beginning clarinetists from eight schools served as subjects in a one-factor, posttest-only design. Students were randomly assigned to one of three treatment groups. The first group received a videotape that included visual and aural models of clarinet performance with accompaniment. The second group received an audiotape that provided only aural clarinet models with accompaniment. The final group, which served as the control, received an audiotape with accompaniment only. Students were evaluated based on visual/physical criteria (embouchure, hand position, instrument position, and posture) and aural/musical criteria (tone quality, intonation, articulation, rhythmic accuracy, and melodic accuracy). Results indicated no significant differences between groups for most dependent measures. The only significant difference among groups occurred in the visual/physical category where the group receiving videotape treatment scored significantly higher than the accompaniment only control group; however, this difference disappeared in a subsequent follow-up test.

Zurcher (1975) studied the results of model-supportive practice on music performance. He randomly assigned 43 beginning brass students to either a control or experimental group for the first week of a 6-week study. Students then alternated between groups for the remainder of the experiment. The subjects, when part of the experimental group, received tapes with instructions, reminders, a pulse provided by a

metronome, and a model "play-along" performances of the lesson on the subject's own instrument. When students were part of the control group, they did not receive the tapes, but *only* practiced. Cumulative results indicated that model-supported practice is more effective than traditional physical practice for gross pitch discrimination, pitch matching, rhythmic discrimination, and time spent in practice, but not for tempo stability, fingering errors, or slide position errors.

Puopolo (1971) created and tested audiotaped self-instructional practice materials for beginning instrumentalists that consisted of 10 weekly instrumental music lessons on audiotape. Fifty-two fifth-grade beginning trumpet players were randomly assigned to either an experimental or control condition. The experimental group practiced 20-25 minutes each day during school hours using the tape. Members of the control group practiced the same amount of time, but without use of the tapes. The audiotapes included a model performance of all lesson material, a piano accompaniment for the performances, along with verbal instructions, explanations, and counted preparatory beats. The researcher found a statistically significant difference in performance achievement in favor of the group that received the audiotapes.

The relationship between the modeling skills of instrumental music teachers and their pupils' performance was investigated by Sang (1987) as an outgrowth of an earlier study (1985). Teachers ($N = 19$) were asked to perform a number of tasks that were recorded on audiotape. They were asked to a) perform musical examples 1-5 from the *Watkins-Farnum Performance Scale* (Watkins & Farnum, 1962) on instruments their students played, b) choose and perform a musical selection on their own instrument

which best demonstrated their current performance ability, and c) perform the *Test of Ear-to-Hand Coordination* (Froseth, 1982). These three tasks were each recorded and evaluated by professional musicians. Additionally, the teachers were asked to complete a visual analysis of musically related performance skills for both major and secondary instruments. The researcher evaluated this task. The teachers randomly selected 10-12 of their students to perform exercises 1-5 of Form B of the *Watkins-Farnum Performance Scale* (Watkins & Farnum, 1962) that was evaluated by a panel of judges. The students' mean score was then calculated for each teacher. The teachers' classrooms were monitored for the amount of time they spent in modeling behaviors. Multiple regression analysis determined that teacher modeling could explain the variance in student performance behaviors. Furthermore, a significant relationship was found between teachers' modeling ability and the amount of time they spent modeling in the classroom.

In a relevant study of vocal modeling, Small & McCachern (1983) sought to examine the effects that female and male vocal models, without feedback, had on first-grade students' ($N = 55$) pitch matching accuracy. In a pretest-posttest design, individual students were initially asked to match pitches with a puppet whose voice was heard as either male or female. Students were then randomly assigned to one of three conditions, a) female vocal model, b) male vocal model, or c) no contact control group. After a number of sessions, students were again tested and recorded on audiotape. Measures were taken to provide no feedback in regards to children's' singing during the treatment. Results indicated that students receiving the female model treatment scored significantly higher than those receiving the male model treatment. More important to the current

study, it was found that there was no significant difference between both vocal model groups and the control group. This indicates that modeling alone, apart from feedback, produced no significant differences in pitch-matching accuracy.

Another study examining the effects of vocal modeling on pitch-matching accuracy of elementary students was undertaken by Green (1990). Under examination in her study was the effect that female, male, and child vocal modeling had on first- through sixth-grade children's ability to match pitch. Students ($N = 282$) enrolled in a university laboratory school were individually tested on three separate occasions (at one-week intervals), each time responding to a different model. Students heard an audiotaped two-beat *sol-mi* pattern sung with each note receiving one beat. Students were then given one opportunity to imitate the stimuli by singing into a microphone that recorded their response. These responses were then evaluated by three trained experts. Results indicated that the most correct responses were elicited from the child model, with the male model having the fewest number of correct responses. Green suggests that the reason for this may be that "a child can most easily match a pitch presented by a voice similar in timbre to his or her own voice" (p. 229).

Summary.

From the literature sited above, it appears that the effectiveness of using models to improve musicians' performance has been mixed, though mostly positive. Some studies have shown that listening to a model may be more effective for music performance than other methods of practice. Models (aural) appear to be more effective than verbal models for improving notes, rhythms, dynamics, and tempo, but not phrasing (Rosenthal, 1984)

and better than singing or not practicing for rhythm and phrasing/dynamics (Rosenthal et al., 1988). When compared to physical practice, listening to a model appears to be as effective or more effective depending upon the amount of "control" assumed in the study. Whereas superior effects of modeling (compared to practice) on performance achievement were found in some studies (Dickey, 1991; Puopolo, 1971; Zurcher 1975), others found modeling to be only "as good as" physical practice (Anderson, 1981; Linklater, 1997; Rosenthal et al., 1988). Studies which found a superior effect for modeling provided treatment to subjects, at least in part, during the regular school day, while those that did not, either provided "take home" treatment (Anderson, 1981; Linklater, 1997) or were relatively short in duration (Rosenthal et al., 1988).

Further, research on effects of music teacher modeling have shown that those who exhibit superior modeling skills tend to have students who are better performers, and these teachers tend to model more often (Sang, 1985, 1987). Research in the area of pitch-matching helps to shed some light on the effects of modeling. It appears that modeling, when separate from feedback, does not assist in performance accuracy (Small & McCachern, 1983). When aural models are used it is important that the timbre be similar between the model and the instrument used to match the model (Green, 1990).

Self-Evaluation

Background.

Much of the time spent by students learning to play a musical instrument takes place in isolation from other musicians, in a "practice room". During these isolated practice sessions, musicians generally rehearse material that has been assigned to them

during a lesson or rehearsal for them to "learn". At later sessions, their learning is evaluated (either formally or informally) by a teacher or ensemble director in regard to its level of performance. During the time meant for learning assigned music (practice sessions) students must regularly determine the strengths and weaknesses of their performances to determine what improvements are needed for the next session. Thus, they must be able to self-evaluate accurately.

Boyle and Radocy (1987) state that evaluation involves "making some judgment or decision regarding the worth, quality, or value of experiences, procedures, activities, or individual or group performances as they relate to some educational endeavor" (p.7). A search of the literature associated with self-evaluation in music education revealed a relatively limited number of studies. The effects of self-evaluation have been investigated in the areas of conductor training (Byo, 1990; Johnston, 1993; Yarbrough, 1987; Yarbrough, Wapnick, & Kelly, 1979), teaching behaviors and effectiveness (Colwell, 1995; Rosenthal, 1985), and instrumental music performance (Aitchison, 1995; Bergee, 1993; Byo & Brooks, 1994; Davis, 1981; Kostka, 1997; Sparks, 1990). Generally, self-assessment does not correlate well to assessments made by teachers or one's peers (Bergee, 1993; Byo, 1990; Kostka, 1997); however, it is necessary if we desire to instill independence in our music students. Bergee (1993) and Kostka (1997) contend that careful structure is needed to make this happen.

Related studies.

Bergee (1993) compared the evaluations of applied brass juries among faculty, peers, and the performers themselves. At the end of two separate academic semesters

brass faculty members evaluated 10 randomly selected brass performances using the Brass Performance Rating Scale (Bergee, 1988, 1989). Immediately following each set of juries, the student performers were asked to evaluate videotapes of the 10 performances that took place during the jury at which they performed, including their own. Data was analyzed for interjudge reliability by comparing mean difference scores. Results indicated that students were able to successfully evaluate their peers' performances as they related to evaluations by the faculty; however, students were unable to precisely evaluate their own performance, as their self-evaluations did not correlate well with that of the faculty. Bergee found no consistent pattern regarding the direction of the inconsistent self-evaluations, as some student self-scores were higher than those of faculty while others were lower.

In a study of undergraduate piano students, Kostka (1997) examined student self-evaluations as they compared to the evaluation of the teacher. Students evaluated their performances at the conclusion of each of five two-week units. The focus of each unit changed every two weeks and included the skills of technique, correct fingering, hand position, musicianship, and sight-reading. The teacher also assessed the students after each unit using Likert-type measurement instruments. Results indicated that students' evaluations did not correlate well with the teacher's evaluations, and that student self-evaluation scores across all five skills were consistently lower than the teacher scores. Over the course of treatment, students' changed the degree to which they "valued" each of the skills. Students valued more highly only the skill of correct fingering, whereas the other four skill areas were all rated lower at the posttest than the pretest.

Comparing junior high school instrumentalists' (n = 48) global performance evaluations to that of music educators (n = 5) was the purpose of a study by Byo and Brooks (1994). Two performance evaluations were undertaken using the Continuous Response Digital Interface (CRDI). The first performance evaluated was a university-level ensemble and the second was a recording of a junior high band whose members were participants in the study. While listening to each performance, subjects manipulated the CRDI providing an overall rating of the piece. The 255 data-point scale provided on the CRDI was divided by the authors into 5 categories. These categories constituted a) bad playing, b) not very good playing, c) just okay playing, d) good playing, and e) very good playing. Results of the study indicated that evaluations of the university-level group were quite similar between the student and educator groups, though not statistically. During evaluation of the junior high ensemble, students rated themselves higher than did music educators. Despite these inflated scores on the part of the students, the authors state that students often concurred with the educators in regard to the high and low points of the performance. In regard to evaluating performances, Byo and Brooks (1994) state that whereas "self-evaluators may not be as objective as independent evaluators...they are capable of being objective to a degree about their performance" (p. 35).

Davis (1981) examined the effects that self-evaluation, coupled with structured singing activities, had on 59 fifth- and 34 sixth-grade wind players' music performance, melodic tonal imagery, attitude, and self-evaluation accuracy. Within each grade, students were placed into one of six treatment groups that met for 40 minutes twice each

week. Three of the groups in each grade received "traditional instruction" and acted as control groups, a fourth group received structured singing exercises during class periods while another incorporated self-evaluation into their lessons. The final group received both structured singing and self-evaluation treatment. Results suggested that self-evaluation practice alone, and when combined with structured singing activities, is effective for developing performance skills during the second year of instruction, yet not the first. All treatment groups helped aid students in the development of self-evaluation skills, but so did the control groups, indicating that students are able to effectively self-evaluate their performances during early stages of learning a musical instrument. Instruction in self-evaluation and self-evaluation/singing activities in the second year of study provided an effective approach to the development of positive attitudes toward music in general.

Sparks (1990) investigated the effects of self-evaluation by beginning band students on their musical achievement, attentiveness in the classroom, and attitudes toward themselves, their school, band, instrument, practice, and their band director. Subjects ($N = 57$) were pretested for music performance (group and individual) using musical excerpts from class as well as attitude prior to treatment. The pretest revealed that the control group (woodwinds) received significantly higher performance scores than the treatment group (brass). Students in the treatment group received instruction in the use of self-evaluation forms, which they used during each class period. The control group received traditional teacher lead evaluation and instruction. All classes were videotaped daily. At the end of two and four weeks of instruction, tests were

administered to examine attitude, individual performance, and group performance. Analysis by a panel of experts revealed that the brass players had a higher percentage of on-task behavior for every category. Data pertaining to students' attitudes revealed that students in the experimental group significantly ($p < .05$) improved from pre- to posttest, while control group scores exhibited little change. Specifically, experimental scores improved in attitude toward band class, frequency of practice at home, and attitude toward the students' band director. Also, the control group's advantage in music performance revealed during the pretest was no longer evident by the posttest as both groups performed similarly. Attitudinal data revealed that student attitudes toward their instrument in the no self-evaluation group decreased. Data from the self-evaluation group revealed statistically significant improvement for attitude toward band class, frequency of home practice, and attitude toward the band director. Sparks concluded that students instructed in self-evaluation made greater progress in performance than students who received no self-evaluation instruction.

The purpose of an investigation by Aitchinson (1995) was to "examine the effects of incrementally increased levels of student self-evaluation involvement, in conjunction with development and time...on performance, self-evaluation accuracy, motivation, and self-esteem" (p. 38). Subjects, 84 seventh- and eighth-grade instrumental music students, were assigned to one of four experimental groups consisting of varying levels of student involvement in the evaluation process (student-only, student-driven, teacher-driven, teacher-only). Students prepared music over an eight-week period and were subjected to pre-, mid-, and posttests of the dependent variables. Results of the study indicated that

subjects in the student-driven mode scored significantly lower than those in the teacher-only mode in most areas including music performance. Performance scores of students with high developmental ability were higher than those of low ability. Over the course of the study, students increased their self-evaluation accuracy, intrinsic interest in music, self-esteem toward their own musical ability (they thought they were better musicians), and critical commentary. Performance scores for student-only evaluation were lower on the posttest than teacher-driven and teacher-only modes for teacher selected music. Finally, students' preference for evaluation was not significantly related to the evaluation mode.

Summary.

Research on the effects of self-evaluation on music performance, perhaps because of limited study, is inconclusive. University students were unable to successfully evaluate their music performances as they related to expert evaluation (Bergee, 1993; Kostka, 1997). Elementary students, upon receiving instruction in self-evaluation, improved their performance ability (Davis, 1981; Sparks, 1990). Middle school students who participated more predominantly in their evaluation saw somewhat mixed results, but generally it is noted that their self-evaluation accuracy increased while their music performance ability did not increase as much as students who were provided teacher feedback (Aitchison, 1995).

Self-evaluation may affect students' attitudes. Piano students lowered the degree to which they valued four of five performance areas after instruction in self-evaluation (Kostka, 1997). Following self-evaluation, attitudes among beginning band students

were more positive toward music in general (Davis, 1981), their own band class, and their attitude toward the band director (Sparks, 1990). Frequency of home practice also increased (Sparks, 1990). Aitchison (1995) found positive influences for intrinsic interest in music and the perception of music performance ability. As a whole, these studies suggest that self-evaluation can affect certain student attitudes.

Self-Listening

No studies to date have investigated the relationship that self-listening has to music performance, but two studies have examined the accuracy of evaluating music performance during live and recorded performances. One purpose for a study by Kepner (1986) was to determine the number of errors detected by high school instrumental students ($N = 50$) while performing live and again when listening to audiotaped versions of the same performance. During the first week of the 4-week study, students received instruction in performing the music. Students also practiced evaluating their performances in regard to pitch accuracy, rhythmic accuracy, articulation, tempo, dynamics, and intonation. The second week involved individual live testing on the prepared material. During week three, students sight-read and evaluated their performances of new material, and in the fourth week, students listened to and evaluated audiotaped versions of their live performances. Students were able to identify more errors while listening to the taped versions of their performances as opposed to the live condition. It is uncertain what types of errors the students detected as the evaluation system used asked that students mark each measure in which they believe *any* type of error was made. Kepner explains the result that more errors were detected while listening

to audiotape with "sensory blocking". He states that high school instrumentalists may be so busy monitoring the many sensory aspects involved in playing a musical instrument that they "block out" what they believe to be less important stimuli including their own performance errors.

In a study similar to that of Kepner, Bundy (1987) sought to determine whether junior high instrumentalists ($N = 29$) could detect their own pitch and rhythm errors more accurately under live conditions or on audiotape. Initially, students performed three etudes after which they immediately noted the measure(s) in which they had made errors. During a second session, students were asked to listen to and evaluate recordings of three etudes, one of which was their own. Results showed that students were more accurate in the performance (live) mode than in the listening mode for detecting pitch errors. Also, within the listening mode, rhythmic errors were identified correctly more often than pitch errors. Students detected significantly fewer errors than existed for both rhythm and pitch in the listening mode, while in the performance mode the difference existed only for rhythm errors. These findings indicate that students were more adept at identifying pitch errors during performance, while students detected rhythm errors equally, though poorly, under both conditions.

Summary.

The two studies examining self-listening reveal no conclusive indications as to whether the process of listening to one's performance on audiotape is a worthwhile task. Kepner indicates that students were better able to detect errors they made while using audiotape than when listening live, but he does not indicate what types of errors (pitch,

rhythm, etc.) these represented. Bundy found that students were able to detect pitch errors more accurately while performing live. Perhaps these findings indicate that students can identify pitch errors more accurately during live performance, but are more accurate at identifying other error types while utilizing audiotape.

Chapter 3

Methodology

Restatement of Purpose

The first two chapters presented an overview of research concerning practice strategies of modeling, self-listening, and self-evaluation. Because there has been little consideration given to the comparison of these various practice strategies, the undertaking of a study examining these areas as they relate to music performance improvement seems warranted. The purpose of this study was to examine the effects that modeling, self-listening, and self-evaluation have on middle school instrumentalists' music performance and practice attitude.

Sample

Participants ($N = 82$) for this study were seventh- ($n = 36$) eighth- ($n = 32$), and ninth-grade ($n = 15$) woodwind, brass, and percussion students from a junior high school in a southwestern state. The school is located in the outlying edge of a large city, and is attended by students whose families maintain a middle to upper-middle class socioeconomic status. Students at the school begin instruction on a wind or percussion instrument in the sixth-grade at one of two distinct elementary schools where both the junior high teacher and an additional part-time teacher provide training in a co-teaching environment. After a year of instruction all students converge into a single band at the junior high school where this study took place. When students complete their seventh-grade year they are placed by audition into either an intermediate or advanced band, consisting of both eighth- and ninth-grade students.

Students ranged in age from 12 to 15 years with a mean of 13.09 (SD = .97) years. The number of females slightly outnumbered males 43 to 40. The students averaged 2.79 (SD = 1.15) years of prior experience performing on their current instrument and had been involved in performing some type of music instrument for 4.70 (SD =2.34) years.

All students (N = 137) enrolled in one of three separate band classes at the school were offered the opportunity to participate in the project. A letter of explanation was sent to parents along with a parental consent form prior to the commencement of the treatment (see Appendix A). Students were allowed to return the consent forms to the researcher or instrumental teacher through the end of the performance pretest period. Students who returned the document by the specified date were then accepted for participation in the study.

Design

A 2 x 2 x 2 factorial design with a repeated measures component was employed for this multivariate study. The dependent variables of music performance (with seven subareas) and practice attitude were addressed in relation to the independent variables of modeling (two levels), self-listening (two levels), and self-evaluation (two levels).

A stratified random sample was used to assign students to treatment groups. Within each grade level, students were randomly assigned to one of eight treatment groups that encompassed two modeling conditions, two self-listening conditions, and two self-evaluation conditions. Treatment groups thus became (A) Model x Self-Listening x Self-Evaluation (n = 11), (B) Model x Self-Listening x No Self-Evaluation (n = 11), (C) Model x No Self-Listening x Self-Evaluation (n = 10), (D) Model x No Self-Listening x

No Self-Evaluation (n = 10), (E) No Model x Self-Listening x Self-Evaluation (n = 10), (F) No Model x Self-Listening x No Self-Evaluation (n = 10), (G) No Model x No Self-Listening x Self-Evaluation (n = 10), and (H) No Model x No Self-Listening x No Self-Evaluation (n = 10). This design is portrayed in Figure 1.

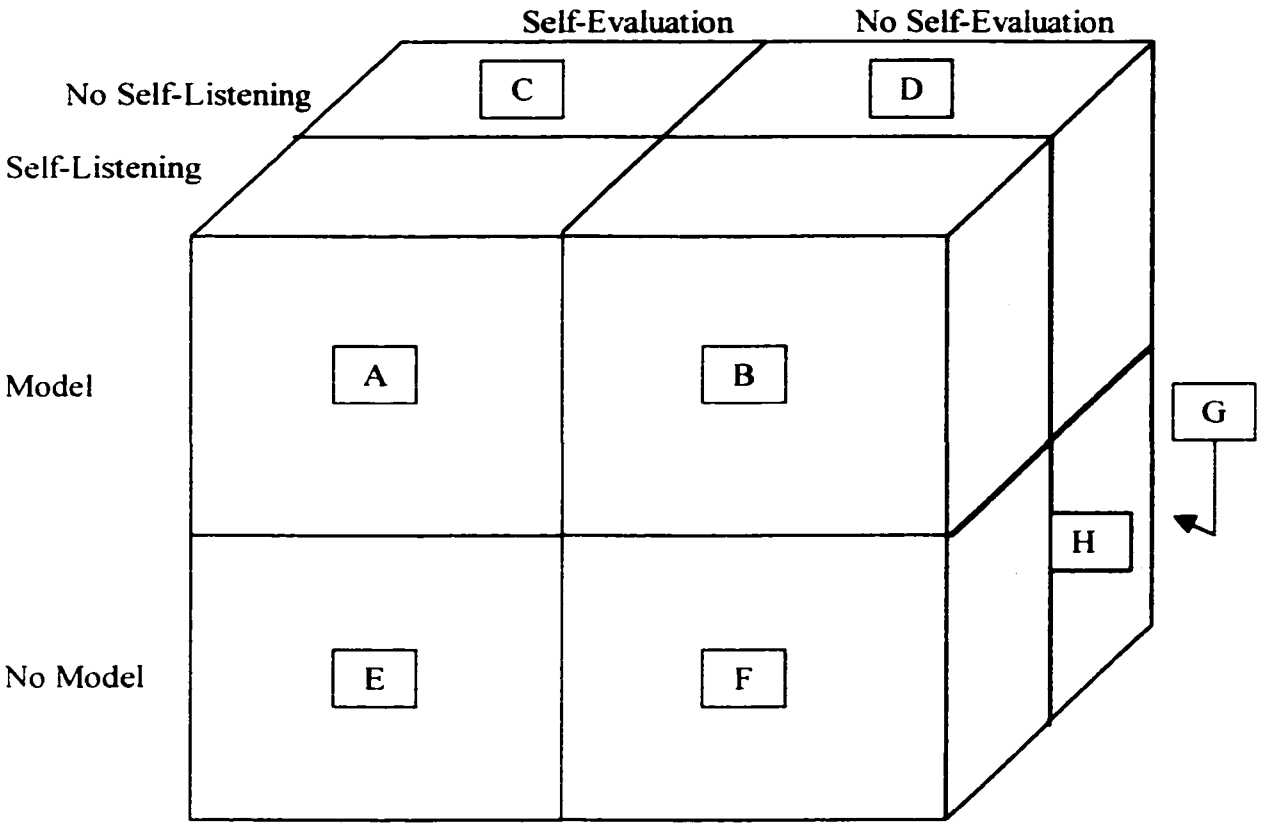


Figure 1. 2 x 2 x 2 pictorial representation of treatment groups.

Assignment to treatment groups was done by grade to negate the possible effect of current music performance ability on the study. Within each grade, students were placed in "score order" and then placed again alphabetically within score order. Thus the

person first in order within each grade was a flute player whose surname came first alphabetically while the person assigned the highest number within the grade was a percussionist whose surname appeared last alphabetically. A random table of numbers was then used to place students into each treatment group beginning with the ninth grade, followed by eighth and then seventh grade. Commencing with group A (chosen also by using a random table of numbers) and continuing through H, selected students were assigned to groups. When rotation through the groups was complete, the next student was assigned to group A and the rotation continued until all students in the grade were assigned to a group. The first eighth grade student chosen was placed in the group immediately following the final ninth-grade student. Similarly, the first seventh-grade student assigned to a group was placed in the group immediately following the last eighth-grade subject.

Time table

This study was carried out between January and March 2000 and constituted three phases over the course of 9 weeks. The first phase, pretreatment, lasted two weeks. During Week 1, students received self-evaluation training and "ran-through" the musical excerpts while in Week 2 they carried out pretest measures for music performance. The treatment period entailed Weeks 3-7 and included self-report attitude measures. The study concluded with two weeks of a post-treatment phase, which included a posttest performance measure during the penultimate week and practice attitude validity videotaping during the final week. This procedure is outlined in Figure 2.

Phase	Week	Procedures
Pre-treatment	Week 1	Self-Evaluation Training, Introduce the *Performance Etude*
	Week 2	Performance Pretest Measure
Treatment	Weeks 3-7	Treatment, with self-report Practice Attitude Questionnaire (PAQ)
Post-treatment	Week 8	Performance Posttest Measure
	Week 9	Videotape for PAQ Validity

Figure 2. Time table for study.

Procedures

Pre-Treatment Phase

All students were trained in the use of the Solo Evaluation portion of the Woodwind/Brass Solo Evaluation Form (WBSEF), adapted for use in this study during the first week of the investigation (see Saunders and Holahan, 1997). All terminology used on the form was explained and discussed to ensure participant understanding. A sample recording was played, and together the students completed the form. Participants then individually practiced completing the adapted WBSEF while listening to additional audiotape recordings. The recordings represented a variety of performances both good and poor and were discussed with the researcher, in conjunction with the WBSEF, after each performance to assess students' understanding and the proper use of the form.

Four "practice rooms" were prepared for the experiment adjacent to the ensemble rehearsal room. Each practice room was identical and equipped with a music stand,

chair, a SONY TCM-929 audiotape recorder, and an Audio-Technica ATR-30 microphone.

Also during Week 1, students were introduced to the music used during the study, identified to them as the *Performance Etude*. Copies of the *Performance Etude* were distributed to each student. Students then participated in two performances led by both the teacher and researcher, each lasting approximately 15 minutes. It is important to point out that these performances included only limited instruction by the teacher and researcher. Specific instructions given included the establishment of correct tempos and admonition of key and meter signatures. Following the 15-minute time period, copies of the music were collected. This procedure was repeated for two additional days during Week 1.

The performance pretest was given in Week 2. All students entered a practice room along with the researcher and then recorded performing the *Performance Etude* in its entirety, stopping only briefly between sections. After this initial performance, students were instructed to keep a copy of the *Performance Etude*, and to practice it throughout the duration of the treatment. Students were told they would not receive a grade for any of their performances, but should nevertheless prepare the piece to the best of their ability. Students were verbally prompted throughout the treatment period, both in class and during treatment, to practice the music. They were also surveyed each week to determine whether a new copy of the music was needed due to loss or damage.

Treatment Phase

Treatment commenced during week 3. After the "in-class" treatment (described below) participants were provided a tape to listen to depending upon their assigned practice condition. Students who received the Model treatment (groups A, B, C, D) were given an audiotape which contained a recording of the *Performance Etude* performed in its "ideal" form. Students receiving the Self-listening treatment (A, B, E, F) were each provided an audiotape which contained a recording of the student's own performance during the treatment. Each week, members of the Self-Listening group recorded and received a new recording of their updated performance dubbed over their previous performance. Students who were assigned to both the Model and Self-Listening groups (A, B) received tapes of both the "ideal" model representation and a tape of their own performance (updated each week). Students who were assigned to neither the Model or Self-Listening treatment (G, H) were provided with an audiotape (for control purposes) of a professional wind ensemble, the contents of which had no direct relationship to the *Performance Etude*. Students in all treatment groups were asked to listen to their tape daily whenever they practiced the piece at home.

Specific treatment instructions were provided to students in both oral and written form during the first treatment session. In subsequent treatments, students were reminded of the procedures to be followed and were monitored for their adherence to them by the researcher.

The written instructions provided to the Model group were to a) listen to the model recording, b) perform the etude matching the recording as closely as possible, c)

practice the piece (if time remained), d) complete the Practice Attitude Questionnaire (PAQ), and e) leave with their assigned tape. The students in the Model group who were also assigned to Self-Evaluation treatment additionally completed the adapted version of the WBSEF form. They did this immediately after step (b), performing the etude, and prior to filling out the PAQ. Members of the No Self-Evaluation group did not complete the WBSEF.

Students in the Self-Listening group were instructed to a) record themselves performing the etude in its entirety, b) rewind the tape and listen to their performance, c) practice the piece (if time remained), d) complete the PAQ, and e) leave with the tape they had just recorded. Members who were also in the Self-Evaluation group also completed the adapted WBSEF immediately after listening to their performance.

Students assigned to both the Model and Self-Listening treatments were directed to a) listen to the model recording, b) perform and record the etude, matching the recording as closely as possible, c) listen to their own performance, d) practice the piece (if time remained), e) complete the PAQ (and the adapted WBSEF if also a member of the Evaluation group), and f) leave with both the tape on which they recorded the etude and the model tape.

Students assigned to neither the Model nor Self-listening group physically practice the etude during the treatment period. They received instructions to a) perform the etude from beginning to end, b) practice the etude on their instrument in any manner they chose, and c) complete the PAQ (and the adapted WBSEF, if an Evaluation group member).

Post-Treatment Phase

Week 9 of the study was used to assess students' individual music performance for a second time. Similarly to the pretest performance, all students individually entered a practice room and recorded themselves. During the final week of the study (Week 10) 24 randomly selected students individually participated in a videotaped interview designed to determine validity of the PAQ.

Dependent Variables

Music Performance

Music performance was evaluated using three independent adjudicators' ratings of students' recorded pretest and posttest performances. The music selected as the performance material for the study (*Performance Etude*) began with the examination of a variety of musical works by the researcher and the participants' teacher to determine whether it fit established criteria. The criteria against which each musical work was evaluated was that it must: (a) incorporate a diversity of technical components appropriate for junior high musicians including a variety of articulations, styles, dynamics, rhythmic patterns, and a moderately wide melodic range; (b) be of appropriate difficulty so that a "ceiling effect" would not be established, and c) be of similar difficulty for each instrument involved in the study. Portions of three musical works were selected to constitute the *Performance Etude*. The *Performance Etude* was then given to three junior high school music educators to determine independently if it met the criteria. It was determined by all of them that it did.

The first excerpt selected as part of the *Performance Etude* was in the concert key of Bb and was slow and legato. The second was a *gigue* in 6/8 time marked *Allegro*. It incorporated various combinations of articulations, changing dynamics and was also written in the concert key of Bb. The third piece was in the concert key of Ab with a meter signature indicating 2/4. The tempo was marked *Allegretto con moto*, and combinations of sixteenth-note patterns were utilized throughout the excerpt. All musical examples were transposed into appropriate keys and ranges to equalize difficulty among instruments.

Performance was evaluated using the WBSEF (Saunders & Holahan, 1997). This instrument uses a 5-point criteria-specific rating scale that independently examines individual tone, intonation, technique/articulation, melodic accuracy, rhythmic accuracy, tempo, and interpretation. Criteria-specific scales typically use written descriptors to characterize the performance qualities necessary to achieve scores at increasingly higher levels. Judges are asked to select the descriptor that most accurately describes the performance they are evaluating. The WBSEF uses continuous performance criteria in six of the seven different categories, while in the technique/articulation section an additive (non-sequential) approach is employed. For the latter category, adjudicators are asked to mark any or all of five separate guidelines listed. The criteria are appropriate and accurate tonguing, appropriate slurs as marked, appropriate accents as marked, appropriate ornamentation as marked, and appropriate length of notes as marked (i.e. legato, staccato).

The WBSEF is a criteria-rated instrument, "designed to diagnose specific levels of instrumental accomplishment and/or deficiency, yielding measurement characteristics comparable to those found for other types of rating instruments" (Saunders & Holahan, 1997, p. 261). Since the present study deals with self-evaluation, an instrument designed to properly diagnose deficiencies and accomplishments is ideal for both assisting students in their self-evaluation while at the same time providing a reliable measurement source.

Internal reliability of the WBSEF has been found to be high, with a reported median alpha reliability of .92; however, interjudge reliability has not been examined. The authors of the WBSEF have suggested that it has strong validity in terms of its diagnostic abilities. The evidence they use to support this claim is the low correlation exhibited among subareas in conjunction with the high independent correlation of each subarea to the overall score. They state that this provides evidence of judges' abilities to isolate the performance characteristics described in the subareas.

As mentioned previously, the Solo Evaluation section of the instrument was the only part of the instrument utilized in the current study. This may cause concern in terms of reliability issues; however, a stepwise multiple regression analysis completed by the authors of the instrument during the initial study of the instrument determined that four subareas (tone, technique/articulation, rhythm accuracy, and interpretation) on the Solo Evaluation portion of the test contributed most to the overall score on the WBSEF. Thus, the Solo Evaluation section contributes highly to the overall reliability figures. Furthermore, each of the subareas mentioned above correlated highly with the overall performance score on the total WBSEF score (r = .72 - .82).

Due to the somewhat advanced terminology used on the WBSEF (it was created for use by professional adjudicators), it was modified for use with junior high students for this study. A small group of junior high age students who were enrolled in a different school than the one used in the study were asked to assist with the adaptation of the form. Guided by the researcher, students were asked to suggest more appropriate ways to state each descriptor. Student descriptors were then examined by music educators to determine whether the new terminology was similar to the original form of the WBSEF. The educators made minor recommendations to the researcher for some modifications, which the researcher considered when developing the final form of this measurement tool seen in Appendix B.

Practice Attitude

Practice attitude was examined to determine whether students in a particular treatment group possessed higher levels of enjoyment and satisfaction with the treatment process. It was felt that this construct was important to measure because it would shed light on whether a particular practice method or self-evaluation method would be used in the students' future practice routines. If students did not enjoy a specific treatment, or felt that it did not help them improve, it seems plausible that it would not be used during their own independent practice. Conversely, when a practice method or self-evaluation was received positively, students would tend to return to that procedure in the future.

Students' practice attitude was measured using the Practice Attitude Questionnaire (PAQ), a researcher created, self-report, Likert-type instrument that was completed by all students immediately following each treatment session. The survey utilized research on

student attitudes as its basis for construction and implementation (Cutietta, 1992). The survey was used to elicit answers regarding students' feelings, beliefs, and values with respect to their experience in their assigned treatment condition. In constructing the survey, a list of questions were generated which sought to reflect as many possible interactions of attitude components (feelings, beliefs, values) and response components (cognitive, behavioral, affective). The final version of the survey appears in Appendix C.

To determine the content validity of the PAQ, one student from each of the eight treatment groups within each grade was randomly selected to be interviewed regarding their attitudes about the study. The interviews were videotaped and then viewed by three independent evaluators who completed a PAQ for each student. The evaluators' scores were each compared to the students' self-reports completed during the posttesting to determine validity.

Independent Variables

The relationships among the three independent variables (modeling, self-listening, self-evaluation) were investigated with respect to instrumentalists' music performance and practice attitude. Using one of two methods, the presence of the variable in the treatment or the absence of the variable, each independent variable was examined.

The audiotape models were prepared at a recording studio at a large southwestern university school of music. University music majors were recruited to perform and record the musical excerpts. Separate recordings were made by individual musicians for each instrument used in the study. Musicians were provided the *Performance Etude* in advance of the recording session and asked to prepare it to "perfection". The flute

recording was made first. Extensive time was taken with this performer to assure that the performance was in "ideal" form as it was to be used as a model for the other studio recordings. While in the studio, each musician had a 30-minute time period to record as many "takes" of the piece as possible. These recordings took place after the performers listened to the flute model. The researcher then guided the musicians to be sure that their performances matched the flutist as closely as possible. Since the *Performance Etude* was actually comprised of three separate etudes, each one was recorded separately. Three professional musicians then selected the "best" recording of each etude to be used in the study.

Null Hypotheses

The following null hypotheses were investigated each with an established alpha level of .05:

Music Performance

1a. There is no significant relationship between modeling, considered alone, and music performance.

1b. There is no significant relationship between self-listening, considered alone, and music performance.

1c. There is no significant relationship between self-evaluation, considered alone, and music performance.

1d. There is no significant relationship between modeling and self-listening, considered in combination, and music performance, independent of the effects of modeling and self-listening alone.

1e. There is no significant relationship between modeling and self-evaluation considered in combination, and music performance, independent of the effects of modeling and self-evaluation alone.

1f. There is no significant relationship between self-listening and self-evaluation considered in combination, and music performance, independent of the effects of modeling and self-evaluation alone.

1g. There is no significant relationship between modeling, self-listening, and self-evaluation considered in combination, and music performance, independent of the effects of modeling, self-listening, and self-evaluation alone.

Practice Attitude

2a. There is no significant relationship between modeling, considered alone, and practice attitude.

2b. There is no significant relationship between self-listening, considered alone, and practice attitude.

2c. There is no significant relationship between self-evaluation, considered alone, and practice attitude.

2d. There is no significant relationship between modeling and self-listening, considered in combination, and practice attitude, independent of the effects of modeling and self-listening alone.

2e. There is no significant relationship between modeling and self-evaluation, considered in combination, and practice attitude, independent of the effects of modeling and self-evaluation alone.

2f. There is no significant relationship between self-listening and self-evaluation, considered in combination, and practice attitude, independent of the effects of modeling and self-evaluation alone.

2g. There is no significant relationship between modeling, self-listening, and self-evaluation considered, in combination, and practice attitude, independent of the effects of modeling, self-listening, and self-evaluation alone.

Analysis

The data were analyzed using the Standard Version of SPSS Base 8.0 software. Specific dependent variables were investigated in the following manner:

Music Performance

Students' individual audiotape recordings from both pretest and posttest performances were re-recorded onto high quality cassette tapes for the purpose of adjudication by three experienced instrumental music educators unfamiliar with objectives of the project. Each tape was made using a different sequence of performances to control for order-effect among judges. Pretest and posttest performances were also commingled on the tapes. Adjudicators were trained in the use of the WBSEF using sample recordings of the music performed by students.

All sub-areas of the WBSEF as well as the overall score were examined using General Linear Model--multivariate statistics with repeated measures. Interjudge reliability was calculated for subareas and the overall score using the Pearson Correlation. As noted in Table 1, correlation between judges ranged from moderate to strong (r = .500 - .827).

Table 1.

Pearson Correlation Matrix for Interjudge Reliability Student Performance Ratings

Pretest Scores	Judge 1 & 2	Judge 2 & 3	Judge 1 & 3	Mean
Tone	.497	.544	.458	.500
Intonation	.531	.539	.485	.518
Technique/Articulation	.631	.581	.543	.585
Melodic Accuracy	.550	.593	.613	.585
Rhythmic Accuracy	.717	.576	.583	.625
Tempo	.466	.634	.500	.533
Interpretation	.648	.589	.593	.610
Overall	.738	.766	.731	.745
Posttest Scores				
Tone	.607	.556	.595	.586
Intonation	.458	.499	.615	.524
Technique/Articulation	.741	.601	.684	.675
Melodic Accuracy	.725	.699	.698	.707
Rhythmic Accuracy	.819	.761	.753	.778
Tempo	.830	.728	.724	.761
Interpretation	.715	.687	.716	.706
Overall	.853	.807	.821	.827

Practice Attitude

Students' attitudes toward practice method were calculated by averaging scores on the PAQ each week for each treatment group. Means and standard deviations were calculated and compared between practice strategies across the span of the treatment. Validity of the researcher created PAQ was investigated during the final week of the study. For each grade level, one student from each treatment group was randomly

selected to orally answer questions (listed in Figure 3) relating to the their experiences during the treatment period. Their answers were videotaped and then viewed by three independent assessors who each completed their own PAQ for that student. These PAQ's were then scored and compared to the other assessors' scores (to test reliability) and then to the students' PAQ from the final treatment of the study to verify validity. Reliability between judges was found to be very strong ($M = .86$), and the correlation between the mean the judges' scores and the students' final self-scores was also strong ($M = .77$).

---

1. Would you say you enjoyed or did not enjoy participating in the practice sessions? Why?
2. What did you like (or not like) about it the most? Why?
3. How did you feel when you were practicing? Why?
4. Did it make you want to practice more? Why?
5. Do you think you might want to practice that way in the future? Why?
6. Did this method of practice help you improve more than you normally would? Why?

Figure 3. List of questions asked to students for validation of the PAQ.

Test Statistic

There has been much discussion in educational measurement regarding which statistic to use when pursuing multivariate analysis, and different sources conflict in their recommendations. Generally, four multivariate tests are produced by most statistical computer programs, Pillai-Bartlett Trace, Wilk's Lambda, Hotelling-Lawly Trace, and

Roy's Largest Root. Roy's Largest Root appears to be the most robust test when assumptions of multivariate analysis are met, though all remain robust (and may even become more conservative) when there are *occasional* (emphasis added) extreme departures from these assumptions (Hand & Taylor, 1987). Pillai's Trace or Wilk's Lambda is generally used when the data do not conform to the test assumptions. Therefore, the data were examined to determine whether the assumptions of normality, multivariate normality, sphericity, and homoscedasticity were met.

In regard to music performance, pretest and posttest scores for all combinations of independent and dependent variables were examined for normality with the Shapiro-Wilk Test of Normality. Though too numerous to list specific groups here, it was determined that 60 of the 384 (15.63%) of groups did not meet normality assumptions. Standardized residual plots for each dependent variable were inspected to ascertain multivariate normality. It appeared that all combinations of variables exhibited moderately strong to very strong linear relationships to one another.

The Levene's test was run on all dependent variables to determine whether groups had equal variances (homoscedasticity). Results indicated that all groups were similar in their variances except for the posttest results for technique/articulation [F (7, 74) = 2.324, p = .034], and melodic accuracy [F (7, 74) = 2.681, p = .016]. Sphericity was assessed using Mauchly's Test of Sphericity. For all dependent variables this assumption was strongly upheld, as the test statistic's value was equal to 1.000 for each variable. In light of these data, it appears that the assumptions for the use of a multivariate repeated measures design have generally been met for all assumptions except normality.

Normality, however, was clearly *not* met for a large number of cases, thus it seems appropriate that Pillai's Trace be used as it is more robust to departures from multivariate assumptions.

The data for Practice Attitude were explored to determine whether the assumptions for using the multivariate test had been met. It was found that the data for 9 of 70 (12.86%) combinations of independent and dependent variables deviated from normality in some manner. Furthermore, assumptions of multivariate normality, homoscedasticity, and sphericity were not met. Thus the Pillai's Trace was used for the analysis of Practice Attitude as well.

Validity

As with most statistical studies in the social sciences, true experimental designs are nearly impossible. The current study is not an exception, though an attempt was made to control as many extraneous variables as possible. The research groups were selected from one school of which 82 of 137 (59.85%) students chose to participate in the study; therefore, the participants are most likely not representative of all junior high students. Generalization to the population of all junior high students should be considered with care. The quasi-experimental design undertaken did control for maturation, history, testing, statistical regression, selection bias, and experimenter bias, all which are threats to internal validity. To further control extraneous variables, students were asked regularly whether they had their copy of the music and/or tape(s) were lost or damaged. If so, they were immediately replaced. Additionally, mortality was not a factor as all students who began the study completed it.

The two measurement instruments used in the study, the PAQ and the adapted WBSEF, also generated concern with validity and reliability issues. The PAQ had not been evaluated in regard to its reliability or validity prior to the study, though its content validity was evaluated during the present study. The content validity was quite high, signifying that it does purport to measure student attitude quite well as it was defined in the present study.

Interjudge reliability of the adapted WBSEF has not been examined, though the authors of the original WBSEF do claim that it has internal reliability and exhibits indirect evidence of internal validity. Nonetheless, the adapted WBSEF also has never been examined in terms of its interjudge reliability or validity until the current study. Mean interjudge reliability scores ranged from moderate to strong (r = .50 - .83) in the present study.

Internal validity may not be as strong on the WBSEF as on the original form of the measure. Scores among the seven performance subareas were highly correlated in the current study, indicating that each field may not have been evaluated separately, but instead the adjudicators may have made global evaluations. This does not necessarily mean that judges evaluated the performances as a whole, but may indicate that junior high students (as opposed to advanced high school students as examined in the original study) may progress equally through all of the performance subareas evaluated in the study.

Assumptions of the statistical tests must also be considered when discussing validity. As discussed, a number of assumptions were not met for the multivariate tests

undertaken; however, cell sizes were greater than 20 for calculations of all main effects and interactions except the four-way interaction of test x model x self-evaluation x self-listening. Generally, multivariate statistics are robust to departures from these assumptions with cell sizes of 20 or more (Jaccard and Becker, 1997). The results of the four-way interactions should be interpreted carefully in light of the small cell sizes.

Chapter 4

Results and Data Analysis

The purpose of this study was to examine the effects that modeling, self-listening, and self-evaluation have on middle school instrumentalists' music performance and attitude about practice.

Music Performance

A general linear model (GLM) repeated measures analysis with multiple dependent variables was performed to determine relationships among two modeling conditions (Model, No Model), two self-listening conditions (Self-Listening, No Self-Listening), two self-evaluation conditions (Self-Evaluation, No Self-Evaluation) and test administration (pretest, posttest) scores of seven performance subareas (tone, intonation, technique/articulation, melodic accuracy, rhythmic accuracy, tempo, interpretation) and overall performance. Overall performance scores were calculated for both pretest and posttest data using sums of each subarea. Tests exhibiting statistically significant effects were succeeded by univariate analyses on each dependent variable. Relationship strength was determined using eta^2, while the nature of relationships was examined using profile plots and descriptive statistics. An α level of .05 was set for all tests.

Table 2 shows effects of the multivariate repeated measures test. A four-way interaction for model, self-listening, self-evaluation and test was found to be not statistically significant [$F(7, 68) = .666$, $p = .700$]. There were two statistically significant three-way interactions. First is a between-subjects interaction of model, self-listening, and self-evaluation [$F(7, 68) = 2.162$, $p = .049$, $eta^2 = .182$]. Second is the

Table 2.

Multivariate Tests of Music Performance Scores

	Hypo. df	Error df	F	Sig.
Between Subjects				
Self-Evaluation	7	68	.985	.450
Self-listening	7	68	.830	.566
Model	7	68	1.890	.085
Self-Evaluation * Self-listening	7	68	1.011	.432
Self-Evaluation * Model	7	68	2.030	.064
Self-listening * Model	7	68	1.852	.091
Self-Evaluation * Self-listening * Model	7	68	2.162	.049
Within Subjects				
Test	7	68	12.560	.000
Test * Self-Evaluation	7	68	.940	.482
Test * Self-listening	7	68	1.073	.482
Test * Model	7	68	2.303	.036
Test * Self-Evaluation * Self-listening	7	68	1.149	.344
Test * Self-Evaluation * Model	7	68	2.185	.046
Test * Self-listening * Model	7	68	.518	.818
Test * Self-Eval * Self-listening * Model	7	68	.666	.700

interaction of model, self-evaluation, and test [F (7, 68) = 2.185, p = .046, eta^2 = .184]. The interaction of model and test [F (7, 68) = 2.303, p = .036, eta^2 = .192] was the only statistically significant two-way interaction. Statistically significant main effects were found only for the test (pretest/posttest) condition [F (7, 68) = 12.560, p = .000, eta^2 = .564]. The interaction effects of between-subjects variables (modeling, self-listening, self-evaluation) were calculated using combined pretest and posttest mean scores, thus providing little meaning to the present study. In addition, subsequent analysis of these between-subjects scores revealed no statistically significant differences for any of the dependent variables, so results of this effect will not be pursued. Furthermore, univariate tests revealed differences between pretest and posttest groups for all subareas of music performance including overall performance for all groups. Statistical analysis revealed that posttest scores were significantly higher than their pretest counterparts for all groups in all subareas except one. The group receiving no modeling, no self-listening, self-evaluation showed a decrease in their rhythmic accuracy score from pretest to posttest. These results will likewise not be explored further.

With regard to the three-way interaction of model, self-evaluation, and test, follow-up univariate analyses revealed statistically significant results for tone [F (1, 74) = 4.029, p = .048, eta^2 = .052], melodic accuracy [F (1, 74) = 3.994, p = .049, eta^2 = .051], rhythmic accuracy [F (1, 74) = 6.489, p = .013, eta^2 = .081], interpretation [F (1, 74) = 9.539, p = .003, eta^2 = .114], and overall performance scores [F (1, 74) = 5.430, p = .023, eta^2 = .068]. The nature of these interactions is displayed using profile plots in Figure 4, while mean pretest/posttest gain scores for all subareas appear in Table 3, and the entire

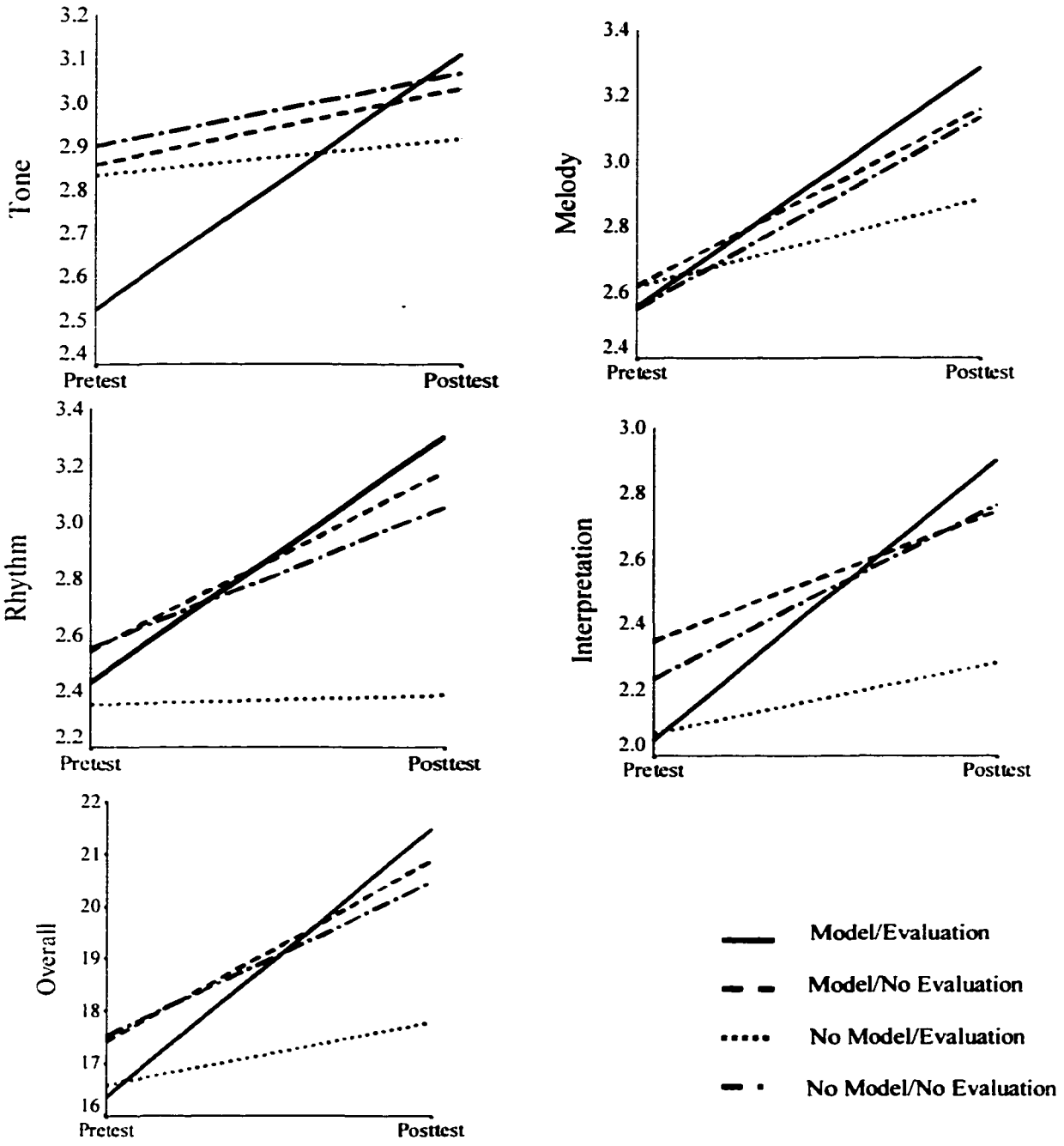


Figure 4. Mean performance gain score interactions for significant model x self-evaluation x test.

Table 3.

Music Performance Pretest/Posttest Mean Gain Scores (and Standard Deviations) for Model x Self-Evaluation x Test Interaction

	Mod—SE	Mod--No SE	No Mod—SE	No Mod--No SE
	n = 20	*n = 21*	*n = 20*	*n = 21*
Tone	.617 (.487)	.175 (.680)	.083 (.457)	.159 (.501)*
Intonation	.433 (.553)	.349 (.662)	.133 (.476)	.191 (.478)
Tech./Artic.	.767 (.542)	.619 (1.050)	.217 (.487)	.429 (.739)
Melody	.750 (.629)	.540 (.654)	.267 (.427)	.571 (.518) *
Rhythm	.917 (.550)	.635 (.623)	.033 (.571)	.476 (.688)*
Tempo	1.017 (.587)	.746 (.649)	.250 (.482)	.460 (.687)
Interpretation	.900 (.553)	.397 (.647)	.217 (.408)	.508 (.564)**
Overall	5.400 (2.832)	3.460 (4.048)	1.200 (2.333)	2.794 (3.255)*

Note. Mod = Model. SE = Self-Evaluation.

*$p < .05$. ** $p < .01$.

univariate analysis table is presented in Appendix D. Scheffé post-hoc tests were performed on the mean gain scores of the test x model x self-evaluation interaction. These tests revealed that the Model/Self-Evaluation group improved more than the No Model/Self-Evaluation group for tone, melodic accuracy, rhythmic accuracy, interpretation and overall performance. Furthermore, the Scheffé test revealed no

differences in scores between the Model/No Self-Evaluation and No Model/No Self-Evaluation for any performance subarea indicating that these groups improved equally.

Univariate analysis disclosed that all performance areas except intonation and melodic accuracy had statistically significant interactions between Model and Test. Significant findings were discovered for tone [F (1, 74) = 4.330, p = .041, eta^2 = .055], technique/articulation [F (1, 74) = 4.083, p = .047 eta^2 = .052], rhythmic accuracy [F (1, 74) = 12.467, p = .001, eta^2 = .144], tempo [F (1, 74) = 11.628, p = .001, eta^2 = .136], interpretation [F (1, 74) = 4.007, p = .049, eta^2 = .051], and overall performance [F (1, 74) = 9.277, p = .003, eta^2 = .111]. The natures of these relationships are displayed in Figure 5 while Table 4 displays the pretest and posttest scores for all subareas.

For each subarea score and overall performance that was significant (and also those that were not), students receiving the Model treatment showed a greater increase in performance scores than did students in the No Model group. These results indicate that certain music performance scores increased more when students listened to a model recording.

Practice Attitude

Each response category on the PAQ was assigned a corresponding numeric value, strongly disagree (1), disagree (2), agree (3), and strongly disagree (4), and the mean score for the seven questions were calculated to obtain a composite score. Mean scores were then calculated for all groups and subjected to a GLM-multivariate repeated measures analysis to determine relationships among two self-evaluation conditions, two

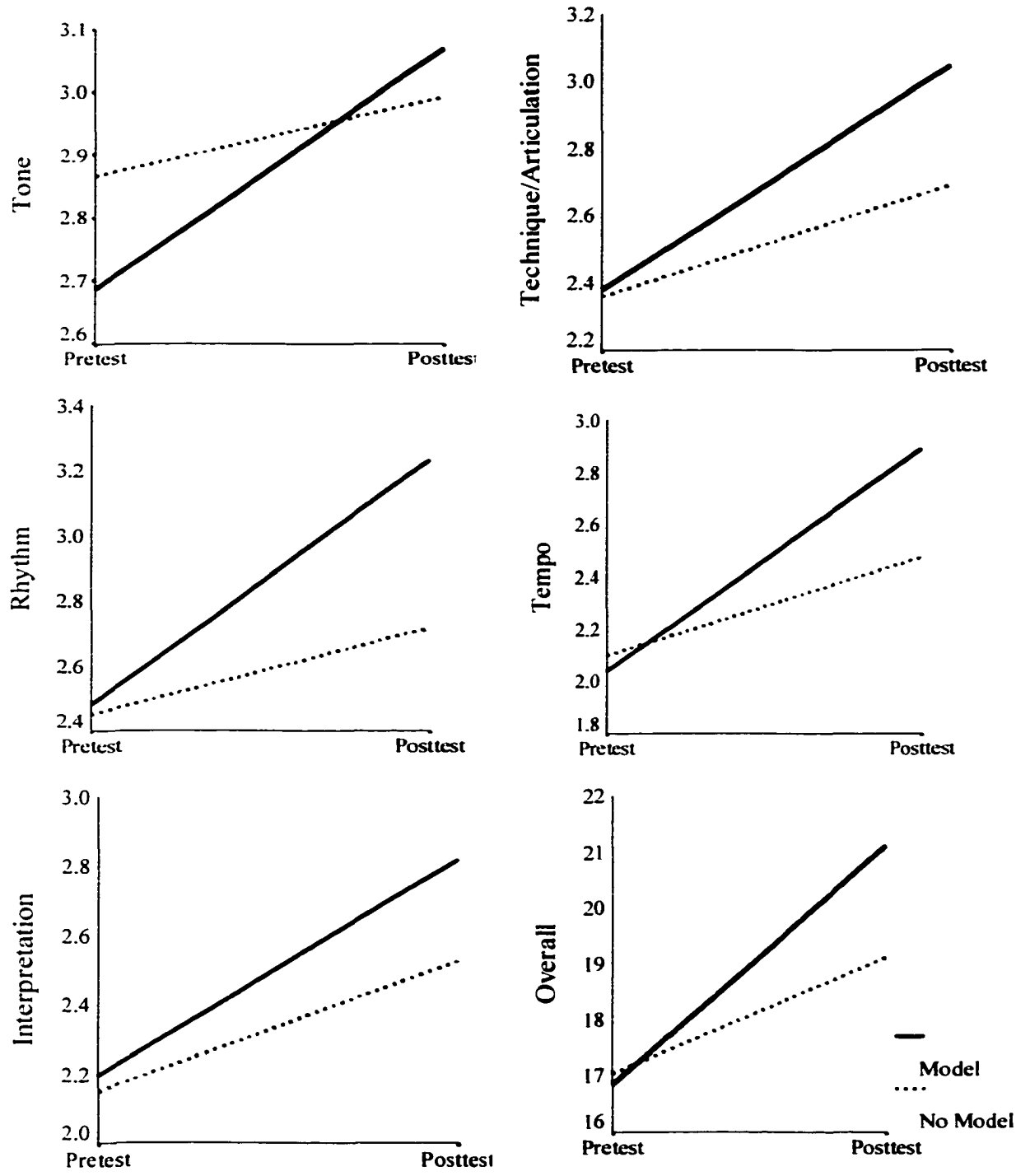


Figure 5. Mean Performance Pretest/Posttest Scores for Significant Model x Test.

Table 4.

Music Performance Pretest/Posttest Scores for Modeling Condition

	Model (n = 20)				No Model (n = 20)			
	Pretest		Posttest		Pretest		Posttest	
	M	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*
Tone	2.686	.105	3.068	.118	2.867	.107	2.992	.120*
Intonation	2.487	.106	2.864	.105	2.542	.109	2.717	.108
Tech/Art	2.379	.145	3.047	.167	2.358	.148	2.692	.171*
Melody	2.578	.114	3.213	.134	2.583	.117	3.008	.137
Rhythm	2.481	.116	3.232	.153	2.450	.119	2.717	.156***
Tempo	2.039	.100	2.889	.151	2.100	.102	2.475	.154**
Interpretation	2.196	.110	2.820	.139	2.150	.112	2.525	.142*
Overall	16.846	.718	21.133	.892	17.050	.735	19.125	.916**

Note. * $p < .05$. ** $p < .01$. *** $< .001$.

modeling conditions, two self-listening conditions, and five practice attitude measures. An α level of .05 was set for all tests.

As indicated in Table 5 and Figure 6, no statistically significant interactions or main effects were found in the data. This indicates that mean practice attitude scores remained constant throughout the duration of the study for all of the groups and also that

Table 5.

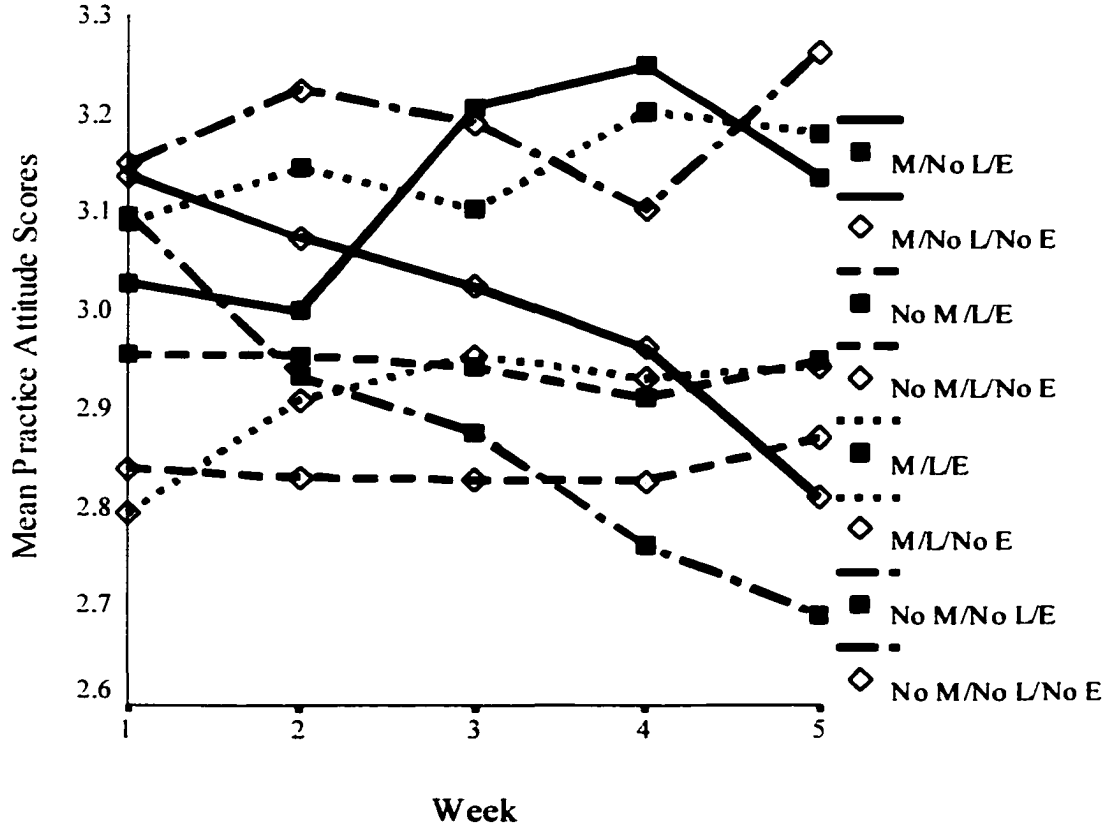
Tests of Within-Subjects Effects for Practice Attitude Scores

	Hypo. df	Error df	*F*	Sig.
Within Subjects				
Test	4	71	.205	.935
Test * Model	4	71	1.736	.152
Test * Self-listening	4	71	1.034	.396
Test * Self-Evaluation	4	71	.561	.692
Test * Model * Self-listening	4	71	.576	.681
Test * Model * Self-Evaluation	4	71	1.212	.313
Test * Self-listening * Self-Evaluation	4	71	.487	.745
Test * Model * Self-listening * Self-Evaluation	4	71	1.411	.239

the treatment groups did not differ from each other in terms of their attitude about the procedure at any time during the treatment period. Furthermore, each group seemed to have a "strong" attitude toward their particular practice strategy. Using a four-point scale, with four representing the highest attitude, the grand mean throughout the study was high [M = 3.003, SD = .061] with scores ranging from 2.692 (.216) to 3.263 (.216).

Summary of Results

A total of fourteen null hypotheses were tested in this study relating to the effects that modeling, self-listening, and self-evaluation have on music performance and practice



Note. M = Model. L = Self-Listening. E = Self-Evaluation

Figure 6. Mean practice attitude scores of model x self-listening x self-evaluation x test.

attitude. The results are summarized as follows:

Music Performance

Null Hypothesis 1a. There is no relationship between modeling, considered alone, and music performance. This null hypothesis was rejected. There were significant interactions between model x test for tone, technique/articulation, melody, rhythm,

interpretation, and overall performance mean scores on the WBSEF.

Null Hypothesis 1b. There is no relationship between self-listening, considered alone, and music performance. The null hypothesis was retained. There were no significant interactions between self-listening and test for any WBSEF performance area.

Null Hypothesis 1c. There is no relationship between self-evaluation, considered alone, and music performance. This null hypothesis was retained. Group mean scores did not differ for any performance area for the interaction between self-evaluation x test.

Null Hypothesis 1d. There is no relationship between modeling and self-listening, considered *in combination*, and music performance, independent of the effects of modeling and self-listening alone. This null hypothesis was retained. There were no significant interactions between model and self-listening for any performance area on the WBSEF or overall performance.

Null Hypothesis 1e. There is no relationship between modeling and self-evaluation considered *in combination*, and music performance, independent of the effects of modeling and self-evaluation alone. This hypothesis was rejected. Statistically significant interactions of model, self-evaluation, and test were found between group mean scores for tone, melodic accuracy, rhythmic accuracy, interpretation, and overall performance. Group gain scores indicated that Model/Self-Evaluation scores were highest for all areas while gain scores were lowest for the No Model/Self-Evaluation group. The Model/No Self-Evaluation and No Model/No Self-Evaluation scores for groups fell between the other two conditions for all performance subareas and overall

performance, however between each other they did not show a consistent order for any of the dependent variables.

Null Hypothesis 1f. There is no relationship between self-listening and self-evaluation, considered *in combination*, and music performance, independent of the effects of modeling and self-evaluation alone. The hypothesis was retained. There were no significant interactions between self-listening, self-evaluation, and test for any subareas on the WBSEF and overall performance.

Null Hypothesis 1g. There is no relationship between modeling, self-listening, and self-evaluation considered *in combination*, and music performance, independent of the effects of modeling, self-listening, and self-evaluation alone. The null hypothesis was retained. There were no significant four-way interactions for any of the dependent variables examined in this study.

Practice Attitude

Seven null hypotheses were examined to determine whether alone, or in combination, the effects of modeling, self-listening, and self-evaluation had on practice attitude. All of these hypotheses were retained. No differences were found between any groups' mean scores on the PAQ over the course of the investigation. Furthermore, no groups' scores changed significantly throughout the duration of the study.

Chapter 5

Discussion of Results and Conclusions

Discussion

The purpose of this study was to examine the effects that modeling, self-listening, and self-evaluation practice strategies have on junior high instrumentalists' music performance and practice attitude. Results of this study suggest that the type of strategy used during practice can affect music performance skill acquisition among junior high wind and percussion students.

Music Performance

Previous researchers have suggested that listening to a model recording can be as effective or more effective than other practice strategies including physical practice for improving music performance. In general, results of the current study support these findings, yet may shed new light on the effects modeling may have on specific areas of performance. Further, modeling appears to be influenced by the presence or absence of self-evaluation as indicated by the significant interaction effects between model and self-evaluation.

Two performance subareas, technique/articulation ($p = .047$) and tempo ($p = .001$) appeared to be significantly affected by only the test x model condition, as the groups who listened to a model improved their performance in these two areas more than those that did not listen to a model. This suggests that students can improve their performance tempo, articulation style, tonguing, slurring, and correct use of accents, staccato and legato by listening to an "ideal" performance.

Melodic accuracy was affected by the interaction of model and self-evaluation (p = .049); however, there was no overall main effect for modeling alone, as the Model group did not significantly differ from the No Model group (p = .099). The interaction effect indicates that listening to a model may not improve melodic accuracy in all situations. It appears that modeling assists in improving melodic accuracy when self-evaluation is involved, but listening to a model is not detectably more effective when there is no self-evaluation. Scores for groups that received the no self-evaluation treatment were not significantly different when examined under different modeling conditions.

The most frequently observed statistically significant phenomenon among the dependent variables were differences among groups for both the main effects of modeling and the interaction of modeling and self-evaluation. This occurred in the areas of tone (p = .041 for main effect, p = .048 for interaction), rhythmic accuracy (p = .001, p = .013), interpretation (p = .049, p = .003), and overall performance (p = .003, p = .023). For the areas of tone and interpretation, it appears that the influence of the main effects can be traced to the large difference between the Model/Self-Evaluation group and the No Model/Self-Evaluation group uncovered in the interaction; as there were no significant differences in mean scores between the Model/No Self-Evaluation and No Model/No Self-Evaluation groups. In regard to rhythmic accuracy, a statistically significant difference was found between the Model/Self-Evaluation and No Model/Self-Evaluation groups along with differences between Model/No Self-Evaluation and No Model/Self-Evaluation groups. Both groups receiving the Model treatment outperformed the No

Model/Self-Evaluation condition. These results seem to indicate that listening to a model performance is effective for improving performance of tone, rhythm, and interpretation when self-evaluation is involved, while it is not detectably more effective when self-evaluation is not included.

The results of these significant findings indicate that listening to a model is either as effective or more effective at improving certain performance areas than not listening to a model. Listening to a model, in and of itself, appears to be better than no model for improving technique/articulation and tempo, as well as the areas of tone, melodic accuracy, rhythmic accuracy, and interpretation when self-evaluation is also involved. Furthermore, listening to a model is not detectably more effective than listening to no model for improving tone, melodic accuracy, rhythmic accuracy, and interpretation when self-evaluation is not included, though it is as effective as no model for these performance areas.

As mentioned, modeling appears to be a viable method for improving technique/articulation and tempo under all conditions. Tone, melodic accuracy, rhythmic accuracy, and interpretation are affected when self-evaluation is undertaken. Without a model to compare to their own performance, students may make inaccurate assumptions regarding their playing ability. This incorrect perception of their competence could alter goals they set for themselves, which in the case of an assumed "good" performance may lessen their desire to practice as they have perceived their targeted objective to already been achieved. When students see their performance as poor, they may "give up" and not try to attain what they perceive to be a good performance. Informal interviews with

students involved in the study who did not listen to a model found that they were more apt to think they had performed adequately rather than poorly.

The area of intonation does not appear to be affected by any examined practice treatment, as there were no differences in main effects or interactions between any groups. Intonation is a difficult area for junior high students to comprehend. Often they are instructed that intonation refers to how well they are "in tune" compared to another musician or electronic tuning device, rather than whether a certain pitch is in tune compared to other pitches they play. Students are not adept at identifying these melodic intonation problems and focus their attention on other areas for improvement. In other words, they do not recognize a need to improve because they do not see a problem, tending to focus on areas where they do realize a problem exists.

In and of itself, self-evaluation as presented in the current study, does not appear to be an effective tool for improving students' music performance. Students in the self-evaluation group improved no differently than those not receiving self-evaluation in all subarea scores and overall performance. This conflicts with previous studies by Davis (1981), and Sparks (1990) who independently determined that self-evaluation was somewhat effective for improving music performance. The reason for this may be traced to the procedures used for self-evaluation. Whereas in the Davis and Sparks studies students received self-evaluation in a group setting coupled with guidance by a teacher, the present study sought to determine whether students could use self-evaluation in isolation. Aitchison (1995) also found that students improve their performance more when evaluative feedback is received from a teacher than when they self-evaluate.

Aitchison also determined that students could increase their self-evaluation accuracy with practice, yet this does not necessarily transfer into greater performance improvement. It appears then that though students are able to make accurate analyses of their presentation, the results of the present study suggest that they are unable to transfer these assessments into meaningful performance improvement.

Listening to a model was more effective than not listening to a model for improving tone, melodic accuracy, rhythmic accuracy, and interpretation when self-evaluation took place, but this was not true when no self-evaluation took place. The question arises as to why this is so. When students had the opportunity to view regularly the areas of performance on which they were evaluated, it may have changed how they approached their practice. The adapted WBSEF, which was used by students in the self-evaluation group to evaluate their performance each week, included verbal descriptors of each performance area. These descriptors may have acted as a guide for students. While listening to a model students may have been reminded of these areas and could focus teir attention, and thus their performance, on these areas. Students who did not participate in self-evaluation, and thus did not have access to the adapted WBSEF, did not have a guide to assist them in identifying areas where they needed to improve. The model did not appear to be more effective among this group because students had nothing outside of themselves to remind them of performance areas they need to improve. Two areas, technique/articulation and tempo improved in all groups that listened to a model regardless of self-evaluation condition. Perhaps these two areas are more easily

identified and improved upon by junior high students who listen to a model regardless of whether they are guided through the improvement process by a self-evaluation form.

In the present study, self-evaluation (unlike previous studies) showed no positive effects for improving music performance. As defined by this study, self-evaluation referred to the completion of the adapted WBSEF during weekly treatment. This limited method of self-evaluating appears not to be an effective method for improving junior high students' music performance. This approach to self-evaluation was selected because it seemed reasonable that the process could and would be enacted by students in their daily practice routines. Perhaps a manner of self-evaluation that more thoroughly and more often engages students in thoughtful reflection on their music performance may be more a more appropriate tool and yield promising results.

There is no direct statistical indication that self-listening is an effective practice strategy for improving music performance. Students who received self-listening treatment scored no differently than those who did not. Although past studies (Bundy 1987; Kepner, 1986) have shown that students' ability to detect musical errors is more accurate when listening to audiotape than during a live performance, these studies did not indicate that this ability transfers to improvement in performance. Detecting errors and finding solutions to correct those errors and thus improve performance appear to be skills that are mostly unassociated with each other.

The combinatorial effects of self-evaluation, self-listening, and modeling together were found to have no relation to music performance. This is somewhat surprising as it was predicted at the start of this investigation that the combination of these three

treatments would provide the greatest effects on music performance. These results also indicate that no practice strategy investigated in this study was any more effective for improving music performance than any other when considered alone. It is important to remember that each of these groups only had 10 or 11 participants, so the possibility of a Type II error does exist with this interaction. Furthermore, five groups exhibited nonnormal distributions, further complicating the analysis of the interaction. Future studies investigating these data should consider increasing the number of participants involved.

As mentioned earlier (Table 1), interjudge reliability was only moderate for some subareas of music performance. This was not foreseen during practice evaluation sessions by the judges. Perhaps the three performance excerpts judges listened to, with time for reflection, comments, and questions, was not enough to adequately gauge reliability. In addition, the WBSEF may be part of the problem with the less than desirable reliability figures. The measurement instrument was created as a tool to evaluate the performance of high school students auditioning for an all-state festival ensemble. Obviously the performance capabilities of this population are different from the junior high students who were participants in the present study. Therefore, the evaluation descriptors on the WBSEF may not be broad enough to encompass all performance attributes of junior high musicians and may include descriptors that cannot be achieved for this level of a student and consequently lead to non-normal distributions. For example, the number of students displaying a "full rich tone...in all ranges and

registers" would almost always be small for this age group while "is not a tone quality characteristic of the instrument" would be higher for junior high students.

Further, the measurement instrument did not possess internal validity as measured by the authors of the instrument. Internal validity was examined in the original study (Saunders & Holahan, 1997) by comparing subarea scores with each other and to the overall scores. Correlations between subareas were quite low in the original study, but between each subarea and the overall score they were high leading the authors to suggest that the instrument acted as a good diagnostic tool. In the present study, similar results were not found. Though correlations between each subarea and overall performance were high in both the pretest (r = .876 - .928) and posttest (r = .834 - .974), they were also high between subareas in both the pretest (r = .651 -. 896) and posttest (r = .691 - .929). This indicates that the WBSEF may not serve its intended purpose with junior high students, as judges did not necessarily isolate subareas of performance that may have needed more attention than others.

Practice Attitude

Students in all treatment groups seemed to enjoy the process of participating in the practice sessions. The processes in which students were involved utilized different practice strategies than they had regularly employed during practice sessions prior to the present study. It was also unlike experiences they have in their regular band classes. This variance from their normal routine may have led to high attitude scores.

Practice attitude scores did not differ between treatment groups during any week of the study. Not only did the groups not differ among each other, they did not change

their attitudes from week to week within their own group. Students do not find alternative practice procedures to be less auspicious than "normal practice" as seen by the fact that there were no differences between groups that received none of the three treatments and those that were subjected to treatment conditions. This implies that students may not be averse to using alternative practice strategies during their own practicing routines as they feel it could be both enjoyable and effective.

Conclusions

Based on the results and discussion of this investigation, the following conclusions are presented:

1. Type of modeling condition (modeling, no modeling) had a significant effect on music performance scores. The scores in the Model group were higher than scores for the No Model group for tone, technique/articulation, rhythmic accuracy, tempo, interpretation, and overall performance. Type of modeling condition had no effect of intonation, melodic accuracy, or practice attitude.
2. Type of self-listening condition (self-listening, no self-listening) had no significant effect on any dependent measure.
3. Type of self-evaluation condition (self-evaluation, no self-evaluation) had no significant effect on any dependent measure.
4. The test condition (pretest, posttest) had a significant effect on all music performance subareas and overall performance, but no significant effect on practice attitude.

5. The interaction of test x modeling x self-evaluation on music performance produced a significant effect. Gain scores of students in the Model/Self-Evaluation condition were higher than those in the No Model/Self-Evaluation condition for tone, melodic accuracy, rhythmic accuracy, interpretation, and overall performance, but not for intonation, technique/articulation, tempo, or practice attitude. Scores of students in the Model/No Self-Evaluation condition were not different than those in the no Model/No Self-Evaluation condition for any dependent variable.
6. There were no significant effects for the interaction of test x modeling x self-listening. Gain scores among all groups did not differ for any dependent variable.
7. There were no significant effects for the interaction of self-evaluation and self-listening. Gain scores among all groups did not differ for any dependent variable.
8. There were no significant effects for the interaction of modeling, self-evaluation, and self-listening. Gain scores among all groups did not differ for any dependent variable.

Implications for Music Education

A number of recent theories regarding music practice have been developed that recommend self-evaluation be included as part of an effective and efficient practice routine. In the present study; however, self-evaluation was found not to be effective for improving performance. Does this mean then that students should not learn to evaluate

their own performances? No. The definition of self-evaluation used in the study was very limited and most likely did not provide students the opportunity for proper assessment of their achievement. Furthermore, other studies (Aitchison, 1995; Davis, 1981; Sparks, 1990) have indicated that self-evaluation has positive effects on music learning.

When self-evaluation is enacted with junior high instrumentalists, a model recording should always be included. Results of the present study indicated that for students receiving self-evaluation treatment, those who listened to a model did significantly better than those who did not for the areas of tone, melodic accuracy, rhythmic accuracy, interpretation, and overall performance. This was not the case for students who did not self-evaluate.

Having stated that listening to a model is an important aspect for improving music performance, it may not be necessary at all times. For students who did not self-evaluate, those who listened to a model did no different for those who not did except in the areas of technique/articulation and tempo. Listening to a model to improve other areas of performance may not be necessary when not self-evaluating. Techniques other than those examined in this study may be more appropriate for helping students improve their tone, melodic and rhythmic accuracy, and interpretation. Teachers should instruct students to identify certain areas in need of improvement and then guide them to appropriate techniques to be used to improve them.

One particular finding by Aitchison (1997) was that students could improve the accuracy of their self-evaluation with practice. The present study showed that students

were unable to improve their performance through only self-evaluation. These findings suggest that students are able to diagnose their strengths and weaknesses, but may be unable to prescribe solutions that would assist in improving their ability. The development of diagnostic and prescriptive skills should constitute a greater portion of the junior high band curriculum. Often junior high and middle school teachers are charged with the goal of preparing their students for upcoming group performances. Classroom time is spent focusing on this endeavor often to the detriment of the individual learner. Teachers should develop strategies to incorporate into their rehearsals/classes that can assist in both developing quality group performances and the development of individual student musicians. Providing time, opportunity, and the proper structure for students to formally reflect on specific individual or group performances seems appropriate for junior high students. Feedback by the teacher regarding assessments followed by discussion of specific remedies to problems that arise may assist in developing students' individual assessment ability. A number of resources (see ASBDA, 1997; Kvet and Tweed, 1997; Labuta, 1997) have been developed in recent years that provide teachers with materials and strategies for implementing these types of strategies.

Social learning theorists and music educators alike have long promoted the advantages of modeling. The present study reaffirms the idea that listening to a model can be a useful approach for improving some aspects of performance. Listening to models should be an integral aspect of musicianship development and students should learn to incorporate the process into their lives. To do so, a number of circumstances are necessary. First, students must be provided opportunities to both live and recorded

models. Teachers themselves should initiate this by performing music being studied on a wide variety of instruments for their students. Students should also be provided opportunities to listen to live performances of individuals who play the same instrument as they do. Not only should they hear professional models, but also students who are only a few years older and are somewhat more advanced in their performing ability. Further, students can also model for each other. If an individual student plays a particular part in a piece well, that student can act as the model for others.

Students should also be led to find recordings of music they are performing. A number of publishing companies are developing method books and music that include recorded performances of the music. When choosing method books, teachers should consider those that include recordings and provide students with the opportunity to invest in the accompanying tape or compact disc and be required to listen to them. Perhaps publishers will eventually make these recordings available over the internet in downloadable files. Some method books provide one general recording for all instruments, using only electronic instruments to generate the sound, while other method books provide an authentic recording of each instrument. Green (1990) suggests that the timbre of the model should match the instrument being used, thus lending credence to the use of recordings using non-electronic sounds. Teachers often direct students to discover professional artists whose recordings demonstrate a high quality of musicianship and who provide truly ideal models for students to emulate. Although this may have motivational, social, and other benefits for the student, there is presently no research that suggests that doing this improves music perfomance.

Recommendations for Further Research

Based on the findings and discussions of this study, the following recommendations for further study are made:

1. Higher interjudge agreement for performance subarea scores should be obtained. Though adequate agreement was met, scores in the .85 - 1.00 range are preferred. More in-depth training of judges could affect this result.
2. This study was limited to a nine-week time period. Given that self-evaluation accuracy has been shown to increase over time, an increase in the length of the study may be appropriate.
3. Follow-up qualitative data (specifically subject interviews) might help to determine the influence of the various practice strategies on the music performance and practice attitude variables. The data could clarify or expand certain conclusions.
4. There was no modeling x self-evaluation x self-listening x time effect. The sample may need to be enlarged to examine this more intricate interaction, as there were only 10 or 11 subjects in the cells for the four-way interaction.
5. The development of self-evaluation music performance skills needs to be explored further. For instance, how do students develop the skills necessary to evaluate their performances, a necessary skill for all musicians? From this study, it seems that junior high school students need a model recording present during self-evaluation to assist them in improving their performance. At what point is this model no longer necessary? University students have had a difficult time accurately evaluating their own performances. Is this possible because they have not been provided a proper

model? Obviously, the development of audiation skills would assist in diminishing the constant use of a model, but is the presence of a model ever unnecessary?

6. Can an effective method for teaching self-evaluation be developed and implemented in a junior high band class?
7. Most secondary music classes today are structured around large group performance. Will self-evaluation, self-listening, listening to a model, and combinations of these variables in a large group ensemble setting assist the group as a whole in improving their performance?
8. How do self-evaluation abilities affect participation in music as individuals progress through their lives? Is it truly something that must be taught so adults will participate independently? Then, can amateur adult musicians self- evaluate as effectively as professionals?
9. The WBSEF was not an effective instrument for evaluating junior high students' performances. Is there an instrument available that is effective, and if not can one be developed and implemented?

APPENDIX A

LETTER TO PARENTS

October 22, 1999

Dear Emily Gray Band Students and Parents,

As a result of having one of the finest band programs in the Tucson area, you (your child) are being invited to participate in a joint project between your school and the University of Arizona. This project will attempt to examine the effects that different practicing have on musical performance.

Your (your child's) band director and I have talked extensively about this project and we feel it is a very worthwhile endeavor for the students at Emily Gray. The results may be able to help us understand better ways to learn new music.

Information regarding the activities you (your child) will be able participate in are described in the enclosed consent form. Should you (your child) choose to participate in the study, the form needs to be signed by both of you and returned to Ms. White by November 19, 1999.

Should you have any questions regarding this project, feel free to call me at my office in the School of Music at the University of Arizona (520) 626-7179. You may also contact me by e-mail at mphewitt@u.arizona.edu Thank you!

Sincerely,

Michael P. Hewitt, M. M.

Julia White, M. M.

APPENDIX B

ADAPTED WBSEF

Evaluating My Performance **Name** _____

Tone
(check only 1)

- ❑ My tone sounded like a very good player on **all** notes (high, low, and middle).
- ❑ My tone sounded like a very good player on **most** notes, but it was distorted sometimes.
- ❑ My tone was made with **some** mistakes (sometimes not enough air, not a full sound, or blasting).
- ❑ My tone was made with **major** mistakes (always not enough air, not a full sound, always blasting).
- ❑ My tone does not sound like the instrument should.

In Tune
(check only 1)

- ❑ I played in tune on all notes of my instrument (high, low, and middle).
- ❑ I played in tune on almost all notes, but on some individual notes I messed up and didn't fix.
- ❑ I played in tune generally but I did mess up a lot. I didn't fix the notes like I should have.
- ❑ I played in tune at a basic level, but I had a lot of problems and didn't try to fix them.
- ❑ I was always out of tune.

Pitches (such as Bb, C, F#, etc.)
(Check only 1)

- ❑ I played all pitches correctly.
- ❑ I played almost all pitches correctly.
- ❑ I played many pitches correctly.
- ❑ I played a lot of wrong notes.
- ❑ I played mostly wrong notes all the time.

Rhythms (such as eighth notes and rests)
(Check only 1)

- ❑ I played all rhythms correct all the way through the piece.
- ❑ I played mostly correct rhythms, but made small mistakes on some of them.
- ❑ I played about half of the rhythms correct, and was kind of close with those I missed.
- ❑ I played many rhythms wrong
- ❑ I played nearly all rhythms wrong.

Tempo
(Check only 1)

- ❑ I played at the correct speed through each number. I did not slow down or speed up except when I was supposed to.
- ❑ I was close to the correct speeds.
- ❑ Sometimes I played a little fast or a little slow, but when I did I kept it steady.
- ❑ I rushed or slowed down a lot of the time, but not all the time.
- ❑ I was always too slow or too fast.

TURN OVER THE PAGE AND FINISH

Dynamics and Phrasing
(Check only 1)

- ❑ I played all dynamics that were written and my phrases were played very musically.
- ❑ I played most of the dynamics that were written, but my phrases were not always played as musically as they should have been.
- ❑ I played with just OK dynamics and phrasing.
- ❑ I played only a few of the dynamics and with hardly any musicality.
- ❑ I played none of the dynamics and did not even think about phrasing.

Check ALL of these that you THINK you did!
(check **all** of these you think you did)

- ❑ I played with good tonguing.
- ❑ I played all the slurs that were marked in the music.
- ❑ I played all of the accents in the music.
- ❑ I played the articulations in the correct style.
- ❑ I played staccato and legato when I was supposed to.

APPENDIX C

PRACTICE ATTITUDE QUESTONNAIRE (PAQ)

Practice Attitude Questionnaire **Name** _____

Place an X in the box that matches the way you feel about each statement.	Strongly agree	Agree	Disagree	Strongly disagree
Rehearsing like this is enjoyable to me.				
This method of practicing stops me from getting better.				
Practicing this way is important to me.				
Practicing would always be fun if I did it this each time.				
Rehearsing like this all the time would be torture.				
I would recommend this way of practicin others.				
I do **not** enjoy practicing this way.				
I feel that practicing this way helps me improve faster.				

APPENDIX D

UNIVARIATE EFFECTS FOR MUSIC PERFORMANCE

Source	Measure	SS	df	MS	F
Test	Tone	2.628	1	2.628	16.863
	Intonation	3.121	1	3.121	20.056
	Tech/Artic	10.264	1	10.264	36.524
	Melodic Acc.	11.494	1	11.494	71.351
	Rhythmic Acc.	10.592	1	10.592	55.068
	Tempo	15.374	1	15.374	77.187
	Interpretation	10.202	1	10.202	64.702
	Overall	414.179	1	414.179	76.734
Test x Self-Evaluation	Tone	.277	1	.277	1.774
	Intonation	.003	1	.003	.024
	Tech/Artic	.038	1	.038	.134
	Melodic Acc.	.044	1	.044	.276
	Rhythmic Acc.	.141	1	.141	.734
	Tempo	.007	1	.007	.034
	Interpretation	.045	1	.045	.288
	Overall	.034	1	.034	.006
Test x Self-Listening	Tone	.001	1	.001	.006
	Intonation	.000	1	.000	.000
	Tech/Artic	.830	1	.830	2.955
	Melodic Acc.	.007	1	.007	.045
	Rhythmic Acc.	.412	1	.412	2.144
	Tempo	.000	1	.000	.003
	Interpretation	.136	1	.136	.861
	Overall	3.834	1	3.834	.710
Test x Modeling	Tone	.675	1	.675	4.330*
	Intonation	.419	1	.419	2.690
	Tech/Artic	1.147	1	1.147	4.083*
	Melodic Acc.	.451	1	.451	2.797
	Rhythmic Acc.	2.398	1	2.398	12.467***
	Tempo	2.316	1	2.316	11.628***
	Interpretation	.632	1	.632	4.007*
	Overall	50.073	1	50.07	9.277**
Test x Self-Evaluation x Self-Listening	Tone	.002	1	.002	.014
	Intonation	.010	1	.010	.063
	Tech/Artic	.061	1	.061	.217
	Melodic Acc.	.101	1	.101	.626
	Rhythmic Acc.	.000	1	.000	.000
	Tempo	.017	1	.017	.086

	Interpretation	.230	1	.230	1.460
	Overall	.036	1	.036	.007
Test x Self-Evaluation x Modeling	Tone	.628	1	.628	4.029*
	Intonation	.042	1	.042	.273
	Tech/Artic	.305	1	.305	1.086
	Melodic Acc.	.643	1	.643	3.994*
	Rhythmic Acc.	1.248	1	1.248	6.489*
	Tempo	.515	1	.515	2.583
	Interpretation	1.504	1	1.504	9.536**
	Overall	29.309	1	29.309	5.430*
Test x Self-Listening x Modeling	Tone	.007	1	.007	.046
	Intonation	.335	1	.335	2.156
	Tech/Artic	.002	1	.002	.008
	Melodic Acc.	.006	1	.006	.035
	Rhythmic Acc.	.044	1	.044	.231
	Tempo	.034	1	.034	.170
	Interpretation	.010	1	.010	.066
	Overall	1.168	1	1.168	.216
Test x Self-Listening x Modeling x Self-Eval.	Tone	.000	1	.000	.000
	Intonation	.002	1	.002	.014
	Tech/Artic	.032	1	.032	.114
	Melodic Acc.	.484	1	.484	3.003
	Rhythmic Acc.	180	1	180	.935
	Tempo	.158	1	.158	.793
	Interpretation	.045	1	.045	.288
	Overall	3.796	1	3.796	.703
Error	Tone	11.534	74	.156	
	Intonation	11.516	74	.156	
	Tech/Artic	20.795	74	.281	
	Melodic Acc.	11.921	74	.161	
	Rhythmic Acc.	14.234	74	.192	
	Tempo	14.739	74	.199	
	Interpretation	11.668	74	.158	
	Overall	339.423	74	5.398	

APPENDIX E

MUSIC PERFORMANCE PRETEST/POSTTEST MEANS AND STANDARD DEVIATIONS FOR MODEL VS. NO MODEL

	Model (n = 42)				No Model (n = 40)			
	Pretest		Posttest		Pretest		Posttest	
Subarea	M	SD	M	SD	M	SD	M	SD
Tone	2.69	.65	3.07	.74	2.87	.69	2.99	.74
Intonation	2.49	.62	2.87	.64	2.72	.68	2.72	.68
Technique/Artic.	2.38	.89	3.06	1.00	2.36	.96	2.69	1.13
Melodic Accuracy	2.59	.69	3.22	.86	2.58	.78	3.01	.84
Rhythmic Accuracy	2.48	.71	3.24	.94	2.45	.76	2.72	1.04
Tempo	2.05	.65	2.90	1.00	2.10	.66	2.48	.94
Interpretation	2.20	.63	2.83	.88	2.15	.76	2.53	.91
Overall	16.88	4.22	21.17	5.59	17.05	4.89	19.13	5.83

APPENDIX F

MUSIC PERFORMANCE PRETEST/POSTTEST MEANS AND STANDARD DEVIATIONS FOR SELF-LISTENING VS. NO SELF-LISTENING

	Self-Listening (n = 42)				No Self-Listening (n = 40)			
	Pretest		Posttest		Pretest		Posttest	
Subarea	M	SD	M	SD	M	SD	M	SD
Tone	2.80	.65	3.06	.66	2.75	.70	3.01	.81
Intonation	2.55	.65	2.83	.64	2.48	.69	2.76	.68
Technique/Artic.	2.34	.81	2.99	1.04	2.40	1.03	2.76	1.12
Melodic Accuracy	2.69	.69	3.24	.72	2.48	.76	2.99	.97
Rhythmic Accuracy	2.44	.67	3.06	.95	2.49	.80	2.90	1.09
Tempo	2.10	.66	2.72	.98	2.04	.64	2.66	1.00
Interpretation	2.17	.66	2.74	.80	2.18	.74	2.62	1.01
Overall	17.10	4.27	20.63	5.36	16.82	4.83	19.69	6.19

APPENDIX G

MUSIC PERFORMANCE PRETEST/POSTTEST MEANS AND STANDARD DEVIATIONS FOR SELF-EVALUATION vs. NO SELF-EVALUATION

	Self-Evaluation (n = 41)				No Self-Evaluation (n = 41)			
	Pretest		Posttest		Pretest		Posttest	
Subarea	M	SD	M	SD	M	SD	M	SD
Tone	2.67	.60	3.02	.70	2.88	.73	3.05	.77
Intonation	2.50	.57	2.77	.67	2.53	.75	2.81	.65
Technique/Artic.	2.27	.77	2.75	1.02	2.47	1.05	3.01	1.13
Melodic Accuracy	2.59	.66	3.09	.91	2.59	.80	3.15	.81
Rhythmic Accuracy	2.39	.63	2.85	1.03	2.54	.83	3.11	1.00
Tempo	1.98	.53	2.59	.95	2.79	1.02	2.29	.76
Interpretation	2.06	.60	2.60	.92	2.29	.76	2.76	.89
Overall	16.46	3.79	19.67	5.81	17.46	5.16	20.67	5.75

APPENDIX H

MUSIC PERFORMANCE PRETEST/POSTTEST MEANS (AND STANDARD DEVIATIONS) FOR MODEL x SELF-EVALUATION INTERACTION

	Model/SE (n = 20)		Model/No SE (n = 21)		No Model/SE (n = 20)		No Model/No SE (n = 21)	
Subarea	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Tone	2.57 (.56)	3.18 (.75)	2.86 (.68)	3.03 (.69)	2.83 (.59)	2.92 (.59)	2.84 (.82)	3.00 (.90)
Intonation	2.53 (.56)	2.97 (.67)	2.48 (.69)	2.83 (.56)	2.50 (.61)	2.63 (.61)	2.56 (.82)	2.75 (.77)
Tech/Artic	2.28 (.76)	3.05 (1.07)	2.49 (1.02)	3.11 (.96)	2.27 (.81)	2.48 (.93)	2.43 (1.08)	2.86 (1.28)
Melodic Accuracy	2.58 (.68)	3.33 (1.05)	2.62 (.71)	3.16 (.66)	2.62 (.66)	2.88 (.71)	2.52 (.89)	3.10 (.95)
Rhythmic Accuracy	2.45 (.63)	3.37 (.90)	2.54 (.81)	3.17 (.96)	2.35 (.65)	2.38 (.93)	2.52 (.85)	3.00 (1.06)
Tempo	2.03 (.58)	3.05 (.97)	2.08 (.73)	2.83 (1.00)	1.95 (.49)	2.20 (.71)	2.22 (.77)	2.68 (1.08)
Interp	2.07 (.57)	2.97 (.96)	2.35 (.67)	2.75 (.80)	2.07 (.65)	2.28 (.76)	2.21 (.85)	2.71 (1.00)
Overall	16.52 (3.76)	21.92 (6.00)	17.41 (4.68)	20.87 (5.05)	16.58 (3.94)	17.78 (4.83)	17.30 (5.68)	20.10 (6.59)

Note. SE = Self-Evaluation.

APPENDIX I

MUSIC PERFORMANCE PRETEST/POSTTEST MEANS (AND STANDARD DEVIATIONS) FOR MODEL x SELF-LISTENING INTERACTION

	Model/SL (n = 22)		Model/No SL (n = 20)		No Model/SL (n = 20)		No Model/No SL (n = 20)	
Subarea	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Tone	2.77 (.66)	3.14 (.65)	2.60 (.64)	3.00 (.84)	2.83 (.64)	2.97 (.67)	2.90 (.75)	3.02 (.81)
Intonation	2.59 (.60)	2.88 (.67)	2.38 (.63)	2.85 (.62)	2.50 (.71)	2.77 (.62)	2.58 (.75)	2.67 (.75)
Tech/Artic	2.42 (.80)	3.23 (1.00)	2.33 (.99)	2.87 (1.00)	2.25 (.83)	2.73 (1.04)	2.47 (1.08)	2.65 (1.24)
Melodic Accuracy	2.77 (.62)	3.41 (.72)	2.38 (.71)	3.02 (.98)	2.60 (.76)	3.05 (.69)	2.57 (.82)	2.97 (.98)
Rhythmic Accuracy	2.55 (.66)	3.36 (.91)	2.42 (.79)	3.10 (.97)	2.33 (.69)	2.73 (.90)	2.57 (.83)	2.70 (1.18)
Tempo	2.23 (.64)	3.05 (.96)	1.85 (.62)	2.73 (1.04)	1.97 (.68)	2.37 (.90)	2.23 (.62)	2.58 (.98)
Interp	2.24 (.57)	2.94 (.81)	2.15 (.70)	2.70 (.96)	2.10 (.75)	2.52 (.74)	2.20 (.78)	2.53 (1.07)
Overall	17.58 (4.01)	22.00 (5.33)	16.12 (4.41)	20.27 (5.86)	16.58 (4.59)	19.13 (5.11)	17.52 (5.24)	19.12 (6.61)

Note. SL= Self-Listening.

APPENDIX J

MUSIC PERFORMANCE PRETEST/POSTTEST MEANS (AND STANDARD DEVIATIONS) FOR SELF-LISTENING x SELF-EVALUATION INTERACTION

	SL/SE (n = 21)		SL/No SE (n = 21)		No SL/SE (n = 20)		No SL/No SE (n = 20)	
Subarea	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Tone	2.73 (.49)	3.08 (.67)	2.87 (.78)	3.03 (.66)	2.62 (.70)	2.95 (.74)	2.88 (.69)	3.07 (.90)
Intonation	2.52 (.50)	2.81 (.65)	2.57 (.78)	2.84 (.66)	2.48 (.65)	2.73 (.71)	2.48 (.75)	2.78 (.67)
Tech/Artic	2.37 (.57)	2.95 (.85)	2.32 (1.01)	3.03 (1.22)	2.17 (.94)	2.53 (1.17)	2.63 (1.08)	2.98 (1.05)
Melodic Accuracy	2.76 (.52)	3.24 (.70)	2.62 (.83)	3.24 (.76)	2.40 (.75)	2.93 (1.08)	2.55 (.78)	3.05 (.86)
Rhythmic Accuracy	2.43 (.42)	3.00 (.93)	2.46 (.87)	3.13 (.99)	2.35 (.81)	2.70 (1.13)	2.63 (.79)	3.10 (1.03)
Tempo	2.02 (.45)	2.65 (.91)	2.19 (.83)	2.79 (1.07)	1.95 (.60)	2.53 (1.02)	2.13 (.68)	2.78 (.99)
Interp	2.08 (.50)	2.76 (.79)	2.27 (.78)	2.71 (.83)	2.03 (.70)	2.43 (1.04)	2.32 (.76)	2.80 (.96)
Overall	16.90 (2.79)	20.49 (5.12)	17.30 (5.44)	20.78 (5.71)	16.00 (4.66)	18.82 (6.48)	17.63 (4.99)	20.57 (5.93)

Note. SE = Self-Evaluation. SL = Self-Listening.

APPENDIX K

MUSIC PERFORMANCE PRETEST/POSTTEST MEANS (AND STANDARD DEVIATIONS) FOR MODEL x SELF-LISTENING x SELF-EVALUATION INTERACTION

	Mod/SL/SE (n = 11)		Mod/SL/No SE (n = 11)		Mod/No SL/SE (n = 10)		Mod/No SL/ No SE (n = 10)	
Subarea	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Tone	2.64 (.57)	3.21 (.79)	2.91 (.75)	3.06 (.49)	2.40 (.60)	3.00 (.83)	2.80 (.63)	3.00 (.89)
Intonation	2.61 (.51)	2.94 (.77)	2.58 (.70)	2.82 (.58)	2.40 (.60)	2.87 (.69)	2.37 (.69)	2.83 (.57)
Tech/Artic	2.52 (.57)	3.36 (.91)	2.33 (1.01)	3.09 (1.11)	2.00 (.85)	2.60 (1.13)	2.67 (1.05)	3.13 (.82)
Melodic Accuracy	2.73 (.51)	3.52 (.79)	2.82 (.74)	3.30 (.66)	2.37 (.81)	3.03 (1.26)	2.40 (.64)	3.00 (.65)
Rhythmic Accuracy	2.48 (.50)	3.48 (.85)	2.61 (.80)	3.24 (.99)	2.37 (.76)	3.10 (1.02)	2.47 (.85)	3.10 (.97)
Tempo	2.15 (.52)	3.15 (.94)	2.30 (.75)	2.94 (1.02)	1.87 (.61)	2.77 (1.11)	1.83 (.65)	2.70 (1.02)
Interp	2.12 (.52)	3.15 (.83)	2.36 (.60)	2.73 (.77)	1.97 (.62)	2.63 (1.09)	2.33 (.77)	2.77 (.86)
Overall	17.24 (3.04)	22.82 (5.59)	17.91 (4.93)	21.18 (5.19)	15.37 (4.34)	20.00 (6.77)	16.87 (4.58)	20.53 (5.14)

Note. SE = Self-Evaluation. SL = Self-Listening. Mod = Model.

APPENDIX K
(Continued)

MUSIC PERFORMANCE PRETEST/POSTTEST MEANS (AND STANDARD DEVIATIONS) FOR MODEL x SELF-LISTENING x SELF-EVALUATION INTERACTION

	No Mod/SL/SE (n = 10)		No Mod/SL/No SE (n = 10)		No Mod/ No SL/SE (n = 10)		No Mod/ No SL/No SE (n = 10)	
Subarea	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Tone	2.83 (.39)	2.93 (.52)	2.83 (.85)	3.00 (.83)	2.83 (.76)	2.90 (.69)	2.97 (.78)	3.13 (.95)
Intonation	2.43 (.50)	2.67 (.47)	2.57 (.89)	2.87 (.76)	2.57 (.72)	2.60 (.75)	2.60 (.81)	2.73 (.78)
Tech/Artic	2.20 (.55)	2.50 (.48)	2.30 (1.07)	2.97 (1.39)	2.33 (1.04)	2.47 (1.26)	2.60 (1.16)	2.83 (1.27)
Melodic Accuracy	2.80 (.55)	2.93 (.44)	2.40 (.91)	3.17 (.89)	2.43 (.74)	2.83 (.93)	2.70 (.91)	3.10 (1.07)
Rhythmic Accuracy	2.37 (.33)	2.47 (.71)	2.30 (.95)	3.00 (1.03)	2.33 (.89)	2.30 (1.14)	2.80 (.74)	3.10 (1.14)
Tempo	1.87 (.32)	2.10 (.47)	2.07 (.93)	2.63 (1.16)	2.03 (.62)	2.30 (.91)	2.43 (.59)	2.87 (1.01)
Interp	2.03 (.51)	2.33 (.47)	2.17 (.96)	2.70 (.92)	2.10 (.80)	2.23 (.99)	2.30 (.79)	2.83 (1.10)
Overall	16.53 (2.60)	17.93 (3.11)	16.63 (6.14)	20.33 (6.49)	16.63 (5.10)	17.63 (6.30)	18.40 (5.50)	20.60 (6.91)

Note. SE = Self-Evaluation. SL = Self-Listening. Mod = Model.

References

Aitchison, R. A. (1995). The effects of self-evaluation techniques on the musical performance, self-evaluation accuracy, motivation, and self-esteem of middle school instrumental music students. (Doctoral dissertation, University of Iowa). Dissertation Abstracts International, 56-10A, 3875.

American School Band Directors Association (ASBDA). (1997). The new ASBDA curriculum guide. Miami, FL: Belwin-Mills.

Anderson, J. N. (1981). Effects of tape-recorded aural models on sight-reading and performance skills. Journal of Research in Music Education, 29, 23-30.

Bandura, A. (1977). Social learning theory. Englewood Cliffs, N. J.: Prentice-Hall.

Bergee, M. J. (1988). Use of an objectively constructed rating scale for the evaluation of brass juries: A criterion-related study. Missouri Journal of Research in Music Education, 5(5), 6-25.

Bergee, M. J. (1989). An objectively constructed rating scale for euphonium and tuba music performance. Dialogue in Instrumental Music Education, 13, 65-86.

Bergee, M. J. (1993). A comparison of faculty, peer, and self-evaluation of applied brass jury performances. Journal of Research in Music Education, 41, 19-27.

Biggs, J. & Moore, P. J. (1993). The process of learning. New York: Prentice Hall.

Boyle, J. D., & Radocy, R. E. (1987). Measurement and evaluation of musical experiences. New York: Schirmer.

Bundy, O. R. (1987). Instrumentalists' perception of their performance as measured by detection of pitch and rhythm errors under live and recorded conditions. (Doctoral dissertation, Pennsylvania State University). Dissertation Abstracts International, 48-10A, 2567.

Byo, J. L. (1990). Recognition of intensity contrasts in gestures of beginning conductors. Journal of Research in Music Education, 38, 157-163.

Byo, J. L. & Brooks, R. (1994). A comparison of junior high musicians' and music educators' performance evaluations of instrumental music. Contributions to Music Education, 21, 26-38.

Colwell, C. M. (1995). Effect of teaching setting and self evaluation on teacher intensity behaviors. Journal of Research in Music Education, 43, 6-21.

Cutietta, R. A. (1992). The measurement of attitudes and preferences in music education. In R. Colwell (Ed.). Handbook of research on music teaching and learning (pp. 295-309). New York: Schirmer.

Davidson, L., & Scripp, L. (1992). Surveying the coordinates of cognitive skills in music. In R. Colwell (Ed.), Handbook of research on music teaching and learning (pp. 392-413). New York: Schirmer.

Davis, L. P. (1981). The effects of structured singing activities and evaluation practice on elementary band students' instrumental achievement. (Doctoral dissertation, Ohio State University). Dissertation Abstracts International, 42-07A, 3051.

Dickey, M. R. (1991). A comparison of verbal instruction and nonverbal teacher-student modeling in instrumental ensembles. Journal of Research in Music Education, 39, pp. 132-142.

Dickey, M. R. (1992). A review of research on modeling in music teaching and learning. Bulletin of the Council for Research in Music Education, 113, 27-40.

Ericsson, K. A. (1996). The acquisition of expert performance: An introduction to some of the issues. In K. A. Ericsson (Ed.), The road to excellence: The acquisition of expert performance in the arts and sciences, sports and games (pp. 1-50). Mahwah, NJ: LEA.

Ericsson, K. A. (1997). Deliberate practice and the acquisition of expert performance: An overview. In H. Jorgensen & A. C. Lehmann (Eds.), Does practice make perfect? Current theory and research on instrumental music practice (pp. 9-52). Norway: Norges musikkhøgskole.

Ericsson, K. A., Krampe, R. T., & Tesch-Romer, C. (1993). The role of deliberate practice in the acquisition of expert performance, Psychological Review 100, 363-406.

Green, G. A. (1990). The effect of vocal modeling on pitch-matching accuracy of elementary schoolchildren. Journal of Research in Music Education, 38, 225-232.

Hallam, S. (1997). What do we know about practicing? Toward a model synthesizing the research literature. In H. Jorgensen & A. C. Lehmann (Eds.), Does practice make perfect? Current theory and research on instrumental music practice (pp. 179-231). Norway: Norges musikkhøgskole.

Hand, D. J., & Taylor, C. C. (1987). Multivariate analysis of variance and repeated measures. London: Chapman and Hall.

Henerson, M. E., Morris, L. L., and Fitz-Gibbon, C. T. (1987). How to measure attitudes. Newbury Park: Sage Publications.

Johnston, H. (1993). The use of video self-assessment, peer-assessment, and instructor feedback in evaluating conducting skills in music student teachers. British Journal of Music Education, 10, 57-63.

Jørgensen, H. (1995). Teaching and learning strategies in instrumental practice: A report on research in progress. In J. A. Taylor (Ed.), Transatlantic roads of music education: World views (pp. 47-51). Tallahassee, FL: Center for Music Research.

Kepner, C. B. (1986). The effect of performance familiarity, listening condition, and type of performance error on correctness of performance error detection by 50 high school instrumentalists as explained through a sensory blocking theory. (Doctoral dissertation, Kent State University, Kent, OH). Dissertation Abstracts International, 47-05A, 1643.

Kostka, M. J. (1997). Effects of self-assessment and successive approximations on "knowing" and "valuing" selected keyboard skills. Journal of Research in Music Education, 45, 273-281.

Kvet, E. J., & Tweed, J. M. (Eds.). (1997). Strategies for teaching: Beginning and intermediate band. Reston, VA: Music Educators National Conference.

Labuta, J. A. (1997). Teaching musicianship in the high school band. (2nd ed.). Ft. Lauderdale, FL: Meredith Music Publications.

Linklater, F. (1997). Effects of audio- and videotape models on performance achievement of beginning clarinetists. Journal of Research in Music Education, 45(3). 402-414.

Madsen, C. K., Greer, R. D., & Madsen, C. H. (1975). Research in music behavior. New York: Teachers College Press.

Minke, A. (January, 1997). Conducting repeated measures analyses: Experimental design considerations. Paper presented at the annual meeting of the Southwest Educational Research Association, Austin, TX.

Neilson, J. (1956). Warm-up procedures for the cornetist. Instrumentalist, 11(3). 53-57.

O'Connell, T. S. (1967). How's your technique? Instrumentalist, 21(9), 46-52.

Ormrod, J. E. (1999). Human learning (3rd Ed.). Upper Saddle River, N. J.: Prentice Hall.

Puopolo, V. (1971). The development and experimental application of self-instructional practice materials for beginning instrumentalists. Journal of Research in Music Education, 19, 342-49.

Reely, R. (1990). Buzzing practice for beginners. Instrumentalist, 44(6), 46-49.

Rosenthal, R. K. (1984). Effects of guided model, model only, guide only, and practice only on performance. Journal of Research in Music Education, 32, 265-273.

Rosenthal, R. K. (1985). Improving teacher effectiveness through self-assessment: A case study. Update: The Applications of Research in Music Education, 3(2), 17-21.

Rosenthal, R. K., Wilson, M., Evans, M., & Greenwalt, L. (1988). Effects of different practice conditions on advanced instrumentalists' performance accuracy. Journal of Research in Music Education, 36, 250-257.

Sang, R. C. (1985). Path analysis of an instructional effectiveness model for instrumental music teaching. Bulletin of the Council for Research in Music Education, 87, 195-206.

Sang, R. C. (1987). A study of the relationship between instrumental music teachers' modeling skills and pupil performance behaviors. Bulletin of the Council for Research in Music Education, 91, 155-159.

Saunders, T. C., & Holahan, J. M. (1997). Criteria-specific rating scales in the evaluation of high school instrumental performance. Journal of Research in Music Education, 45, 259-272.

Shehan, P. K. (1985). Transfer of preference from taught to untaught pieces of non-western music genres. Journal of Research in Music Education, 33, 149-158.

Slavin, R. E. (1991). Educational psychology (3rd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Sloboda, J., Davidson, J. W., Howe, M. J. A., & Moore, D. G. (1996). The role of practice in the development of performing musicians. British Journal of Psychology, 87, 287-309.

Small, A. R., & McCachern, F. L. (1983). The effect of male and female vocal modeling on pitch-matching accuracy of first-grade children. Journal of Research in Music Education, 31, 227-233.

Sparks, G. E. (1990). The effect of self-evaluation on musical achievement, attentiveness, and attitudes of elementary school instrumental students. (Doctoral dissertation, Louisiana State University). Dissertation Abstracts International, 51-09A, 3009.

Tait, M. J. (1992). Teaching strategies and styles. In R. Colwell (Ed.), Handbook of research on music teaching and learning (pp. 525-534). Washington: Music Educators National Conference.

Van Develde, J. (1981). Doing double duty on the lip and waistline. Instrumentalist, 36(4), 54-58.

Watkins, J. G., & Farnum, S. E. (1962). The Watkins-Farnum performance scale (Form B). Milwaukee, WI: Hal Leonard Publishing.

Yarbrough, C. (1987). The relationship of behavioral self-assessment to the achievement of basic conducting skills. Journal of Research in Music Education, 35, 183-189.

Yarbrough, C., Wapnick, J., & Kelly, R. (1979). Effect of videotape feedback techniques on performance, verbalization, and attitude of beginning conductors. Journal of Research in Music Education, 27, 103-112.

Zimmerman, B. J. (1998). Developing self-fulfilling cycles of academic regulation: An analysis of exemplary instructional models. In D. H. Schunk, & B. J. Zimmerman (Eds.), Self-regulated learning: From teaching to self-reflective practice. New York: Guilford.

Zurcher, W. (1975). The effect of model-supportive practice on beginning brass instrumentalists. In C. Madsen, R. D. Greer, & C. H. Madsen (Eds.), Research in music behavior: Modifying music behavior in the classroom. New York: Teachers College Press.