Order Number 9426218

# Error detection abilities of conducting students under four modes of instrumental score study

Crowe, Don Raymond, D.M.A.

The University of Arizona, 1994



U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

# ERROR DETECTION ABILITIES OF CONDUCTING STUDENTS UNDER FOUR MODES OF INSTRUMENTAL SCORE STUDY

by

Don Raymond Crowe



A Dissertation Submitted to the Faculty of the

SCHOOL OF MUSIC

In Partial Fulfillment of the Requirements
For the Degree of

DOCTOR OF MUSICAL ARTS
WITH A MAJOR IN MUSIC EDUCATION

In the Graduate College

THE UNIVERSITY OF ARIZONA

1 9 9 4

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Final Examination Committee, we certify that we have read the dissertation prepared by Don Raymond Crowe

entitled ERROR DETECTION ABILITIES OF CONDUCTING STUDENTS UNDER FOUR MODES OF INSTRUMENTAL SCORE STUDY

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

| | |
|---|---|
| Steven K. Hedden | 4/29/94 |
| Steven K. Hedden | Date |
| John Fitch | 4/29/94 |
| John Fitch | Date |
| Nancy Ferguson | 4/29/94 |
| Nancy Ferguson | Date |
| | 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.

| | |
|---|---|
| Steven K Hedden | 4/29/94 |
| Dissertation Director | 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 acknowledgement 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 copyright holder.

SIGNED: Don R. Crowe

## ACKNOWLEDGEMENTS

The author would like to express his gratitude to his thesis supervisor, Dr. Steven Hedden, for his guidance and instruction, both formal and informal, from the beginning of doctoral study through the completion of this project. His support and encouragement have been of immeasurable help. Also to the other members of the thesis committee, Dr. John Fitch and Professor Nancy Ferguson, for their work as teachers and mentors in so many ways. Their availability to and interest in their students set high standards.

A great debt of gratitude is owed to the educators whose students took part in the study. Their willingness to supervise the project and encourage their students to complete the sessions made the study possible. Also, thanks are due the many friends and colleagues who provided support, encouragement and occasional advice throughout the project.

The author cannot express adequately his deep appreciation for his family's willingness, indeed eagerness, to support this effort in every possible way. Finally, to my wife, Linda, whose help, faith and inspiration were unending, goes all credit for this study. Without her, this would not have been possible.

# Dedication

This is dedicated to my wife and our family, without whom none of this would have been possible. Thank you for your faith in me and your willingness to support this effort in so many ways. This is more your project than mine.

TABLE OF CONTENTS

TABLE OF CONTENTS (CONTINUED)

# LIST OF TABLES

LIST OF TABLES (CONTINUED)

## ABSTRACT

This study investigated the effects of four score study styles--no score study, study with score alone, study with score and a correct aural example, and score study at the electronic keyboard--on the pitch and rhythm error detection abilities of beginning conducting students. Subjects were 30 members of undergraduate beginning conducting classes at three midwestern universities.

Four tests were developed, each having 31 four- to six-measure excerpts from band literature. Each excerpt contained only one error. Excerpts were grouped according to difficulty and assigned to tests in a modified random manner to facilitate equality of difficulty between sets. Within each test, excerpts were arranged in order of increasing difficulty and rescored to contain from one to eight parts.

A counterbalanced design was utilized featuring a Latin Square into which the four score study styles were entered. Over the course of four sessions subjects received all four styles and all four tests. The orders in which subjects received score study styles were assigned on a rotational basis. Each subject within a university received the tests in the same order, but this order varied between universities. Six Hypercard© (Atkinson, 1987-90) stacks were developed on a Macintosh LC computer for presentation of the tests, management of the study, and data collection.

Excerpts were played through MIDI keyboards using sampled wind instrument sounds.

Study with the score and a correct aural example was found to be significantly more effective than either study with the score alone or no study. No significant difference was found between score study at the keyboard and any other score study style.

There were significant differences in test scores attributable to the number of parts in examples. Generally, error detection became more difficult as the number of parts in examples increased.

There were no significant differences in test scores attributable to the order of presentation of score study styles, individual example sets, or groups/order of presentation of example sets. There were significant differences in mean score study time per session attributable to score study style, and in mean total time per session attributable to session number.

# Chapter I

# Introduction

## Introduction and Need For The Study

The ability to detect errors in performance is a basic skill for all conductors. If the director of an ensemble cannot readily discern when and where errors are being made, rehearsal time becomes less effective and the ensemble's level of performance suffers. Sidnell (1971) states that

> This skill is deemed a necessary behavioral competence of teachers of instrumental music. These teachers are faced with detecting and identifying errors during rehearsal class. Highly developed aural/visual skill, commonly referred to as score-reading ability, can have positive effect on the judicious use of available teaching time. (p. 85)

In a survey of music teachers, music consultants, and university music education faculty, Taebel (1980) identified and ranked 51 musical competencies for music teachers. Taebel states that "teachers considered aural skills to be the most important competency used in their teaching. Of the highest 11 competencies, 9 were aural skill competencies, which involved error detection or evaluation of tonal features rather than sightsinging or dictation." (p. 188) The two highest rated competencies were "Detects errors or problems in pitch or intonation" and "Detects errors or problems in rhythm performance" (p. 191). That error detection is regarded as a crucial skill is demonstrated by its inclusion in such standardized commercial tests as the <u>Aliferis-Stecklein Music Achievement Test, College Midpoint</u>

Level (1962), in one section of which subjects are presented music visually and aurally and required to determine whether the two match.

According to Larson (1977), "Ear training instruction in college music programs often attempts to develop aural discrimination abilities through experience in dictation and sightsinging, to the neglect of developing abilities to detect performance errors" (p. 264). Jones (1990) states that two types of programs for aural skill development exist, neither of which "provides sufficient experience for the conductor's development of error detection skills" (p. 2).

> The first may be described as a concentrated listening program devoted to aural recognition of traditional harmonic progressions and chord voicings. Listening examples are usually performed on a piano in chorale style, using 18th century harmonic progressions, while the student responds by notating the music heard. The second program emphasizes listening to and conducting appropriate instrumental ensembles. Emphasis is usually placed on conducting technique instead of the process of listening for and correcting errors. (p. 2)

Jones identifies three deficiencies in traditional music dictation classes:

> Musical examples seldom approximate music traditionally performed by school instrumental organizations. Students also expect no errors in the performance of examples and consequently have no need to listen for errors and no need for analysis of error types. A third deficiency concerns the type of instrument performing the musical examples. It is important for students who are developing error detection skills to hear musical examples performed by a representative instrumental ensemble. (pp. 2-3)

A number of studies have shown that this preparation is inadequate to prepare students for rehearsal and performance situations in which they must detect and correct errors in music which they may or may not have seen or heard before, while reading a score of up to 64 parts and dealing with many different timbres. The preponderant evidence of these studies, which have considered predictors, correlates, and training of error detection ability, seems to indicate that the skills developed in the traditional ear training approach described above do not transfer to this new situation (Gonzo, 1970; Simard, 1982; Larson, 1977; Boyer, 1974; Shaw, 1971; Brand and Burnsed, 1981; Doane, 1989).

Sidnell (1971) affirms that

> Experience in aural music theory classes is very often unrelated to the aural discrimination skills needed by instrumental teachers. Rarely, if ever, are future teachers asked to detect and identify errors produced in a music texture involving various instruments. Aural skills are developed in ear training classes but there is little evidence of transfer to the problems that confront the conductor. (p. 85)

Boyer (1974) found no significant relationship between error detection ability and music theory study. Shaw (1971) found no significant correlation between error detection ability and either sophomore music theory grades, piano study, or vocal study. Brand and Burnsed (1981) found that grades in theory, sightsinging, and ear training were not significant predictors of error detection skills; this suggested to the authors that the ability to detect music errors may exist independently of other music abilities.

Doane (1989) found that performance in aural skills classes was "not useful in predicting the development of an individual's error detection skills" (p. 13). Doane's results supported the view that "conductors' aural error detection skills seem to be developed independently of skills taught in traditional aural skills classes" (p. 14).

A number of investigations of error detection have produced results with implications for the present study, which focuses on the effects of score study styles on error detection ability. Collings (1973) investigated the effects of three error detection training techniques—no prestudy, prestudy followed by error examples, and prestudy followed by correct recorded examples—on the pitch error detection abilities of second semester undergraduate conducting students. Examples were six-measure segments of Baroque period works scored for five- and six-part brass ensemble in concert pitch. No significant difference was found.

Grunow (1980) utilized four- to sixteen-measure examples of two- to five-part brass, woodwind and string ensemble pieces in an investigation of the effects of four modes of score study on the development of "visual-aural discrimination skills." Subjects used the score alone, score plus a recorded example, recorded example alone, or no study. No score study style was found to be significantly better than any other, but all subjects improved significantly over the course of the study.

Hopkins (1991) investigated the effects of four modes of score study--score alone, score at the piano, sightsinging from the score, and score plus correct recorded example--on the abilities of undergraduate music students to detect and notate five errors of pitch and rhythm in complete choral works. Each subject studied a score for 50 minutes before being tested. Score study with a correct recorded example was found to be significantly better than score study at the piano. No other significant differences were found.

Ramsey (1979) investigated the effects of programmed instruction on the error detection abilities of upper level undergraduate students. Three programs of different lengths were developed with full band excerpts from standard band literature. All three programs resulted in significant gains in the tested abilities, with the longer programs better than the shorter one.

Deal (1983) compared the effects of Ramsey's long program with those of a computer-assisted program based on it. The examples for the computer-assisted program were limited to four voices, three timbres, and four measures by the technology available. Both programs resulted in significant gains in error detection ability.

Jones (1990) designed and tested a computer-assisted error detection instruction program for undergraduate and graduate instrumental music education students. Twenty full-band excerpts, four to eight measures in length, were presented in a printed score booklet. The computer program randomly selected the performance of each excerpt from one correct and four incorrect performances on tape. The incorrect performances included one randomly-inserted error of either pitch, rhythm, articulation or style, or ensemble. Jones found significant changes in mean test scores from pretest to posttest, between some of the ten sessions for errors of articulation/style and errors of ensemble, and in mean test scores from session to session.

Gruner (1993) investigated the effects of a computer-assisted training program on the pitch and rhythm error detection abilities of undergraduate instrumental music education majors. Excerpts were taken from existing band literature of varying difficulty and consisted of from two to five parts with synthesized wind instrument timbres. The excerpts were printed in a score booklet, and subjects heard examples and provided responses at the computer. The treatment group improved significantly in overall error detection ability, and both the treatment and control groups improved significantly in rhythm error detection. Neither showed significant improvement in pitch error detection. Gruner concluded that computer-assisted instruction in error detection skills using synthesized sound sources is viable.

Those who have investigated error detection have made a number of recommendations for research in this area. Collings (1973) recommended that further research in the training of error detection include rhythmic errors and ensembles of various sizes and combinations of instruments in written pitch. Grunow (1980) recommended that future studies utilize examples with one part and examples with more than five parts. Ramsey (1979) suggested that programs be developed which would include small ensembles. Deal (1983) recommended the use of more than four voices with key signatures, longer examples, computer-generated timbres, and combinations of such timbres. Sidnell (1968) recommended that further studies utilize a wider range of timbres and an increasing number of staves. Jones (1990) suggested enhanced audio quality through the use of digital recording and playback technology, as well as the use of solo and small ensemble examples. Gruner (1993) recommended that further research include correct versions of excerpts for comparison with error versions, on-screen full scores with errors highlighted for reinforcement after incorrect answers, and an expanded number of voices per excerpt.

According to Bailey (1992), "The best method to improve aural skills is to improve score preparation. Without an accurate expectation of the sound, a conductor cannot expect to detect errors." (p. 32) Score study styles whose effects on error detection ability have been investigated include sightsinging from the score, study of the score alone, study of the score with a correct recorded example, and study of such an example without the score. Two studies (Stwolinski, Faulconer, and Schwarzkopf, 1988; Hopkins, 1991) have investigated the effects of score study at the keyboard on error detection ability. These utilized simple piano works and complete choral works, respectively, as their materials. No study has attempted to compare the effects of score study at the keyboard with other methods of score study on the ability to detect errors in instrumental performance.

Three studies have incorporated computer-assisted instruction in error detection (Deal, 1983; Jones, 1990; Gruner, 1993). These were limited by the technology available. The first was limited in the number of parts (four), timbres (three), length of excerpts (two to four measures), and score type (a small grand staff in concert pitch, without key signatures). The second utilized a score booklet for subject responses and a random access audio device for presentation of materials. The use of this device resulted in an audio quality which proved distracting and tiring to the subjects. The third study also utilized a booklet for score presentation, and was limited as to the choice of timbres. None of these studies investigated the comparative effects of score study styles.

The technology is now available for the incorporation of a greater number of timbres, parts, and measures in excerpts presented on-screen in a computer-assisted format. The development of MIDI (Musical Instrument Digital Interface) technology also allows for score study at the electronic keyboard of excerpts presented in such a format, and for the presentation of digitally sampled timbres. Such sampled timbres are much closer to the original than are are those which are computer-generated and thus result in a more realistic performance.

The recommendations and limitations of the studies cited above demonstrate the need for a study of error detection which would compare score study at the keyboard with other score study styles and which would incorporate computer-assisted presentation of materials and scoring of subjects' responses, a variety of instrumental timbres with scores in written pitch, a varied and increasing number of parts in error examples, and the use of pre-planned errors of pitch and rhythm only in "clean" performances.

Purpose

It was the purpose of this study to investigate the effects of four different modes of score study on the abilities of undergraduate conducting students to detect errors of pitch and rhythm in four- to six-measure excerpts drawn from band literature. The study utilized a program written for the Macintosh computer as the medium for presenting scores and examples and for data collection. Examples used sampled sounds presented through MIDI (Musical Instrument Digital Interface) keyboards.

Under the "no prestudy" condition, the scores and error examples were presented to the subjects simultaneously.

Under the "score only" condition, subjects were allowed to study the score as presented on the computer screen as long as desired and then hear the error example. Under the "score plus correct example" condition subjects were presented the score and correct example simultaneously, and were then allowed to study each score as long as desired before hearing the error example. Finally, under the "score plus keyboard" condition the subjects were encouraged to play the score on the electronic keyboard provided before hearing the error examples.

The excerpts ranged from one part to eight parts of music in written pitch and with sampled timbres, and consisted of natural combinations of wind instruments. Four sets of excerpts, based on those used in a study by Ramsey (1979), were developed. There was only one error per excerpt. Utilizing a counterbalanced design, every subject in each of three intact groups received every mode of score study and every example set.

Problem

The problem of the study is this: Which of four styles of score study--no study, study with the score alone, study with the score and a correct example, and study with the score at an electronic keyboard--best prepares undergraduate conducting students to detect errors of pitch and rhythm in musical examples featuring an increasing number of parts and sampled wind instrument timbres?

In addressing this problem the following subproblems, stated as null hypotheses, were identified:

There will be no difference in test scores on the Tests in Error Detection attributable to:

$H_o1$: score study style.

$H_o2$: order of presentation of score study styles.

$H_o3$: example set/individual test.

$H_o4$: intact group/order of presentation of example sets.

$H_o5$: the number of parts in the items.

# Chapter II

## Related Literature

There have been a number of studies on the error detection abilities of music students. The investigators have considered both the relationship of error detection ability to other factors and the effects of specific methods of training on error detection ability.

These studies have involved a wide range of subjects, error types, excerpt types, presentation media, and environment. This disparity has resulted in a clouded picture of error detection ability and training and in many cases has produced contradictory conclusions.

### Studies of error detection ability and its relationship to other factors

Simard (1982) investigated the relationship between instruction in ear training and the location and correction of errors of pitch and rhythm in a melodic line. Subjects were enrolled in ear training courses and were given tests in sightsinging, melodic dictation, and error detection. He found that test scores on the melodic dictation and sightsinging tests were very good predictors of error detection ability, and that ear training grades, years of formal keyboard study and self-confidence also contributed significantly to such prediction. Error location was considerably easier than error correction, and more variables predicted ability in error location than in error correction.

Pagan (1970) developed a test to measure the abilities of basic musicianship students to form aural images of chords in a score. A test of 45 items, some with single chords, some with a series of three chords, was developed. The four-voice chords on a grand staff were briefly projected on a screen, and an aural version was then played. Subjects were asked to indicate whether the aural presentation was the same as or different from the visual presentation. Alterations consisted of the change of one note by a step or half-step or a change in chord position. Pagan concluded that keyboard study was significantly related to the tested ability, and that the test effectively measured the ability targeted.

Larson (1977) studied the relationships between melodic sightsinging, melodic dictation, and melodic pitch error detection among junior and senior music majors. Test scores on three researcher-constructed tests were compared. Twelve melodies were used in the error detection test. Larson found that the relationship between error detection and melodic dictation test scores was stronger than that between error detection and sightsinging test scores.

As part of their studies involving programmed instruction in score reading and rhythmic error detection respectively (discussed below), two investigators studied the relationships between error detection ability and other factors. Boyer (1974) investigated the relationships between error detection ability and music achievement, formal music study, keyboard study, music theory study, and perfect pitch. No significant relationships were discovered. Shaw (1971) found that subjects' number of years of private instrumental study, number of years in band, and score on the Math section of the Scholastic Aptitude Test correlated significantly with their test scores of initial error detection ability, while age, sex, measures of piano and vocal instruction, and sophomore music theory grades did not. The low reliability (an average of .6465 over four administrations) of the error detection test used in Shaw's study must bring these results into question.

Brand and Burnsed (1981) examined the predictive validity of previous music abilities and experiences in regard to skill in music error detection among undergraduate instrumental music education majors. Predictor variables included the number of instruments played; ensemble experience; ability in music theory, sightsinging and ear training; and years of pre-college private instrumental instruction. Results indicated no statistically significant relationships between any of the predictor variables and error detection, though these results may be suspect due to the low reliability (.59) of the error detection test instrument. These results suggested to the authors that the ability to detect music errors may exist independently of other music abilities, and that the benefits accrued through participation in one kind of music activity may not contribute to other areas of music development.

Gonzo (1970) compared the pitch error detection abilities of undergraduate choral music education majors with those of experienced secondary choral music teachers. There was no significant difference between the two groups, but significant differences were found between a) undergraduates and those teachers with six to ten years' experience, b) subjects with masters degrees and those without, and c) students with A averages over two years of theory instruction and those with lower averages.

Kepner (1986) attempted to determine the types of errors (pitch, rhythm and expression) produced by high school instrumentalists during prepared and sightread performances as well as the number of those errors which were perceived by the subjects while performing live as against listening to tapes of their performance. Subjects made more rhythmic errors than either pitch or expression errors and identified more errors while listening to tape recordings of their performance than while performing.

In a similar study, Bundy (1987) examined junior high instrumentalists' detection of pitch and rhythm errors in live and recorded performance. In contrast to Kepner's findings, Bundy found pitch error detection in performance significantly better than that in the listening condition, and rhythm error detection significantly better than that of pitch errors in the listening mode. The frequency of perceived errors was significantly lower than actual error frequency, except for pitch response frequency in performance mode.

In a study involving college band instrumentalists, Blocher (1986) used brass trio excerpts with randomly placed errors of articulation, dynamics, intonation, note accuracy, phrasing, and rhythm accuracy. Subjects were randomly assigned to either a conducting or non-conducting group and were required to circle and identify detected errors in group settings. Blocher found no significant difference in the overall error detection performance across error types and no significant effect due to the conducting condition. Music majors performed better than non-majors, brass instrumentalists performed better than woodwind players or percussionists, and upper level and graduate groups performed better than sophomores or freshmen.

Hayslett (1991) investigated peripheral hearing as it relates to error detection. In his study, the attention of undergraduate instrumental music majors was drawn to one part of a three-part excerpt through one or more of the following means; verbal instruction, highlighting in the score, and the insertion of obvious errors. During the playing of the excerpt, another part was deleted. The excerpts were in concert pitch, without tempo, dynamic or interpretation markings which might distract the subjects from the error detection task. The sophomore, junior, and senior subjects were tested for their ability to recognize the deleted part in excerpts drawn from one familiar and one unfamiliar piece.

Significantly fewer deleted parts were identified when the subjects' focus was externally directed, whether focus reinforcers (obvious errors) were or were not present, than when such focus was not externally directed. Whether the focus was externally directed or not, there were significantly fewer deleted parts identified when reinforcers were present than when they were not. Significantly fewer reinforcers were identified when focus was directed externally than when it was not. When the excerpt was from the familiar piece, significantly more deleted parts were identified regardless of whether focus was externally directed and significantly more reinforcers were identified. The most accurate identification of deleted parts occurred when the melody was deleted, then the bass part, and last the harmony part.

Byo (1993) investigated the effects of timbre and texture on the pitch and rhythm error detection ability of graduate and undergraduate music majors. Two excerpts from standard band literature, one polyphonic and one homophonic in nature, were edited to produce 20 taped examples of six measures each. Half of these examples utilized a single timbre and the other half were multi-timbral. Sixteen of the examples, eight of each type, held two errors in separate measures and voices. Thus there were four examples of each excerpt in each timbral condition.

Examples were rescored into four voices, and scores were in concert key with a separate staff for each part. Subjects were given 1 minute to study the score for each of the two excerpts, then heard each example three times, with 5 seconds between hearings, and circled the errors on the score booklet. There were 10 seconds between each set of three repetitions of an example and the next example.

Results indicated that all subjects were significantly better at identifying rhythm errors than pitch errors and at identifying errors in the single-timbre condition than in the multi-timbral one. No effects were observed for degree status (graduate or undergraduate) or for voice placement of errors (soprano, alto, tenor or bass.) There was a significant three-way interaction between timbre, textural placement of errors, and error type. Approximately one-half of the subjects' responses were incorrect.

Byo concluded that the study illustrated the complexity of the error detection process. He stated that "A pitch error in one context...may have a different degree of detectability than the same error in a different context" (p. 165). The multi-timbral condition seemed to hinder the error detection process, and he recommended that instructors in ear training work to improve the transfer of aural ability from a focus on the piano to a focus on multiple timbres. He suggested that future research isolate single and multiple timbres as well as error type, timbre, and voice placement of errors while controlling for the function of the error voice (for example melody, bass line, harmony).

Collings (1973) investigated error detection in a two-phase study. In the first phase, he compared the existing error detection skills of doctoral and senior undergraduate students, using five- and six-part brass quintet examples in concert pitch. In the six-part examples, a baritone horn part was inserted on the same staff as, but an octave above, the tuba part: this baritone horn part was not played. Examples consisted of complete works, and only pitch errors were included. In phase two, discussed below, he compared the effects of three error detection training techniques on error detection ability.

In the first phase, doctoral students performed significantly better than seniors, and seniors had significantly more trouble when there were a larger number of errors in a segment. Collings recommended that a study be made which included rhythmic errors and that the musical examples for such a study be recorded by ensembles of various sizes and combinations of instruments with parts indicated in written pitch.

Mount (1982) studied the differential effects of listening to one voice part, paired parts and all four parts of a Bach chorale on error detection accuracy among graduate choral music and church music majors. Twenty-five errors were inserted into the chorale, and one phrase was presented at a time. Mount found that there were significant differences in error detection test scores between a) the "parts alone" and "paired parts" listening conditions, b) the "parts alone" and "all voices" conditions, and c) the "paired parts" and "all voices" conditions.

The preponderant evidence from these studies seems to indicate that no factor other than experience is related to error detection ability. Detection of errors appears to be more difficult in multi-timbral music than in music with only one timbre, in multi-part music than in solo music regardless of timbral considerations, and in excerpts with more than one error.

Studies investigating methods of training students in error detection

There have been a number of studies on the improvement of error detection skills. Some have employed live performers, others taped performances, and still others computer-generated performances. Error detection instruction has been treated as a part of a regular class and in a self-instructional format, and has been combined with podium-based experiences and with instruction in other, generic, teaching skills. Subjects have included high school students, undergraduates, graduate and doctoral students as well as experienced teachers. Some studies have utilized heterogeneous groups such as conducting classes, while others have focused on choral or instrumental students. The following will serve as a brief overview of the studies which have been done in this area and includes a discussion of the instruction, types of errors, subjects, and instructional approach (class or individual) involved in each.

Boyer (1974), in a quasi-experimental study, investigated the effects of a program featuring one to eight parts of music, in a variety of timbres and with one or more randomly-placed errors of pitch, rhythm, and articulation taped by student musicians. Subjects were members of two undergraduate conducting classes. Scores were in concert pitch, with F and G clefs. The excerpts were from two to twelve measures in length and were taped. The program was linear in format. A control group which received no error detection instruction was tested for comparison with the treatment group.

Excerpts were heard four times following a minute of silent score study. After the first two hearings, the tape was rewound and repeated and then an error-free rendering of the excerpt was presented. After this "clean" recording, subjects compared their responses with the correct answers printed on the back of each response sheet.

Boyer found a significant difference in gain scores in favor of the experimental group, which he taught. Subjects found it easier to determine an error's location than its specific nature. He suggested that his materials be tried on a self-instructional basis and made available for a greater amount of time. He also recommended that such instruction be adapted to several different media (including computer-assisted instruction) and that a branching program be developed.

Costanza (1971) developed a Self-Instructional Program in Score Reading and tested its efficacy in teaching melodic and harmonic score reading skills. Undergraduate music majors who had completed all required theory and ear training courses tested the program. The excerpts included music from the Baroque to the contemporary period, arranged and scored for brass quartet (two trumpets and two trombones) and clarinet quartet (two sopranos, one alto and one bass).

The final form of the program consisted of eight sets of 40 frames each. Subjects received instruction twice weekly for eight weeks. The pretest/posttest took the same form as the instructional program, with 32 items. There was a significant difference between pretest and posttest scores.

Behmer (1984) investigated the effects of a program in score reading on the ability of undergraduate music students to detect errors in instrumental performance. The program consists of three levels: two-stave, three-stave, and four- and five-stave, with increasing numbers of examples at each level. After being given time to study the score, subjects heard the excerpts twice and identified errors of tempo, balance and dynamics, articulation, phrasing, intonation, and note accuracy. They then compared answers with the answer sheet and heard the excerpt for a third time. Weekly one-hour sessions were held for six weeks for the experimental group.

The <u>Visual-Aural Discrimination Skills Test</u> (VADS) developed by Grunow (1980) (discussed below) served as pretest and posttest, with a revised scoring key developed by Behmer. Subjects were members of two conducting classes, which were compared by pretest scores and scores on the Colwell (1970) <u>Music Achievement Test</u> and the <u>University of Illinois Aural Skills Test</u>. Matched pairs were assigned on the basis of test scores on these latter two measures. At the end of the study, subjects completed Likert-scale questionnaires about the instructional program and the VADS test.

The experimental group performed significantly better on the posttest than did the control group. The two groups were significantly different on test scores for Grunow's "Specific Musical Criteria" but not on those for "General Musical Criteria." Behmer concluded that the program was effective in improving the subjects' VADS test scores, although item difficulties and item discriminations on the VADS test were lower than commonly recommended.

The subjects were generally favorable in their responses to the program. However, comments about the program materials indicated some subjects' confusion concerning the categories of error types, the general vs. specific criteria, the difference between "pitch accuracy" and "note accuracy," the varying degrees of acceptability which must be indicated, and the key.

In an attempt to improve the reliability of the VADS test, Behmer appended to it some exercises from the instructional program. No appreciable improvement was noted. He states that as the testing procedure for specific criteria on the VADS test was open-ended, there was no limit to the number of non-existent errors which could be indicated. This resulted in item difficulties in some cases greater than 1.00. Behmer developed a revised scoring key to take the detection of non-existent errors into account and compared it with the original scoring method, which considered only undetected errors. He found that subjects detected items erroneously more often than correctly.

Behmer states that "the (VADS) test structure, character of examples, and scoring methodology hinder the ability of the test and materials to effectively record subject responses" (p. 103). He suggests further research and test development in the employment of discriminant error criteria, pre-planned errors in excerpts with valid musical content and clarity of performance, various subject response methods including a forced-answer format, various scoring techniques, and computer-graded media.

Dolbeer (1969) studied the effects of linear programmed instruction in error detection skills which featured the simulation of an aural teaching environment on the skills of junior and senior music education students (mainly choral majors) as well as the interaction of background factors and any change in those skills. The aural environment simulated was that of a high school band, and this simulation resulted in a number of unplanned errors. Excerpts were presented in condensed score form.

Errors of pitch and rhythm only were used in this study. The error examples were grouped according to their estimated difficulty, and subjects were allowed 90 seconds for prestudy of the score before the excerpts were presented.

There was no statistical treatment of data in this study. Dolbeer concluded that the programmed material was not a detriment to the acquisition of error detection skills, although the majority of his subjects were unable to perceive 50% of the errors on the posttest. He also found that past and present conducting experience had no effect on change in those skills. There was "some evidence that subjects who had completed the formal conducting course, or had participated in varied applied instrumental music study and band, or whose achievement level in music theory was relatively high, showed greater improvement...than those subjects lacking this background" (p. 83). Also, there was no difference on the posttest due to the number of sessions or total hours spent with the program. His subjects reported that melodic rhythm errors were the easiest and pitch errors in the chordal structures of contemporary music the most difficult to detect.

Dolbeer recommended that the difficulty levels assigned to his examples be re-evaluated; that consideration be given to the elimination of the simulated aural environment for purposes of example clarity; and that further attempts be made to identify the background factors which might affect change in error detection skills.

Other recommendations and comments made have specific interest to the present study. Dolbeer's subjects suggested that a correct performance precede each error example. He also stated that "the students should be allowed to study the score at the piano before listening to the recorded material as this is a procedure many teachers follow in preparing their music before rehearsal" (p. 111).

Rizzolo (1969) investigated the effectiveness of tape recordings as opposed to traditional classroom training in teaching undergraduate conducting students sensitivity to intonation errors in triads and chords. Subjects were required to identify the voice in which errors in major and minor thirds, minor sevenths, fifths, major and minor triads and dominant sevenths occurred, and to describe the direction (sharp or flat) of the error. No significant difference was found between the two methods of instruction.

Collings (1973), in the second phase of a two-part study (see above), compared the effects of three error detection training techniques--no prestudy, prestudy followed by error examples, and prestudy followed by correct examples--on the pitch error detection abilities of members of a second semester conducting class, all of whom were junior music majors or minors. These programs utilized random rather than harmonic placement of errors and divided the complete works into six-measure segments with a random number of errors per segment. The investigator controlled the time spent on the programs, and their content validity was inferred. Students were given little time to compare their responses with the corrected examples.

All groups showed significant improvement as well as a significant difference attributable to the number of errors per segment. The investigator states that the large intra-group differences evidenced on the pretest may have masked statistical variability on the posttest. The music used in the pretest and posttest was all Baroque.

Collings recommended that a study be made which included rhythmic errors, and that the musical examples for such a study be recorded by ensembles of various sizes and combinations of instruments with parts indicated in written pitch.

Malone (1985) developed and tested a pitch error detection instructional approach for choral music education majors. Errors related to perfect fourths, perfect fifths, keys, chromatic pitches and intonation of sustained notes, repeated notes, and the thirds of major chords were inserted into musical examples and prepared and performed by a small choral ensemble. These performances were taped, and a pretest, a posttest, and six instructional modules were developed.

Subjects were presented the materials in three settings; one at a time, in small groups, and in large groups. Scores were provided for use during the presentation of the taped materials. Posttest scores were significantly higher than pretest scores, and Malone concluded that the materials were effective with choral music education students.

Michels (1972) also developed and evaluated a program of instruction in pitch error detection for student choral conductors. There were no significant differences between subjects' scores on pretests, midtests, and posttests. Michels concluded that pitch error detection can be taught using visual-aural materials in programmed instruction.

In a study involving 25 members of a choral conducting class, Shaw (1971) developed a self-instructional program in rhythmic error detection. Two groups were formed, matched on a pretest of error detection ability, with voice parts balanced as much as possible. Group A used the program for two weeks, and a midtest was given. Group B then used the program for two weeks, and both groups took a posttest together. The pretest and posttest were identical, and the midtest utilized the same items, rearranged. There were 29 taped examples with 162 errors.

The program utilized a separate series of examples. These were organized into seven tapes, with 15 to 22 examples per tape. Each example was presented twice, once with errors and once without. Subjects were to circle errors on a score.

Both groups exhibited significant improvement from pretest to posttest. The number of years of private instrumental study, the number of years in bands, and test scores on the Math section of the Scholastic Achievement Test correlated significantly with initial error detection ability. Subjects' age, sex, measures of piano and vocal instruction, and sophomore music theory grades did not.

Hopkins (1991) investigated the relative effectiveness of four approaches to score study on undergraduate music students' ability to detect and notate pitch and rhythm errors in choral music. The pitch errors were a minor second or more away from the notated pitch, and rhythm errors were half or two times the value of the notated pitch. The four approaches used were score alone, score study at the piano, sightsinging from the score, and score plus correct recorded example.

Different complete compositions were used for each of four sessions. Each subject experienced all four score study styles, with groups of four subjects going through the styles in different orders. The compositions were Four Motets by Christopher Tye. Subjects were given 50 minutes to study the score. After a short break, all subjects were gathered together and given three hearings of the performance with five errors each of pitch and rhythm. They were instructed to listen only during the first hearing, circle detected errors during the second, and notate the errors as they were performed during the third. Scores were collected one minute after the third hearing.

Hopkins found that the subjects were more adept at detection of errors than at notating them, and that they detected and notated more rhythm errors than pitch errors. The use of score plus correct recorded example resulted in significantly higher test scores than the study of the score at the piano, but no other significant differences due to score study style were found. Pianists received higher test scores than non-pianists, but this difference was not significant.

Stwolinski et al. (1988) compared keyboard sightreading with listening to recorded examples as methods of learning to detect harmonic alterations in simple piano works. The pieces were characteristic of those which would appear in college-level class piano courses. Four short excerpts of 6 to 12 measures in length were utilized in the test, with four to five harmonic errors inserted in each. Usually only one chord tone was altered, changing the quality or function of the chord, or changing both.

Subjects in the tape-recording group heard three correct renditions of each excerpt and then one incorrect one. Those in the sightreading group were allowed three readings and were then presented with the incorrect performance. The researchers found that the group which used tape recordings was significantly more accurate in identifying harmonic alterations on the posttest.

Stuart (1977) studied the effects of error detection training in conjunction with training in three "generic teaching skills." The subjects were undergraduate members of orchestra clinic classes, with approximately half of each class designated "students" and the other half designated "teacher trainees." In both classes the teacher trainees worked on developing conducting-rehearsal techniques for four weeks, and those in one class worked concurrently on developing error detection skills.

The training utilized videotapes, slides, textual materials, and class discussions, and focused on aural perception, bowing, positional aspects, intonation, and their combinations. The pretest consisted of videotaping the subjects in the process of rehearsing-conducting the "students," all of whom were beginners. The posttest was comprised of three sections: a videotape during which subjects were to identify and locate errors of specific types; a multiple-choice test concerning string bowing functions; and specific questions concerning string scores. Stuart's two groups--that which received error detection training (EDT) and that which did not (NEDT)--exhibited no significant differences on the pretest. On the posttest there was a significant difference in mean test scores in favor of the EDT group, although both groups improved significantly.

Stuart stated that "the training procedures for error detection were inadequate" (p. 193) in her study. She recommended that error detection training should begin with single parts using various timbres and focusing on intonation and rhythm and should gradually become more complex through the introduction of more parts, more complex rhythmic patterns, textural difficulties, and/or harmonic inaccuracies.

Tromblee (1972) developed and tested a program for teaching intonation discrimination. Subjects were high school and college band students, assigned to four groups by test scores on the Pitch, Timbre, and Tonal Memory sections of the Seashore Measures of Musical Talents. There was one experimental and one control group for high school students and one of each for college students. Tromblee developed an intonation discrimination test for use as pretest/posttest. The test consisted of 100 taped harmonic intervals and triads, each with one out-of-tune note. C major and minor triads and their constituent intervals were the only materials, and the items were presented by clarinet, trumpet, and bassoon. The first section of the test contained 60 harmonic items. Triadic items made up the second and third sections, with 24 and 16 items respectively.

In the first and second sections, the out-of-tune notes were identified, and the subjects were required to determine whether the indicated note was sharp or flat. In the third section, the subjects were asked which of two possible tones was out-of-tune and in which direction. At intervals of 10 items, perfectly tuned "reference chords" were presented.

The experimental groups received eight pre-recorded lessons, progressing from easy to difficult items and using the same materials as the test. Tromblee found a significant difference in favor of the experimental groups on sections I and II of the test, but not for section III. He concluded that the intonation problems presented in the third section were excessively difficult.

DeCarbo (1981) compared the effects on undergraduate instrumental music students' conducting and error detection skills of programmed error detection instruction and error detection instruction given while students conducted live musicians. He used the same materials for both groups: eight complete compositions from various style periods, of two, three, and four parts in full score (concert pitch) and performed by brass instrumentalists (two trumpets, one French horn, and one trombone). Both groups were given the opportunity to study the scores outside of class time.

The conducting skills addressed in the study were basic ones. The glossary of errors used was obtained from structured interviews with active conductors and included dynamics, ensemble considerations, intonation, note accuracy, rhythm accuracy, and style (articulations). These errors were assigned randomly by measure, part, error type, and style of excerpts. The 80 items were generally four to eight measures long.

In the conducting experience condition subjects were asked a series of four questions relating to each excerpt and repeated the excerpt until they could answer all four correctly. These questions concerned the measure, type, and performing instrument of the error, and its precise nature. Those in the programmed learning group were given the correct answers after responding to the first hearing of the excerpt, then rewound the tape and replayed the excerpt as many times as needed until they understood the reasons for the correct answers. Each group received this training as a part of the regular class period.

The written test used taped excepts which were heard twice. In the conducting test, subjects stopped the group as soon as they heard an error, responded to the four questions, then repeated the excerpt and had the opportunity to change their answers.

DeCarbo found no significant difference between his two groups on the written test, but a significant difference in favor of the conducting experience group on the conducting test. He concluded that the error-selection lists were effective and efficient in assigning errors to the conducting experience format during an ongoing class; that the use of concert key made score reading, score study, and error detection easier; that detecting errors in two, three, and four-part scores was difficult for his subjects; and that his approach provided diverse areas for error detection without over-complicating the task. The latter conclusions led him to suggest that the use of full-band examples would make the error detection task even more difficult and might not be beneficial.

Sidnell (1968) studied the differential effects of non-programmed instruction and a linear program with a four-stage response format on the pitch and rhythm error detection skills of 26 junior instrumental music education majors. In the program each excerpt was presented four times with each successive presentation focusing more narrowly on the error until, having identified the measure, type, and performing instrument of the error, the student heard only that instrument and wrote the error as performed. In the non-programmed format, each excerpt was heard twice with no narrowing of the focus of attention, and full identification of each error was required immediately after the second hearing.

Using college musicians, Sidnell re-recorded errors gleaned from performances by high school and junior high bands. Each excerpt was four measures long and involved randomly-placed errors. The excerpts utilized all standard orchestral and band instruments in concert pitch, in three, four, or five parts on a four-line staff with only G and F clefs. He employed a matched-pairs two-group experimental design, with the matching done on the basis of the results from various standardized tests, an investigator-produced test of score reading ability, and grade-point average in second year theory courses.

Results of Sidnell's study indicated that while both groups improved significantly, the group receiving programmed instruction had a significantly higher mean gain that did the group which received the non-programmed instruction. He found no significant correlation between the matching variables and change in error detection ability. He recommended drills involving a wider range of timbres, an increasing number of staves in the conductor's scores, and a branching rather than linear format.

In studies involving band directors as well as graduate and undergraduate music students, Grunow (1980) studied the relative effectiveness of four modes of score preparation--score only, score and recorded example, recorded example only, and no preparation--on the development of "visual-aural discrimination skills," using typical instrumental combinations. His examples ranged from two to five parts and from 4 to 16 measures in length, and were performed by woodwind, brass, and string ensembles. The music used ranged in difficulty from elementary to high school level. He also investigated the correlations between gain scores and the effects of years of teaching experience, grade level taught, and degree earned.

The errors used in Grunow's program were suggested by the student musicians and conductor who recorded them, as well as by the investigator and his advisor. The subjects were required to respond to the excerpts by deciding whether the tempo, balance, articulation, tone quality, and intonation ("General Musical Criteria") were acceptable, questionable, or unacceptable, and by indicating on the score errors of rhythm, note accuracy, pitch accuracy, articulation and ornamentation, phrasing, and dynamics ("Specific Musical Criteria"). On the pretest and posttest they were allowed 30 seconds before each of two hearings to study the score, progressing from the general to the specific criteria. Finally, a correct version of the excerpt was presented and subjects were given 30 seconds to compare their answers with the answer sheet and compute their test score.

Subjects in the "score only" prestudy condition were allowed one minute to study the score before each excerpt of the program was presented. Those in the "score plus recorded example" group were also allowed one minute total, but after the first 10 seconds, a correct performance was heard, with the remaining time also allotted to score study. Those in the "recorded example only" condition heard the correct performance twice, with no score. Those with no prestudy went straight to the error examples.

All four groups made significant gains from pretest to posttest, but there was no significant correlation between test scores on these measures and any of the four modes of score preparation. The subjects' background characteristics were not related to pretest, posttest, or gain scores. The subjects scoring highest on the posttest were correct on only 50% of the Specific Musical Criteria items. Also, the answers on the items were combined artificially into three categories--expressive, rhythmic, and melodic components--for statistical treatment, which could have masked significant differences between groups. This criticism is borne out by the low reliability of the test. As a result of his investigation, Grunow recommended a study to research and utilize materials less demanding (e.g., single-part scores with fewer errors) or more demanding (e.g., contemporary idioms and more complex instrumental and vocal combinations) than those used in his program.

Doane (1989) compared the effects of the Grunow and Froseth (1979) MLR Instrumental Score Reading Program with those of an experimental instructional program developed by the investigator. The subjects were 30 instrumental music education students in a two-semester course sequence. During the first semester, students were pretested with the MLR Instrumental Score Reading Test (Grunow and Froseth, 1979), and then underwent a six-week instructional program using the MLR Instrumental Score Reading Program. The MLR Instrumental Score Reading Test was administered as a posttest. The experimental evaluation instrument was administered one week later so as to correlate its results with those of the MLR Test: A correlation between the two tests of .47 was established.

This administration of the experimental evaluation instrument also served as a pretest for the experimental instructional program. Subjects were instructed to conduct alternate examples on this evaluation instrument. Examples for both the test instrument and the experimental instructional program included from two to six errors of pitch, rhythm, articulation, or precision. The examples, which were taken from wind band literature, were from 8 to 37 measures in length and contained from four to six parts in written pitch.

During the second semester, subjects used the experimental program for a period of six weeks. One hour per week was spent in a classroom ensemble situation with errors inserted into the instructional examples. Every subject received two opportunities to conduct the ensemble during the six-week period. Subjects spent from 15 minutes to 3 hours per week in individual practice, during which time they either conducted a video monitor as the videotaped examples were presented or listened to the examples on an audio tape. Subjects were instructed to identify the measure in which an error occurred, the instrument performing the error, the general type of error, and the specific nature of the error. Four audio and four video tapes were produced, each with examples containing errors of one error type only. At the end of this six-week session, the experimental evaluation instrument, with error examples on audio tape, was administered as a posttest.

Gain scores for both semesters were significant at the <.01 level. Subjects were not as accurate in detecting errors when conducting, but this difference was not significant. Doane concluded that the act of conducting did not interfere with the error detection task.

The improvements in test scores on the MLR Instrumental Score Reading Test (Grunow and Froseth, 1979) were compared with a numerical index of each student's accomplishment in aural theory. The correlation between these variables was .11, which indicated to the investigator that grades in aural theory classes are not useful in predicting the development of an individual's error detection skills. Doane states that "Results of the study did support the conclusions of earlier research which implied that conductor aural error detection skills seem to be developed independently of skills taught in traditional aural skills classes" (p.14) and that success in such classes does not indicate that a student is more likely to develop skill in error detection. He recommended that further research investigate a hierarchy of errors by difficulty of detection.

Ramsey (1979) investigated the effects of an individualized programmed instruction approach of three different lengths on the error detection skills of 77 upper-level undergraduate music students. The three forms of his program had 114, 76, and 38 frames and utilized taped full band excerpts with pitch and rhythm errors only. The errors utilized were those reported by conductors as typical of those which they corrected in their rehearsals.

The excerpts were grouped according to their estimated difficulty and were recorded by a college group of the wind ensemble type. There was one error per excerpt, and subjects were asked its location, then the instrument performing it, and finally its nature. On the tests, subjects were given one minute to study the score and then heard the excerpt three times, with decreasing intervals of time between subsequent hearings. While working through the program, subjects had to answer all three questions but could hear each excerpt as many times as they wished; this resulted in a hybrid between linear and branching program styles.

When posttest scores were adjusted for differences on both the pretest and the Aliferis-Stecklein Music Achievement Test, College Midpoint Level, all experimental groups (combined data from three state universities) showed significant gains while the control groups (combined) did not. Groups using the longer forms exhibited significant differences from the control groups on the posttest, and the group using the 114-frame form performed significantly better on the posttest than did that using the 38-frame form.

Ramsey suggested a program beginning with small instrumental ensembles (trios, quartets, sextets) and progressing to large ones (full band or orchestra) with a variety of score types. Other recommendations included the investigation of different drill time periods and of the retention of error detection skills.

Deal (1983) compared the effects of Ramsey's (1979) program with those of a computer-assisted program based on it. The subjects were junior and senior instrumental music majors. His Computer-Assisted Program in Error Detection (CA-PED) was written for the Apple II+ and had several limitations due to the technology available at the time. He used excerpts of two to four measures, in four voices, with three timbres, no key signatures or beaming of 8th or 16th notes, and with some examples rescored in order to avoid the crossing of alto and tenor voices on the small grand staff which was used. He based his program on Ramsey's 114-frame program, discarding 16 excerpts due to the unavailability of percussion timbres. Also, memory limitations of the computer program limited the number of excerpts which could be included in each section.

Both groups made significant gains. There were significant differences between the two groups on the pretest. After posttest scores from the three participating schools had been combined and adjusted for these differences, no significant differences were found.

Deal concluded that both programs were effective. He recommended that the program be expanded to include more than four voices, record-keeping functions, key signatures, and longer excerpts (through the use of scrolling on the computer screen). He also recommended the investigation of various computer-generated timbres and their combinations in order to determine which are most effective in teaching error detection skills, as well as further study of the effects of repeated hearings and of response feedback as opposed to prompts.

Jones (1990) designed and tested a computer-assisted error detection instruction program for undergraduate and graduate instrumental music education students. Examples were from four to eight measures of in length, were taken from representative band literature, and were recorded on reel-to-reel tape by a college wind ensemble. One correct and four incorrect performances were recorded. The incorrect performances included one error of either pitch, rhythm, articulation or style (including dynamics and performance practice), or ensemble (including synchronization, imprecise attacks and releases, and miscounting of rests resulting in an early or late entrance). Errors were inserted into parts and measures randomly.

The program consisted of 20 excerpts in a printed score booklet. For each, the computer randomly selected and presented one of the five possible versions through the use of a random access audio device. Subjects were to identify the type of error, then the measure, and finally the part in which the error occurred, with a limit of three attempts at each level. The subject could hear the error example again or hear a correct version for comparison.

Each subject completed 10 sessions with the same excerpts. Jones found that the mean item response time decreased the most during the first five sessions. There were significant changes in mean test scores from the pretest to the posttest and between some sessions for errors of articulation/style and errors of ensemble. There were also significant positive changes in mean test scores from session to session. There was no correlation between subjects' perception of either the effectiveness of the system or of personal improvement and variables of error type, measure location, instrument performing the error, and mean total response time.

Jones concluded that the program was effective in providing instruction in error detection and score reading, but that enhancing the audio quality of the examples would minimize distractions and help avoid the subjects' tiring of the program. He suggested that his program should eventually use digital recording and playback technology, specifically on compact disk when such become editable. Jones also recommended the use of solo and small ensemble examples.

Gruner (1993) investigated the effects of a computer-assisted error detection skill development program using two- to five-part synthesized excerpts on the pitch and rhythm error detection abilities of volunteer undergraduate instrumental music education majors. These subjects had completed or were enrolled in a beginning conducting course.

The investigator developed the Acoustic Error Detection Skills Inventory (AEDSI) and the Computerized Error Detection Skills Program (CEDSP) for the study. Both the inventory and the program utilized a printed score booklet and examples taken from wind band literature of medium difficulty.

After all subjects had been given the AEDSI as a pretest, the experimental group received approximately eight hours of exposure to the CEDSP over a period of eight weeks. The instructional program used synthesized sound sources and consisted of examples of four levels of difficulty arranged according to the difficulty of the literature from which they came and the number of errors included. Some synthesized timbres were excluded due to poor sound quality. Subjects were allowed to hear the examples as many times as they wished before identifying the measure, instrument, and exact nature of the error performed, and were ultimately guided to the correct answer through a forced-answer approach.

Gruner found that composite gain scores were significant for the treatment group but not for the control group. Both groups improved significantly in rhythm error detection, but neither improved significantly in pitch error detection. He found a significant difference between the composite test scores of the two groups on the posttest, but not on the pitch and rhythm subtests. Gruner concluded that

> a) error detection skills can improve with training...b) programed, self-instructional error detection training formats are effective...c) computer-assisted instruction is a viable means by which error detection skills can be improved...and d) computer-assisted instruction in error detection skills can be effective with enhanced sound quality and faster computer response times. (p. 45)

Gruner suggested that a tutorial component be added to his program which would train students in focusing their listening habits and aid in the development of their error detection skills. He also recommended adding a correct version of each example for comparison, adding on-screen full scores with the rhythm or pitch error highlighted for reinforcement after an incorrect answer is given, improving some of the synthesized sounds, acoustically panning the synthetic sounds to different locations within the speaker to emulate actual rehearsal situations, and expanding the number of voices in the examples to full band instrumentation.

Summary

The wide variety of subjects, conditions, training procedures, and error types featured in the above studies, and the recommendations made by the investigators, indicate the need for further investigation into error detection. Each of the studies cited above has recommendations for such investigation. These recommendations include the use of multiple timbres (Byo, 1993; Collings, 1973; Sidnell, 1968), ensembles of various sizes (Collings, 1873; Grunow, 1980; Deal, 1983; Jones, 1990; Gruner, 1993), and scores in written pitch (Collings, 1973) with an increasing number of parts and timbres (Stuart, 1977; Sidnell, 1968; Ramsey, 1979). A focus on errors of pitch and rhythm only (Stuart, 1977; Collings, 1973) in clear performances with pre-planned errors (Dolbeer, 1969; Gruner, 1993) is suggested, along with score study at the piano (Dolbeer, 1969). Finally, an investigation of computer-assisted instruction (Boyer, 1974; Behmer, 1984) utilizing computer-generated timbres (Deal, 1983; Gruner, 1993) with scores presented on the computer screen (Gruner, 1993) is recommended.

Several studies have featured the investigation of prestudy and its effects on error detection skills, using different time lengths and definitions of such study. Only two studies have involved prestudy at the keyboard. One of these investigated error detection in simple piano works: The other allowed 50 minutes of prestudy for a complete four-part choral piece into which a number of errors had been inserted.

Those studies which utilized taped excerpts and/or hard copies of program booklets all exhibit the inevitable clumsiness of such materials. In some cases, students had to rewind a tape before they could hear an excerpt a second time. In others, unwanted errors crept into recorded examples, or planned errors were covered by extraneous noise or inadequate recording balance. Some programmed texts required students quickly to compare their answers with corrected scores which appeared on the backs of their answer sheets.

Three studies utilized computer-assisted instructional programs (Deal, 1983; Jones, 1990; Gruner, 1993). None of the three investigated the comparative effects of score study styles. Each of these studies was limited by the technology available or its application.

In the first study, the investigator was limited in number of parts (four), timbres available (three), length of excerpts (two to four measures), and score type (a small grand staff in concert pitch, without key signatures, where the middle voices had to be rescored to avoid having their lines cross on the score). The second study utilized a random access audio device as source for its examples. The audio quality of the recorded examples was such that the subjects were easily distracted from the planned errors and grew tired of the program. Also, the randomly-inserted errors included not only pitch and rhythm but also errors of articulation and style and of ensemble. Excerpts in either full or condensed score were presented in a printed score booklet in which the subjects recorded their responses. The last of these studies also presented excerpt scores in a printed booklet, and was limited in its choice of timbres.

# Chapter III

# Procedures

## Experimental Design

The focus of the study was the effects of four modes of score study on the pitch and rhythm error detection abilities of undergraduate beginning conducting students. A counterbalanced design, as described in Campbell and Stanley (1963), was employed. This experimental design incorporated four computer-assisted tests and intact groups at four universities. Subjects were 30 members of beginning conducting classes.

Two Latin Squares were used in an attempt to control for possible presentation order effects for the score study styles and the individual tests. Each intact group received the tests in a different order. Within intact groups, each subject received a different score study order from the previous three subjects.

## Excerpts

### Selection and assignment to example sets

The excerpts for the four example sets were chosen from those developed by Ramsey (1979). There were 134 possible excerpts, including the 114 excerpts in the long form of his instructional Program in Error Detection and the 20 items on his Test in Error Detection. These were screened in order to remove any excerpts with errors in percussion parts which could not be transferred to other parts. Percussion timbres were not available for use in this study. Those percussion errors which could be rescored into other parts, including doubled rhythmic sections and timpani parts doubled in bass parts, were retained. Ten excerpts were eliminated, with 124 retained. Each of the four computer-assisted tests contained 31 of these excerpts plus the two-part practice excerpt, for a total of 32 excerpts.

The pieces from which the excerpts were taken are listed in Appendix A. Permission was obtained from the appropriate publishers for their use. A sample permission request letter appears in Appendix B.

Ramsey determined difficulty coefficients for the excerpts used in his program and test. On the test, these ranged from .9231 for the least difficult to .2009 for the most difficult items. On the program, the difficulty coefficients ranged from 1.0000 for the least difficult to .0000 for the most difficult items.

Those excerpts which had been selected for inclusion in the example sets were grouped according to these difficulty coefficients for assignment to specific sets. All items of one difficulty level were assigned before any excerpt of the next, more difficult level was assigned. That is, all items with a difficulty coefficient of 1.0000 were assigned before moving to those with a difficulty coefficient of .9333.

In order to improve the chances for equivalent example sets, a modified random assignment of items to sets was used.

Items were assigned on the basis of a double coin toss. Table 1 shows how example sets were related to coin tosses.

Table 1

Relation of Coin Tosses to Example Sets

| Toss | Set |
|---|---|
| Heads-Heads | A |
| Heads-Tails | B |
| Tails-Heads | C |
| Tails-Tails | D |

Each Ramsey item number and its difficulty coefficient were written on a slip of paper, and these slips were grouped by difficulty coefficient. On the first double coin toss, one slip was drawn from the group with a difficulty coefficient of 1.0000, and was assigned to the appropriate example set. After the next double toss, another was drawn and assigned. When this toss indicated the assignment of a second excerpt to an example set before a full rotation of the four sets had occurred, the toss was void. Thus, no set received a second item before all four had received their first. This procedure was adopted so as to ensure that no one set would receive an inordinate number of items of any particular difficulty. Had the assignment been made by coin toss without modification, one set conceivably could have received a majority of the easiest items and another a majority of the most difficult (conceivable but unlikely). Table 2 shows the median difficulty coefficient of each example set which resulted from this procedure.

Table 2

Median Difficulty Coefficients of Example Sets

| Set | Median Difficulty Coefficient |
|---|---|
| A | .4666 |
| B | .5333 |
| C | .4666 |
| D | .5333 |

Rescoring of the excerpts and adjustments to the example sets

The excerpts within each set then were rescored into the appropriate number of parts, with the first four (after the practice example) being one part, the next four two, and so on, until the last three excerpts had eight parts. The excerpts were then rescored in the appropriate number of parts in natural groupings using Finale® (1989) and a Korg DS-8 electronic keyboard. The Finale® software program allows the user to set up a score in any fashion desired, with parts in written pitch, and to assign each staff of a score to a specific MIDI (Music Instrument Digital Interface) channel.

The instruments/timbres used in the program were flute, clarinet, oboe, bassoon, French horn, trumpet, trombone, baritone/euphonium and tuba. The same voice setting was used for the baritone/euphonium and tuba, with range being the only difference between the two. Each staff carried one part. If two or more parts required the same instrument, each was assigned a different staff but the same MIDI channel. Site managers at the four participating universities assigned the sampled timbres of their individual MIDI keyboards to the specified channels. As a result, the author was assured that subjects would hear the examples as intended without having to program a separate set of instructions for every possible keyboard.

The excerpts were rescored using natural groupings, that is, combinations of instruments commonly found in instrumental music. These included brass and woodwind quartets and quintets, double quartets, duets and trios of like instruments, and others. For purposes of this study, the French horn was included in both brass and woodwind excerpts.

Once the excerpts were rescored, each example set was checked to ensure that two brass and two woodwind excerpts were included at each level from single-part through seven-part excerpts and that there were one brass and two woodwind excerpts at the eight-part level. Where this was not the case, the excerpts were rescored as needed. Finally, a check was made to ensure that no excerpt was repeated in any one example set. In one instance, two excerpts were exchanged in order to avoid such a repetition and produce equivalence of instrument families. This did not affect the median difficulty coefficients of the example sets. The specific items in each example set and their Ramsey difficulty coefficients are shown in Appendix C.

Once all error excerpts and their correct counterparts had been completed and checked visually and aurally, the correct scores and a segment of each score showing the error as performed were transferred to the computer program example stacks using Screenshot™ (Whittingham, 1990-91) (see below). The correct and error excerpts were also stored as MIDI files so that the program could present them aurally. These files were played using the HyperMIDI™ (Redmon, 1990) software application. This application interfaces between HyperCard© (Atkinson, 1987-90) and MIDI files, allowing the control of such files from within HyperCard©.

Score Study Order

Four score study styles were investigated in this study: no score study, study with the score alone, study with the score and a correct aural example, and score study at the electronic keyboard. Each subject received these score study styles in one of four possible score study orders. The assignment of score study order to subjects was done on a rotational basis, with the four orders entered into a Latin

Square so that the fifth and ninth subjects in any one intact group received the same score study order as the first. This Latin Square arrangement, illustrated in Table 3, was designed to control for the effects of presentation order of score study styles.

Table 3:

Latin Square for Score Study Order Within Groups

| | Session | | | |
|---|---|---|---|---|
| Student | 1 | 2 | 3 | 4 |
| 1, 5, 9... | X1 | X2 | X3 | X4 |
| 2, 6, 10... | X2 | X1 | X4 | X3 |
| 3, 7, 11... | X3 | X4 | X1 | X2 |
| 4, 8, 12... | X4 | X3 | X2 | X1 |

Note. X1 = no score study, X2 = study with score alone, X3 = study with score plus correct example, and X4 = study with score at the keyboard.

Development of the Program

A set of six Hypercard© stacks was developed for the presentation of examples and the management of the study. Each example set stack was capable of presenting the materials under every score study condition. These stacks served as the computer-assisted tests for the study. An introductory stack assigned a score study style order to each subject upon his or her first session. There were four possible orders, and each subsequent subject was assigned a different order until all four had been assigned, when the series was repeated. This stack also initiated the correct score study style for each session according to the order assigned to each subject, and opened the correct example set stack for each session. At the end of each session, data recorded for the session were transferred to a separate statistics stack.

The tests were developed for the HyperCard© software application on a Macintosh LC computer with a 40-megabyte hard drive. HyperCard© is a software development package which emulates stacks of index cards. A stack consists of a series of screens which represent the cards. The stack developer designs the background for the stack and incorporates buttons, fields, menus, and graphics, all of which can be specific to one card in a stack or common to all. "Scripting" is the term used for programing the various elements of the stack. Stacks, backgrounds, individual cards, buttons, and fields may all be "scripted" by the developer.

The program incorporated two interface methods. The subjects used the computer keyboard to identify themselves to the program and the "mouse" to respond to the error example questions. The Macintosh utilizes a graphic interface, with icons representing various activities and documents. In order to make a choice, the user moves the mouse, which moves an arrow around the computer screen. When this arrow is pointing to the desired icon or spot on the screen, a button on the mouse is "clicked" and the choice is made.

For this study, three types of stacks were developed (see Appendix D). The introductory stack was the first seen by the subjects. A new card was produced for each use of the program, and subjects gave their name, the instructor's name, the time or section of their class, and the number of the session. The program provided the date and time, assigned a score study order on the first session by any one subject, and kept track of that order and the particular score study style for each session. Thus, a complete record of each subject's use of the program, the score study order assigned to that subject, and the score study style implemented for each session was kept in this stack.

The program for the introductory stack moved the subject into the correct example set, engaged the script for the score study style assigned to that subject for that session, and eliminated access to the programming environment in order to prevent tampering. Each intact group received the example sets in a different order, so that all four possible orders of the Latin Square shown in Table 4 were used. Every member of a group received the example sets in the same order. The script of each example set stack included instructions for the subjects for each score study style, which were triggered by information from the introductory stack. All example sets began with the same two-part practice example.

Table 4:

Latin Square for Example Set Order Between Groups

| | Session | | | |
|---|---|---|---|---|
| School | 1 | 2 | 3 | 4 |
| | Example Set | | | |
| I | A | B | C | D |
| II | B | A | D | C |
| III | C | D | A | B |
| IV | D | C | B | A |

Under the "no study" condition, the score for each excerpt was shown and the error example played immediately. The subject was asked to indicate, by pointing and clicking with the mouse, first the part, then the measure, and finally the specific note which was in error. Each mouseclick was recorded as an attempted answer. Upon correct identification of the error note, or when an incorrect answer was given after nine attempts, the program ended the subject's access to the example and presented the error as performed, and the subject was prompted to move to the next example. The same questions and limits were utilized under all score study conditions.

Under the "score only" condition, subjects were allowed to study the score as long as they desired before playing the error example by clicking on a button. Under the "score plus correct example" condition, the program automatically presented a correct rendition of the excerpt when the score was shown. Again, the subject controlled the time of presentation of the error example. Finally, under the "score plus keyboard" condition, before hearing the error example the subjects were encouraged to play through the excerpt on the electronic keyboard provided.

The program recorded the following data on each session: subjects' names, their instructor's names and class time or section, the session number, the assigned score study order, the score study style for each particular session, the example set, the date of the session, and the starting and ending times of the session. The program also recorded the time (in seconds) spent studying each excerpt, the time taken to answer correctly each of the three questions, the total time per excerpt, the number of attempts (in mouseclicks) at each of the three questions, and the total number of attempts per excerpt. If a student failed to answer a particular question, no attempts were recorded for it. However, the total attempts data formed a record of such an event. It will be seen that under the "no study" condition, there was no time recorded for score study. If a subject chose to look over the score <u>after</u> the error example was played, such time was recorded as spent on the first question asked about the excerpt.

The third stack developed for the study was the statistics stack. Upon the subject's completion of each example set, the program copied the data recorded, opened the statistics stack, and placed the data on a new card. During this copying process, the computer screen showed the original statistics card. When the copying was complete, the subject was returned to the screen from which the session was begun. Thus the statistics stack was never available to the subjects. The original card remained in the example set stack and was erased when the set was next used, with the new subject's data replacing the previous subject's.

Validity and reliability of the tests

The content validity of the tests/example sets was inferred from the validation procedures used by Ramsey in his study, since all excerpts came from his program and test. Instrumental music teachers supplied examples of pitch and rhythm errors which they felt were typical or which they found themselves correcting frequently. These typical errors were then inserted into excerpts from commercially available band literature of medium difficulty, with care taken to avoid altering the music in any way other than the desired error. Recordings of these error excerpts were then evaluated by a panel of experts to determine whether the errors were typical, and to validate the recordings themselves. Finally, a pilot test was undertaken with a group of instrumental music majors in order to develop a continuum of difficulty using item analysis.

As a check of the reliability of the example sets used in this study, one set was chosen at random and tested for internal consistency using a split-half approach. A Pearson product-moment correlation was calculated for Example Set D and corrected using the Spearman-Brown prophecy formula. A reliability of .81 was established for this set.

## Pilot Testing

### First pilot test

The example sets were pilot tested in early fall of 1992 by members of an undergraduate conducting class. For the pilot test a Roland SC-55 "Sound Canvas" Tone Generator was used to provide the sampled sounds, as the available keyboard, a Korg DS-8, did not have sampled-sound capability. Subjects heard the excerpts through headphones. Subjects were asked to keep a record of any difficulties incurred while using the program, complete a questionnaire (Appendix E), and provide any comments they wished about the program. As a result of the pilot testing, and in response to reported difficulties and other comments, minor changes were made to the computer program, and problems with some excerpts were corrected.

The program was delivered to four midwestern universities for their use in the fall semester of 1992 with undergraduate conducting classes. Difficulty arose with the data collection part of the program due to incorrect information from the software supplier, and the study was postponed for a semester.

Second pilot test

With the data collection problems corrected, the program was issued to the schools again in the spring of 1993. However, one university suffered a break-in with the loss of much of its hardware and all data for the study, another suffered a hard disk crash with the loss of all data, and a third was unable to complete the program. The data from the one university which was able to complete the study were used to test the data treatment plan for the final study.

The study

In the fall of 1993, new copies of the program were sent out. Samples of the enclosed cover letter, instructions for installing the program, and suggested student instructions appear in Appendix F. Three universities were able to complete the program. Data disks were received from these participants in November and December 1993.

In periodic phone conversations during the semester, the fourth participant indicated that all was proceeding as planned. However, despite repeated attempts by the investigator to obtain the data, none was forthcoming (see Appendix G). On February 12, 1994, the original program disks were returned to the investigator with no data and no explanations.

Treatment of the data

Error detection ability was measured by the total number of attempts made in answering the error identification questions per session and the number of such attempts per excerpt. The total time taken per session and score study time taken per session were considered to be functions of score study style assigned. Means were calculated for each of these measures and tabulated by subject, by group, by score study style, by example set, by session, and by number of parts in excerpts. These data were then analyzed using one-way analyses of variance and, where indicated, post-hoc *t*-tests.

# Chapter IV

## Presentation and Interpretation of Data

### Introduction

The investigator mailed out computer disks containing the program stacks in August of 1993 to four universities. Three went directly to the instructors of the conducting classes taking part in the study and one went to a music computer laboratory coordinator. Participants were contacted by phone at least monthly except in one case, where a graduate assistant was in charge of program oversight. The investigator visited one university during the course of the study at the participating faculty member's request. Difficulty accessing the program had arisen after it had been in use for some time. The problem was traced to a change in the start-up routine of the computer effected by another faculty member, which change required a different procedure for entering the error detection program. The investigator had no contact with the subjects at this university, and the problem was solved in one morning.

Three participants sent study data to the investigator during November and December 1993. The investigator contacted the fourth university repeatedly from October 1993 through early February 1994 with no effect (see Appendix G). Finally, through contact with other faculty members at this university, the original program disks were returned on February 12, 1994 with no data.

Descriptive statistics were obtained and analyses of variance performed on the data received. Where indicated by analysis of variance, post-hoc *t*-tests were also performed.

The responses obtained included incomplete sets (less than four sessions per subject) and duplicated score study styles and example sets. This indicated that a number of subjects did not follow directions at the beginnings of sessions, or failed to complete all four sessions. Duplicated sessions were dropped from the study. Where no designed score study order could be attributed, the sessions involved were omitted from score study order data analyses.

Each intact group received the four example sets in a different order. These orders were: for Group I, with 0 subjects, Sets A, B, C, and D; for Group II, with 11 subjects, Sets B, A, D, and C; for Group III, with 6 subjects, Sets C, D, A, and B; and for Group IV, with 13 subjects, Sets D, C, B, and A. The number of cases in each group for each score study style and example set is presented in Table 5.

The computer program assigned a score study order for the four sessions to the subjects in each group on a rotating basis. Sixteen sessions were completed under score study order 1 (X1, X2, X3, X4), 21 under order 2 (X2, X1, X4, X3), 12 under order 3 (X3, X4, X1, X2), and 11 under order 4 (X4, X3, X2, X1).

Table 5

Number of cases for each group by score study style and example set

| | Group | | | Totals |
|---|---|---|---|---|
| | II | III | IV | |
| Score Study Style | | | | |
| X1 | 7 | 6 | 11 | 24 |
| X2 | 9 | 4 | 11 | 24 |
| X3 | 6 | 6 | 13 | 25 |
| X4 | 6 | 6 | 12 | 23 |
| Example Set | | | | |
| A | 7 | 5 | 11 | 23 |
| B | 11 | 6 | 13 | 30 |
| C | 4 | 6 | 10 | 20 |
| D | 6 | 5 | 13 | 24 |

The data collected for each session consisted of the number of attempted answers, operationalized as the number of mouseclicks, recorded during each session and example; the total time taken for each session and example; and the time spent in studying the score for each session and example.

The number of attempted answers per session was taken as a measure of the difficulty of the task. The null hypotheses for the study focus on this measure.

Total time per session and study time per session were considered to be functions of the score study style used during the session and not measures of difficulty. Thus, it was presumed that the total time and study time taken for sessions under "Score Study Style X4: Score plus Keyboard" should have been greater than that taken for "Score Study Style X1: No Score Study." This is an important consideration, as the results of the study could be called into question if the subjects did not utilize the correct score study styles for each session. Analyses of the data pertaining to score study time and total time were performed in order to validate the procedures of the study. The results of these analyses follow the discussion of those dealing with the null hypotheses.

## Presentation and Analysis of Data

### $H_O1$

Mean test scores per session, operationalized as mouseclicks, according to score study style are presented in Table 6. Analysis of variance was performed on the dependent variable, test scores per session. The independent variable was score study style. The results of the analysis appear in Table 7.

For the 97 cases, the multiple R was .294. An *F*-ratio of 2.929 was obtained, yielding a *p*-value of 0.038. As a result, the null hypothesis was rejected. There was a significant difference in answers attempted attributable to score study style.

Post-hoc *t*-tests were performed to identify where the differences in means lay. The results of these tests are summarized in Table 8. Significant differences occurred between X1 and X3 ($p = .011$) and X2 and X3 ($p = .018$). Score study with a correct recorded example resulted in significantly higher test scores than either no score study or study with the score alone. There was no significant difference between score study at the piano and any other score study style.

Table 6

<u>Mean and Standard Deviation of Test Scores per Session by Score Study Style</u>

| Score Study Style | n[a] | Mean | Standard Deviation |
|---|---|---|---|
| X1 | 24 | 251.458 | 23.694 |
| X2 | 24 | 255.917 | 21.090 |
| X3 | 26 | 239.038 | 21.564 |
| X4 | 23 | 247.652 | 16.792 |

[a]Number of cases

Table 7

The effects of score study style on test scores

| Source | Sum-of-Squares | df | Mean-Square | *F*-ratio | *p* |
|---|---|---|---|---|---|
| S.S.Style | 3871.473 | 3 | 1290.491 | 2.929 | 0.038* |
| Error | 40969.971 | 93 | 440.537 | | |

$*p<.05$

Table 8

Mean Differences, Standard Deviation Differences, and *t*-tests for Test Scores by Score Study Style

| Styles | n[a] | Mean Difference | Standard Deviation Difference | *t* | df | *p* |
|---|---|---|---|---|---|---|
| X1 vs. X2 | 18 | -0.056 | 19.489 | 0.012 | 17 | .990 |
| X1 vs. X3 | 22 | 13.864 | 23.163 | 2.807 | 21 | .011* |
| X2 vs. X3 | 21 | 13.667 | 24.369 | 2.570 | 20 | .018* |
| X1 vs. X4 | 18 | 2.222 | 27.113 | 0.348 | 17 | .732 |
| X2 vs. X4 | 18 | 10.944 | 24.199 | 1.919 | 17 | .072 |
| X3 vs. X4 | 21 | -9.857 | 30.145 | 1.498 | 20 | .150 |

[a]Number of cases

$*p<.05$

$H_O2$

Mean test scores per session according to order of presentation of score study style are presented in Table 9. Analysis of variance was performed on the dependent variable, test scores per session. The independent variable was score study style order. The results of the analysis appear in Table 10. It should be noted that due to the subjects' failure to follow directions, no score study order could be determined for 14 subjects. The data for these subjects were therefore not included in this analysis.

For the 16 subjects included in the analysis, the multiple R was .573 and an *F*-ratio of 1.951 was obtained yielding a *p*-value of 0.175. The null hypothesis was retained. There was no significant difference in test scores attributable to score study style presentation order.

Table 9

Mean and Standard Deviation of Test Scores per Session by Score Study Style Presentation Order

| Score Study Style Presentation Order | n[a] | Mean | Standard Deviation |
|---|---|---|---|
| 1 (X1, X2, X3, X4) | 4 | 244.438 | 17.222 |
| 2 (X2, X1, X4, X3) | 6 | 256.847 | 14.277 |
| 3 (X3, X4, X1, X2) | 3 | 265.333 | 12.665 |
| 4 (X4, X3, X2, X1) | 3 | 242.389 | 9.419 |

[a]Number of cases

Table 10

The effects of score study style presentation order on test scores

| Source | Sum-of-Squares | df | Mean-Square | *F*-ratio | *p* |
|---|---|---|---|---|---|
| SSOrder | 1173.933 | 3 | 391.311 | 1.951 | 0.175 |
| Error | 2407.164 | 12 | 200.597 | | |

$H_O3$

Mean test scores per session according to example set are presented in Table 11. Analysis of variance was performed on the dependent variable, test scores per session, with example set number as the independent variable. The results of the analysis appear in Table 12. There were 97 cases, and the multiple R was .234. An *F*-ratio of 1.800 was obtained yielding a *p*-value of 0.153. The null hypothesis was retained. There was no significant difference in test scores attributable to the individual tests/example sets.

Table 11

Mean and Standard Deviation of Test Scores per Session by Individual Test/Example Set

| Test/Example Set | n[a] | Mean | Standard Deviation |
|---|---|---|---|
| A | 23 | 251.783 | 21.411 |
| B | 30 | 245.467 | 20.377 |
| C | 20 | 255.750 | 21.205 |
| D | 24 | 242.417 | 22.564 |

[a]Number of cases

Table 12

The effects of example set on test scores

| Source | Sum-of-Squares | df | Mean-Square | *F*-ratio | *p* |
|---|---|---|---|---|---|
| Example Set | 2460.480 | 3 | 820.160 | 1.800 | 0.153 |
| Error | 4230.963 | 93 | 455.709 | | |

$H_O4$

Mean test scores per session according to intact group/order of presentation of individual tests are presented in Table 13. Analysis of variance was performed on the dependent variable, test scores per session, with group number as the independent variable. The results of the analysis appear in Table 14. There were 97 cases, the multiple R was .032, and an *F*-ratio of 0.050 was obtained yielding a *p*-value of 0.952. The null hypothesis was retained: There was no significant difference in test scores attributable to intact groups/order of presentation of example sets.

Table 13

Mean and Standard Deviation of Test Scores per Session by Intact Group/Order of Presentation of Example Set

| Group/Set Order | n[a] | Mean | Standard Deviation |
|---|---|---|---|
| II (B, A, D, C) | 28 | 247.750 | 17.884 |
| III (C, D, A, B) | 22 | 249.591 | 17.546 |
| IV (D, C, B, A) | 47 | 248.085 | 25.407 |

[a]Number of cases

Table 14

The effects of intact group/order of presentation of example sets on test scores

| Source | Sum-of-Squares | df | Mean-Square | *F*-ratio | *p* |
|---|---|---|---|---|---|
| Set Order | 47.216 | 2 | 23.608 | 0.050 | 0.952 |
| Error | 44794.228 | 94 | 476.534 | | |

$\underline{H_O5}$

Mean test scores per session according to number of parts in the examples are presented in Table 15. Analysis of variance was performed on the dependent variable, test scores per example, with the number of parts per example as the independent variable. The results of the analysis appear in Table 16.

There were 776 cases, with a multiple R of .618. An *F*-ratio of 67.815 was obtained, yielding a *p*-value of 0.000; the null hypothesis was rejected. There was a significant difference in test scores attributable to the number of parts in the examples.

Post-hoc *t*-tests were performed to identify where the differences in means lay. The results of these tests are summarized in Table 17. Significant differences occurred between every pair of part numbers except between examples with three parts and those with four and between examples with five parts and those with six.

Table 15

Mean and Standard Deviation of Test Scores per Example by Number of Parts in Example

| Number of Parts | n[a] | Mean | Standard Deviation |
|---|---|---|---|
| 1 | 97 | 6.044 | 0.830 |
| 2 | 97 | 7.327 | 1.238 |
| 3 | 97 | 7.773 | 1.276 |
| 4 | 97 | 7.649 | 1.017 |
| 5 | 97 | 8.456 | 1.323 |
| 6 | 97 | 8.531 | 1.473 |
| 7 | 97 | 9.057 | 1.628 |
| 8 | 97 | 9.478 | 1.415 |

[a]Number of cases

Table 16

The effects of number of parts per example on test scores

| Source | Sum-of-Squares | df | Mean-Square | *F*-ratio | *p* |
|---|---|---|---|---|---|
| No. of Parts | 798.406 | 7 | 114.058 | 67.815 | .000** |
| Error | 1291.691 | 768 | 1.682 | | |

$^{**}p<.01$

Table 17

Mean Differences, Standard Deviation Differences, and *t*-tests for Test Scores by Number of Parts in Examples (97 cases)

| Styles | Mean Difference | Standard Deviation Difference | *t* | df | *p* |
|---|---|---|---|---|---|
| 1 vs. 2 | -1.284 | 1.227 | 10.305 | 96 | .000** |
| 1 vs. 3 | -1.729 | 1.238 | 13.755 | 96 | .000** |
| 2 vs. 3 | -0.445 | 1.489 | 2.945 | 96 | .004* |
| 1 vs. 4 | -1.606 | 1.194 | 13.245 | 96 | .000** |
| 2 vs. 4 | -0.322 | 1.317 | 2.408 | 96 | .018* |
| 3 vs. 4 | -0.123 | 1.519 | 0.797 | 96 | .427 |
| 1 vs. 5 | -5.407 | 1.322 | 17.940 | 96 | .000** |
| 2 vs. 5 | -1.124 | 1.503 | 7.365 | 96 | .000** |
| 3 vs. 5 | -0.679 | 1.670 | 4.002 | 96 | .000** |
| 4 vs. 5 | -0.802 | 1.685 | 4.685 | 96 | .000** |
| 1 vs. 6 | -2.487 | 1.493 | 16.409 | 96 | .000** |
| 2 vs. 6 | -1.204 | 1.601 | 7.405 | 96 | .000** |
| 3 vs. 6 | -0.759 | 1.755 | 4.258 | 96 | .000** |
| 4 vs. 6 | -0.881 | 1.811 | 4.793 | 96 | .000** |
| 5 vs. 6 | -0.080 | 1.819 | 0.433 | 96 | .666 |
| 1 vs. 7 | -3.013 | 1.718 | 17.270 | 96 | .000** |
| 2 vs. 7 | -1.729 | 1.809 | 9.415 | 96 | .000** |
| 3 vs. 7 | -1.284 | 1.769 | 7.152 | 96 | .000** |
| 4 vs. 7 | -1.407 | 1.655 | 8.373 | 96 | .000** |

(table continues)

| Styles | Mean Difference | Standard Deviation Difference | *t* | df | *p* |
|---|---|---|---|---|---|
| 5 vs. 7 | -0.606 | 2.003 | 2.979 | 96 | .004** |
| 6 vs. 7 | -0.526 | 2.118 | 2.445 | 96 | .016* |
| 1 vs. 8 | -3.434 | 1.543 | 21.918 | 96 | .000** |
| 2 vs. 8 | -2.150 | 1.762 | 12.016 | 96 | .000** |
| 3 vs. 8 | -1.705 | 1.624 | 10.342 | 96 | .000** |
| 4 vs. 8 | -1.828 | 1.508 | 11.941 | 96 | .000** |
| 5 vs. 8 | -1.027 | 1.785 | 5.664 | 96 | .000** |
| 6 vs. 8 | -0.947 | 1.839 | 5.071 | 96 | .000** |
| 7 vs. 8 | -0.421 | 1.865 | 2.222 | 96 | .029* |

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

Validation of procedures

In order to validate the procedures used in the study, analysis of mean score study times and mean total times taken per session was undertaken with regard to the effects of score study style, example set, and session number. It was assumed that the time required for score study would be greater when subjects were playing the scores at the keyboard than when hearing correct examples, greater when hearing correct examples while studying the score than with the score alone, and greater when studying the score than when hearing the error example simultaneously with the presentation of the score. Thus it was assumed that differences found in these measures which were attributable to score study style and which followed the trend outlined above (Time X4 > Time X3 > Time X2 > Time X1) would indicate that subjects were indeed using the styles prescribed in each session.

Any differences found in these measures attributable to example set would seem to indicate that the four sets were not of equal difficulty. Differences attributable to session number could be the results of many different factors, including familiarity with the program procedures, subjects' tiring of the program, and increasing facility in the error detection task due to prior sessions.

Mean score study times taken according to score study style are presented in Table 18. Analysis of variance was performed on the dependent variable, score study time per session, with score study style as the independent variable. The results of the analysis appear in Table 19. For the 97 cases, a multiple R of .367 was calculated and an *F*-ratio of 4.837 was obtained, yielding a *p*-value of 0.004. There was a significant difference in score study time according to score study style.

Post-hoc *t*-tests were performed to identify where the differences in means lay. The results of these tests are summarized in Table 20. Significant differences occurred between score study times under X1 (no prestudy) and each of the other three styles. The means for each score study style show an increasing score study time moving from X1 to X2 (score alone) to X3 (score and correct example) to X4 (score at the piano). This indicates that subjects were indeed using the assigned score study styles.

Table 18

Mean and Standard Deviation of Score Study Times by Score Study Style

| Score Study Style | n[a] | Mean | Standard Deviation |
|---|---|---|---|
| X1 | 24 | 4.833 | 2.078 |
| X2 | 24 | 194.542 | 163.561 |
| X3 | 26 | 283.346 | 219.989 |
| X4 | 24 | 356.522 | 630.565 |

[a]Number of cases

Table 19

The effect of score study style on score study time

| Source | Sum-of-Squares | df | Mean-Square | *F*-ratio | *p* |
|---|---|---|---|---|---|
| SSStyle | 1649621.744 | 3 | 549873.915 | 4.837 | .004** |
| Error | 105728E+08 | 93 | 113685.515 | | |

$^{**}p<.01$

Table 20

Mean Differences, Standard Deviation Differences, and *t*-tests for Score Study Times by Score Study Style

| Styles | Mean Difference | Standard Deviation Difference | *t* | df | *p* |
|---|---|---|---|---|---|
| X1 vs. X2 | -183.444 | 164.652 | 4.727 | 17 | .000** |
| X1 vs. X3 | -289.455 | 230.049 | 5.902 | 21 | .000** |
| X2 vs. X3 | -54.333 | 250.170 | 0.995 | 20 | .331 |
| X1 vs. X4 | -365.500 | 689.074 | 2.250 | 17 | .038* |
| X2 vs. X4 | -155.167 | 727.770 | 0.905 | 17 | .378 |
| X3 vs. X4 | -46.333 | 603.395 | 0.352 | 20 | .729 |

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

Mean total times taken per session according to score study style are presented in Table 21. Analysis of variance was performed on the dependent variable, total time per session. The independent variable was score study style. The results of the analysis appear in Table 22. With 97 cases, the multiple R was .196. An *F*-ratio of 1.239 was obtained yielding a *p*-value of 0.300. There were no significant differences in mean total times taken attributable to this factor. However, the means for total time per session according to score study style, presented in Table 21, indicate the same general trend of increased times attributable to score study style as appeared for score study times. This trend would seem to support the conclusion that subjects did indeed use the score study styles indicated for each session.

Table 21

Mean and Standard Deviation of Total Times of Sessions by Score Study Style

| Score Study Style | n[a] | Mean | Standard Deviation |
|---|---|---|---|
| X1 | 24 | 962.583 | 302.819 |
| X2 | 24 | 1044.625 | 316.061 |
| X3 | 26 | 1129.385 | 390.253 |
| X4 | 24 | 1260.696 | 962.580 |

[a]Number of cases

Table 22

The effect of score study style on total time

| Source | Sum-of-Squares | df | Mean-Square | *F*-ratio | *p* |
|---|---|---|---|---|---|
| SSStyle | 1142887.003 | 3 | 380962.334 | 1.239 | .300 |
| Error | 285984E+08 | 93 | 307509.962 | | |

Mean score study times taken per session according to example set are presented in Table 23. Analysis of variance was performed on the dependent variable, score study time per session, with example set as the independent variable. The results of the analysis appear in Table 24.

For the 97 cases a multiple R of .174 was calculated. An *F*-ratio of 0.969 was obtained yielding a *p*-value of 0.411. There were no significant differences in mean score study times attributable to this factor.

Table 23

Mean and Standard Deviation of Score Study Times of Sessions by Example Set

| Example Set | n[a] | Mean | Standard Deviation |
|---|---|---|---|
| A | 23 | 152.565 | 206.757 |
| B | 30 | 171.933 | 216.836 |
| C | 20 | 209.500 | 240.301 |
| D | 24 | 312.292 | 606.156 |

[a]Number of cases

Table 24

The effect of example set on score study time

| Source | Sum-of-Squares | df | Mean-Square | *F*-ratio | *p* |
|---|---|---|---|---|---|
| Example Set | 370471.186 | 3 | 123490.394 | 0.969 | .411 |
| Error | 118519E+08 | 93 | 127439.822 | | |

Mean total times taken per session according to example set are presented in Table 25. Analysis of variance was performed on the dependent variable, total time per session. The independent variable was example set. The results of the analysis appear in Table 26. With 97 cases, the multiple R was .190. The obtained $F$-ratio was 1.166, yielding a $p$-value of 0.327. There were no significant differences in mean total times taken attributable to example set. The lack of significant differences here and in mean score study times taken attributable to example set would seem to indicate that the individual sets were of equivalent difficulty and thus played no significant part in determining session scores.

Table 25

Mean and Standard Deviation of Total Times of Sessions by Example Set

| Example Set | n[a] | Mean | Standard Deviation |
|---|---|---|---|
| A | 23 | 964.609 | 429.285 |
| B | 30 | 1038.300 | 380.263 |
| C | 20 | 1186.300 | 646.769 |
| D | 24 | 1228.000 | 736.267 |

[a]Number of cases

Table 26

The effect of example set on total time

| Source | Sum-of-Squares | df | Mean-Square | *F*-ratio | *p* |
|---|---|---|---|---|---|
| Example Set | 1077697.506 | 3 | 359232.502 | 1.166 | .327 |
| Error | 286636E+08 | 93 | 308210.924 | | |

Mean score study times taken per session according to session number are presented in Table 27. Analysis of variance was performed on the dependent variable, score study time per session, with session number as the independent variable. The results of the analysis appear in Table 28. A multiple R of .236 was calculated for the 97 cases. An *F*-ratio of 1.826 was obtained yielding a *p*-value of 0.148. There were no significant differences in mean score study times taken attributable to session number.

Table 27

Mean and Standard Deviation of Score Study Times of Sessions by Session Number

| Session Number | n[a] | Mean | Standard Deviation |
|---|---|---|---|
| 1 | 30 | 326.800 | 567.109 |
| 2 | 22 | 196.182 | 209.391 |
| 3 | 24 | 165.083 | 198.583 |
| 4 | 21 | 108.095 | 139.270 |

[a]Number of cases

Table 28

The effect of session number on score study time

| Source | Sum-of-Squares | df | Mean-Square | *F*-ratio | *p* |
|---|---|---|---|---|---|
| Session | 679960.944 | 3 | 226653.648 | 1.826 | .148 |
| Error | 115424E+08 | 93 | 124111.975 | | |

Mean total times taken per session according to session number are presented in Table 29. Analysis of variance was performed on the dependent variable, total time per session, with session number as the independent variable. The results of the analysis appear in Table 30. There were 97 cases, and the multiple R was .433. An *F*-ratio of 7.153 was obtained yielding a *p*-value of 0.000. There was a significant difference in total time taken per session according to session number.

Post-hoc *t*-tests were performed to identify where the differences in means lay. The results of these tests are summarized in Table 31. Significant differences occurred between every pair of session numbers except 3 and 4. There was a steady decrease in total time per session from the first through the fourth. This decline may be attributable to the subjects' familiarity with the procedures involved. It should be noted that not all subjects completed four sessions: It is possible that those subjects who did complete them were more adept at the error detection task, although no significant difference was found in test scores attributable to session number. It is also possible that subjects who completed all four sessions were tiring of the program and spent little time considering their answers, although the lack of significant differences in mean score study times attributable to session number would seem to refute this.

Table 29

Mean and Standard Deviation of Total Times of Sessions by Session Number

| Session Number | n[a] | Mean | Standard Deviation |
|---|---|---|---|
| 1 | 30 | 1445.100 | 755.124 |
| 2 | 22 | 1044.136 | 345.178 |
| 3 | 24 | 911.583 | 401.762 |
| 4 | 21 | 872.905 | 265.926 |

[a]Number of cases

Table 30

The effect of session number on total time

| Source | Sum-of-Squares | df | Mean-Square | *F*-ratio | *p* |
|---|---|---|---|---|---|
| Session | 5576226.551 | 3 | 1858742.184 | 7.153 | .000** |
| Error | 241651E+08 | 93 | 259839.644 | | |

$^{**}p<.01$

Table 31

Mean Differences, Standard Deviation Differences, and *t*-tests for Total Times by Session Number

| Session Number | Mean Difference | Standard Deviation Difference | *t* | df | *p* |
|---|---|---|---|---|---|
| 1 vs. 2 | 513.318 | 704.772 | 3.416 | 21 | .003** |
| 1 vs. 3 | 408.667 | 754.543 | 2.653 | 23 | .014* |
| 2 vs. 3 | 123.471 | 227.069 | 2.242 | 16 | .039* |
| 1 vs. 4 | 569.095 | 864.678 | 3.016 | 20 | .007** |
| 2 vs. 4 | 252.143 | 233.048 | 4.048 | 13 | .001** |
| 3 vs. 4 | 76.050 | 323.393 | 1.052 | 19 | .306 |

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

Summary

Several important findings resulted from the data analyses presented above. Significant differences in error detection test scores occurred which were attributable to the score study style implemented in test sessions. Score study with a correct recorded example was found to be significantly more effective than either no study or study with the score alone. There was no significant difference between score study at the keyboard and any of these three score study styles, but the raw data suggested that study at the keyboard might be the second most effective style behind study with a correct example.

There were significant differences in test example scores attributable to the number of parts in the examples. Generally, error detection became increasingly more difficult as the number of parts in the excerpts rose. No significant differences were found attributable to the order of presentation of the four score study styles; the individual example set; the order of presentation of these sets, or intact group; or the number of the session.

Significant differences were found in the mean score study time per session attributable to score study style, and in the mean total time per session attributable to session number. No other significant differences were found on these measures.

# Chapter V

# Summary, Results, Conclusions, and Recommendations for Further Research

## Summary of the design

The purpose of the study was to investigate the effects of four modes of score study on the pitch and rhythm error detection abilities of undergraduate beginning conducting students. The four score study styles were no score study (X1), study with score alone (X2), study with the score and a correct recorded example (X3), and study with score and piano keyboard (X4).

Four universities agreed to participate in the study. Beginning conducting classes at three used the materials in the fall of 1993. The example sets were presented in a different order to each group.

Four sets of examples were developed which varied from four to six measures in length and from one to eight parts. These examples were arranged by the investigator from full-score examples developed for an earlier study of error detection by Ramsey (1979). The examples were then assigned to sets, using a modified random assignment to ensure equivalence between sets.

Each example was then rescored into a specific number of parts according to the difficulty of the originals as reported by Ramsey, with the most difficult being given eight parts. There were 16 examples each of from one to seven parts, four in each set, and 12 examples with eight parts, three in each set.

A set of six Hypercard© stacks was developed for the presentation of these examples and the management of the study: an introductory stack, four example set stacks and a statistics stack. Using a counterbalanced design, the four possible score study styles were entered into a Latin Square. The introductory stack assigned one of the four resulting score study style orders to each subject upon his or her first session. These were assigned on a rotating basis. Each subsequent subject was assigned a different order until all four had been assigned, then the series was repeated.

This stack also initiated the correct score study style for each session according to the order assigned to the subject, and opened the correct example set stack for each session. The order of example sets was the same for every subject in any one group, but different for each group through the application of another Latin Square.

Each example set stack was capable of presenting the materials under every score study condition. Each opened with instructions for the particular score study style assigned for the session and a practice example. The 31 test examples in the set were then presented, and the subject was required to indicate the part, the measure, and the specific note which were incorrect. The program recorded the total time spent on each example, the amount of time spent studying each, and the number of mouseclicks, or attempted answers, taken for each. Upon successfully completing an example, or having reached the limit of allowed answer attempts, the subject was shown the error as it was played before the next example was presented. When the example set was complete, the data recorded during the session were copied to a statistics stack and the example set stack readied for the next subject.

The number of attempted answers for each example was taken as the test score for that example. The total time and study time for each were considered a function of the score study style applied.

The following five subproblems, stated as null hypotheses, were identified for the study:

There will be no difference in test scores on the Tests in Error Detection attributable to:

$H_O1$: score study style.

$H_O2$: order of presentation of score study styles.

$H_O3$: example set/individual test.

$H_O4$: intact group/order of presentation of example sets.

$H_O5$: the number of parts in the items.

Results and discussion

There was a significant difference in test scores attributable to the mode of score study implemented. Study of the score with a correct aural example (X3) was found to be significantly more effective than either no score study

(X1) or study with the score alone (X2). Although no significant difference was found between any of these three styles of score study and score study at the keyboard (X4), the raw data suggested that score study at the keyboard might be the second most effective score study style behind study with the score and a correct recorded example.

These results directly contradict the findings of other investigations of the effects of score study style on error detection ability. This contradiction may be attributable in part to the difference in error types, excerpt types, procedures, and subjects involved in the studies.

Collings (1973) found no significant difference in the error detection abilities of second-semester undergraduate conducting students following training using three methods of score study; no study, score study followed by error examples, and score study followed by correct examples. The subjects in his study received six one-half hour instructional sessions, and only pitch errors were used, with from one to five being randomly inserted into brass quintet and sextet excerpts of six measures' length. Excerpts used on the pretest and posttest were randomly arranged with regard to the number of errors in each.

Grunow (1980) investigated the effects on error detection ability of training featuring four modes of score study--no study, study with the score alone, study with score and a correct recorded example, and study with the recorded example alone--and found no significant differences in the developed error detection abilities. The errors investigated were placed into one of two categories. "General Musical Criteria" included tempo, balance, style of articulation, tone quality, and intonation, while "Specific Technical Criteria" included note accuracy, pitch accuracy, phrasing, dynamic contrast, and ensemble. The low reliability of the "General Musical Criteria" portion of the test ($r$ = .40) and of the overall test ($r$ = .66) call his results into question.

Stwolinski (1988) investigated the differential effects of keyboard sight-reading and listening to recordings on the abilities of undergraduate music students to detect harmonic errors in piano excerpts. The subjects in this study had "demonstrated minimum keyboard proficiency in the third semester of college-level class piano" (p. 84). Four excerpts of from 6 to 12 measures' length and of equivalent difficulty were chosen, and four to five harmonic alterations were inserted into each. Subjects either heard three correct repetitions of each excerpt or sight-read the excerpt at an electronic piano through three times in succession before hearing the error excerpt. All four excerpts were completed within 40 minutes. Subjects marked the printed score on the beats containing perceived errors. Both missed errors and incorrectly identified beats were considered in scoring. The listening condition was found to be significantly more effective in preparing the subjects for the error detection task than the sight-reading condition.

Hopkins (1991) compared the effects of score study with the score alone, with the score at the piano, with the score and a correct recorded example, and sightsinging from the score on the error detection abilities of undergraduate choral music education students. Score study with a correct recorded example was found to be significantly better than score study at the piano, but no other significant differences were identified. The subjects completed one score study session under each condition, with a 50-minute time period allotted for studying the full score of a four-part motet. Subjects heard each error version three times.

Two of the above studies were concerned with error detection training, and used either multiple errors of pitch alone, or multiple error types artificially combined into two categories, in their excerpts. One compared the effects of sightreading of unfamiliar piano excerpts by class piano students of minimal competence with repeated hearings of correct versions of the excerpts. The last, while using a one-shot approach to score study as did the present investigation, allowed 50 minutes for the study of a single whole work of choral music. Further investigation into this area is needed to resolve the contradictions between the findings of these studies and those of the present one.

No significant differences were found in test scores attributable to the order of presentation of score study styles, the individual example set, or the intact group/order of presentation of example sets. This indicates that the example sets were equivalent and that no one order of presentation of either score study styles or example sets resulted in better test scores. The lack of significant differences attributable to intact groups seems to indicate that in this study institutional differences had no effect on test scores. The results of this study may therefore be generalizable across a wide range of institutions.

There was a significant difference between test scores on examples due to the number of parts in the examples. The results of post hoc *t*-tests show significant differences between every pair of part groups except those with three and four parts and those with five and six parts. The test scores on this measure show that subjects found the error detection task increasingly more difficult as the number of parts increased, regardless of score study style, example set or session.

These results corroborate the findings of the only other study that has investigated the effects of the number of parts in musical excerpts on error detection ability. Mount (1982) investigated the differential effects of listening to one voice part, paired parts, and all four parts of a Bach chorale on error detection accuracy among graduate choral music and church music majors. Twenty-five errors were inserted into the chorale, and one phrase was presented at a time. Mount found that there were significant differences in error detection test scores between a) the "parts alone" and "paired parts" listening conditions, b) the "parts alone" and "all voices" conditions, and c) the "paired parts" and "all voices" conditions.

Conclusions and recommendations for further research

In this study, score study at the keyboard and score study with a correct example produced better test scores than study with the score alone and no study. Score study with a correct example was found to be significantly more effective than the latter two, but score study at the keyboard was not. The example sets developed for the study were valid, reliable and equivalent. The results of the study may be generalizable to other undergraduate conducting classes, as there was no significant difference in test scores attributable to intact group. Finally, the number of parts in each example did have an effect on the difficulty of the error detection task. The detection of errors became increasingly difficult as parts were added to excerpts.

Further research into error detection ability is needed. The author suggests research in the following areas:

1. Investigations of the score study styles using both longer and shorter examples and longer, more structured score study times.

2. It was observed in the course of data analysis for this study that at some point during a session the total time and score study time per excerpt began to decrease, where up until that point they had been increasing with the difficulty and complexity of the excerpts. This seemed to indicate that at that point subjects began to tire. Therefore, investigations into the optimum length for tests such as those developed for this study are recommended.
3. An investigation of the interaction between score study styles and both example complexity and error type.
4. An investigation of the effectiveness of these four score study styles when used by subjects with different levels of experience, i.e., undergraduate advanced conducting students and graduate conducting students.
5. A replication of this study involving only instrumental students.
6. The development and testing of a set of HyperCard© stacks for training conducting students in score study techniques at the electronic keyboard.
7. An investigation including percussion and string timbres.

8. An investigation comparing the effects of error detection training with the example sets developed for this study with training using the Ramsey (1979) Program in Error Detection.

Several suggestions for the training of error detection arise from the results of this study. Such training should utilize a variety of timbres, in solo parts and in natural combinations. The use of multiple and increasing parts provides for increasing difficulty, and could be used further to develop such skills. Correct aural examples should be incorporated into error detection training in order to develop score reading and audiation skills: Excerpts without these examples should gradually be introduced, to encourage the development of students' aural imaging.

# Appendix A: Sources of Excerpts

Cacavas, J. (1961). Aria for Winds. New York: Bourne Company. Used by permission.

Erickson, F. (1957). Toccata for Band. New York: Bourne Company. Used by permission.

Erickson, F. (1970). Mexican Folk Rhapsody. Melville, NY: Belwin-Mills Publishing Corporation. Used by permission.

Frangkiser, C. (1957). Hickory Hills. Cleveland, OH: Ludwig Music Publishing Company. Used by permission.

Frangkiser, C. (1960). Monte Vista. Cleveland, OH: Ludwig Music Publishing Company. Used by permission.

Giovanni, C. (1966). Overture in Bb. Palm Desert, CA: Sam Fox Publishing Company. Used by permission.

Gordon, P. (1960). Three Mendelssohn Chorales. New York: Bourne Company. Used by permission.

Grundman, C. (1948). American Folk Rhapsody No. 1. New York: Boosey and Hawkes, Incorporated. Used by permission.

Grundman, C. (1955). Kentucky 1800. New York: Boosey and Hawkes, Incorporated. Used by permission.

Hanson, E. (1954). Pleasant Valley. Cleveland, OH: Ludwig Music Publishing Company. Used by permission.

Maillart, L. A. (1966). The Dragoons of Villars. Cleveland, OH: Ludwig Music Publishing Company. Used by permission.

Sladek, P. and Cantizone, J. (1968). Alcala. Melville, NY: Belwin-Mills, Publishing Corporation. Used by permission.

Appendix B: Sample Copyright Permission Request Letter

July 27, 1992

CPP/Belwin, Inc.
15800 NW 48th Avenue
PO Box 4340
Miami, FL 33014

Dear Sirs:

For my doctoral dissertation, which will investigate the effects of score study style on the error-detection abilities of conducting students, I am writing an evaluation program for Hypercard on the Macintosh computer. In the four example sets for this program, I plan to adapt six to eight non-consecutive four-measure excerpts from the band literature listed below. These examples would consist of from one to eight lines of score. I would appreciate your cooperation in granting me permission to use this piece. The music used in the program will not be reproduced for commercial use, nor will it be reproduced in the microfilm copy of the dissertation. Credit will be given to the composer and to your company in the final document.

Thank you very much for your help in this matter.

Mexican Folk Fantasy - Erickson
Alcala - Sladek, Paul and Cantizone, Joseph
(published by Pro Art Publications)

Yours,

Don R. Crowe

# Appendix C: Example Set Items, Ramsey Program/Test Numbers and Difficulty Coefficients

Example Set A

| Example Number | Number of Lines | Ramsey Number Overall | Ramsey Number Program | Ramsey Number Test | Ramsey Difficulty Coefficient |
|---|---|---|---|---|---|
| 1 (practice) | 2 | | | | |
| 2 | 1 | 9 | 6 | | 1.0000 |
| 3 | 1 | 8 | | 2 | 1.0000 |
| 4 | 1 | 13 | 10 | | .8666 |
| 5 | 1 | 15 | | 3 | .8666 |
| 6 | 2 | 26 | 21 | | .7333 |
| 7 | 2 | 31 | 25 | | .7333 |
| 8 | 2 | 29 | | 5 | .7333 |
| 9 | 2 | 37 | 30 | | .6666 |
| 10 | 3 | 36 | | 6 | .6666 |
| 11 | 3 | 50 | | 8 | .6000 |
| 12 | 3 | 43 | | 7 | .6000 |
| 13 | 3 | 44 | 36 | | .6000 |
| 14 | 4 | 42 | 35 | | .6000 |
| 15 | 4 | 59 | 49 | | .5333 |
| 16 | 4 | 57 | | 9 | .5333 |
| 17 | 4 | 66 | 55 | | .4666 |
| 18 | 5 | 65 | 54 | | .4666 |
| 19 | 5 | 67 | 56 | | .4666 |
| 20 | 5 | 75 | 63 | | .4666 |
| 21 | 5 | 82 | 69 | | .4000 |
| 22 | 6 | 90 | 76 | | .4000 |
| 23 | 6 | 101 | 85 | | .3333 |
| 24 | 6 | 99 | | 15 | .3333 |
| 25 | 6 | 91 | 77 | | .3333 |
| 26 | 7 | 111 | 94 | | .2666 |
| 27 | 7 | 104 | 88 | | .2666 |
| 28 | 7 | 114 | 96 | | .2000 |
| 29 | 7 | 115 | 97 | | .2000 |
| 30 | 8 | 122 | 103 | | .1333 |
| 31 | 8 | 128 | 108 | | .0600 |
| 32 | 8 | 130 | 110 | | .0000 |

Example Set B

| Example Number | Number of Lines | Ramsey Number | | | Ramsey Difficulty Coefficient |
|---|---|---|---|---|---|
| | | Overall | Program | Test | |
| 1 (practice) | 2 | | | | |
| 2 | 1 | 7 | 5 | | 1.0000 |
| 3 | 1 | 3 | 1 | | 1.0000 |
| 4 | 1 | 16 | 12 | | .8666 |
| 5 | 1 | 17 | 13 | | .8666 |
| 6 | 2 | 25 | 20 | | .8000 |
| 7 | 2 | 21 | 17 | | .8000 |
| 8 | 2 | 32 | 26 | | .7333 |
| 9 | 2 | 27 | 22 | | .7333 |
| 10 | 3 | 38 | 31 | | .6666 |
| 11 | 3 | 40 | 33 | | .6666 |
| 12 | 3 | 45 | 37 | | .6000 |
| 13 | 3 | 47 | 39 | | .6000 |
| 14 | 4 | 54 | 45 | | .6000 |
| 15 | 4 | 61 | 51 | | .5333 |
| 16 | 4 | 63 | 53 | | .5333 |
| 17 | 4 | 64 | | 10 | .5333 |
| 18 | 5 | 77 | 65 | | .4666 |
| 19 | 5 | 71 | | 11 | .4666 |
| 20 | 5 | 74 | 62 | | .4666 |
| 21 | 5 | 85 | | 13 | .4000 |
| 22 | 6 | 81 | 68 | | .4000 |
| 23 | 6 | 93 | 78 | | .3333 |
| 24 | 6 | 94 | 79 | | .3333 |
| 25 | 6 | 95 | 80 | | .3333 |
| 26 | 7 | 105 | 89 | | .2666 |
| 27 | 7 | 108 | 91 | | .2666 |
| 28 | 7 | 113 | | 17 | .2000 |
| 29 | 7 | 117 | 99 | | .1333 |
| 30 | 8 | 119 | 101 | | .1333 |
| 31 | 8 | 124 | 105 | | .0600 |
| 32 | 8 | 131 | 111 | | .0000 |

Example Set C

| Example Number | Number of Lines | Ramsey Number | | | Ramsey Difficulty Coefficient |
|---|---|---|---|---|---|
| | | Overall | Program | Test | |
| 1 (practice) | 2 | | | | |
| 2 | 1 | 2 | Practice | | 1.0000 |
| 3 | 1 | 1 | | 1 | 1.0000 |
| 4 | 1 | 14 | 11 | | .8666 |
| 5 | 1 | 12 | 9 | | .8666 |
| 6 | 2 | 18 | 14 | | .8000 |
| 7 | 2 | 23 | 18 | | .8000 |
| 8 | 2 | 24 | 19 | | .7333 |
| 9 | 2 | 28 | 23 | | .7333 |
| 10 | 3 | 34 | 28 | | .6666 |
| 11 | 3 | 53 | 44 | | .6000 |
| 12 | 3 | 48 | 40 | | .6000 |
| 13 | 3 | 49 | 41 | | .6000 |
| 14 | 4 | 51 | 42 | | .6000 |
| 15 | 4 | 56 | 47 | | .5333 |
| 16 | 4 | 55 | 46 | | .5333 |
| 17 | 4 | 69 | 58 | | .4666 |
| 18 | 5 | 73 | 61 | | .4666 |
| 19 | 5 | 68 | 57 | | .4666 |
| 20 | 5 | 76 | 64 | | .4666 |
| 21 | 5 | 86 | 72 | | .4000 |
| 22 | 6 | 83 | 70 | | .4000 |
| 23 | 6 | 98 | 83 | | .3333 |
| 24 | 6 | 96 | 81 | | .3333 |
| 25 | 6 | 92 | | 14 | .3333 |
| 26 | 7 | 107 | 90 | | .2666 |
| 27 | 7 | 110 | 93 | | .2666 |
| 28 | 7 | 109 | 92 | | .2666 |
| 29 | 7 | 121 | 102 | | .1333 |
| 30 | 8 | 118 | 100 | | .1333 |
| 31 | 8 | 123 | 104 | | .0600 |
| 32 | 8 | 129 | 109 | | .0000 |

Example Set D

| Example Number | Number of Lines | Ramsey Number | | | Ramsey Difficulty Coefficient |
|---|---|---|---|---|---|
| | | Overall | Program | Test | |
| 1 (practice) | 2 | | | | |
| 2 | 1 | 6 | 4 | | 1.0000 |
| 3 | 1 | 11 | 8 | | .9333 |
| 4 | 1 | 10 | 7 | | .9333 |
| 5 | 1 | 20 | 16 | | .8000 |
| 6 | 2 | 22 | | 4 | .8000 |
| 7 | 2 | 19 | 15 | | .8000 |
| 8 | 2 | 30 | 24 | | .7333 |
| 9 | 2 | 35 | 29 | | .6666 |
| 10 | 3 | 39 | 32 | | .6666 |
| 11 | 3 | 33 | 27 | | .6666 |
| 12 | 3 | 52 | 43 | | .6000 |
| 13 | 3 | 41 | 34 | | .6000 |
| 14 | 4 | 46 | 38 | | .6000 |
| 15 | 4 | 60 | 50 | | .5333 |
| 16 | 4 | 58 | 48 | | .5333 |
| 17 | 4 | 62 | 52 | | .5333 |
| 18 | 5 | 72 | 60 | | .4666 |
| 19 | 5 | 70 | 59 | | .4666 |
| 20 | 5 | 79 | 66 | | .4000 |
| 21 | 5 | 89 | 75 | | .4000 |
| 22 | 6 | 87 | 73 | | .4000 |
| 23 | 6 | 102 | 86 | | .3333 |
| 24 | 6 | 97 | 82 | | .3333 |
| 25 | 6 | 100 | 84 | | .3333 |
| 26 | 7 | 103 | 87 | | .2666 |
| 27 | 7 | 106 | | 16 | .2666 |
| 28 | 7 | 116 | 98 | | .2000 |
| 29 | 7 | 120 | | 18 | .1333 |
| 30 | 8 | 125 | 106 | | .0600 |
| 31 | 8 | 126 | 107 | | .0600 |
| 32 | 8 | 127 | | 19 | .0600 |

# Appendix D: Scripts and Sample Cards

Scripts from the IntroStack

<u>Stack script of IntroStack</u>

```
on openStack
  hide menubar
  go card 1
  repeat with x = 1 to 6
    put empty into field x
  end repeat
  answer "Copyright © 1993 Error-Detect and its licensors.
All rights reserved."
  put the date into bg field 7
  put the time into bg field 8
  set visible of button "OK" to true
  set visible of button "finished" to false
  hide field 9
  hide field 10
  hide field 11
  Put "Welcome to ERROR-DETECT!." &&¬
  "Fill in the information requested above, using the TAB "
&&¬
  "key to move from one field to the next." &&¬
  "The cursor will flash to show you where you are typing."
&&¬
  "  Be sure to separate your last name, first name, and" &&¬
  "middle initial by pressing the TAB key.  Do NOT put a
period after" &&¬
  " your initial.  Press RETURN or" &&¬
  "click on OK when finished.  (To click on a button, " &&¬
  "move the mouse around until the arrow point or fingertip "
&&¬
  "is on top of the button, then press and release the mouse
button.)"¬
  into field "Instructions"
  select text of field 1
end openStack

on okFin
  if the visible of button "OK" is true
  then
    send mouseUp to button "OK"
  else
    if the visible of button "Finished" is true
    then
      send mouseUp to button "Finished"
    end if
  end if
end okFin
```

Script of Card Button id = 7 "OK" from IntroStack

```
on mouseUp
  set the visible of me to false
  set the visible of button "Finished" to true
  put "Is all the information correct?" &&¬
  "If you need to change anything, press the TAB key" &&¬
  "until either 1) the incorrect information is highlighted,
or 2) " &&¬
  "the cursor flashes on the line you need to fill in (if you
left " &&¬
  "a line blank.)  Then type in the correction. " &&¬
  "When all the information is correct, click on FINISHED. "
&&¬
  "You will then be given INSTRUCTIONS FOR THIS SESSION: read
these " &&¬
  "carefully, as they change for each session."¬
  into field "instructions"

end mouseUp
```

Script of Card button id = 8 "Finished" from IntroStack

```
on mouseUp
  infocheck
end mouseUp

on finished
  session
  if field "Session Number" = 1
  then
    if field "last score study" = 0
    then
      SSOrder1
      add 1 to field "last score study"
    else
      if field "last score study" = 1
      then
        SSOrder2
        add 1 to field "last score study"
      else
        if field "last score study" = 2
        then
          SSOrder3
          add 1 to field "last score study"
        else
          if field "last score study" = 3
          then
            SSOrder4
```

Script of Card button id = 8 "Finished" from IntroStack
(continued)

```
            add 1 to field "last score study"
          else
            if field "last score study" = 4
            then
              SSOrder1
              put 1 into field "last score study"
            end if
          end if
        end if
      end if
    end if
  else
    put field "LastName" into repeatbus1
    put field "FirstName" into repeatbus2
    put field "Initial" into repeatbus3
    go next
    repeat until repeatbus1 = field "LastName" and repeatbus2
      = field "FirstName" and repeatbus3 = field "Initial"
      find repeatbus1
    end repeat
    put field "SSOrder"into whichOrder
    go first
    if whichOrder = 1
    then SSOrder1
    else
      if whichOrder = 2
      then SSOrder2
      else
        if whichOrder = 3
        then SSOrder3
        else
          if whichOrder = 4
          then SSOrder4
        end if
      end if
    end if
  end if
  doMenu "Copy Card"
  doMenu "Paste Card"
  go first
  set visible of me to false
  global exampleNum
  if field "Session Number" = 1
  then
    put "MacintoshHD:HyperCard:hmfinale:D:D" into exampleNum
    go stack "Example Set D"
    put "D" into cd field "Example set"of card "stats"
```

Script of Card button id = 8 "Finished" from IntroStack
(continued)

```
  else
    if field "Session Number" = 2
    then
      put "MacintoshHD:HyperCard:hmfinale:C:C" into
exampleNum
      go stack "Example Set C"
      put "C" into cd field "Example set"of card "stats"
    else
      if field "Session Number" = 3
      then
        put "MacintoshHD:HyperCard:hmfinale:B:B" into
examplenum
        go stack "Example Set B"
        put "B" into cd field "Example set"of card "stats"
      else
        if field "Session Number" = 4
        then
          put "MacintoshHD:HyperCard:hmfinale:A:A" into
exampleNum
          go stack "Example Set A"
          put "A" into cd field "Example set"of card "stats"
        end if
      end if
    end if
  end if

end finished

on SSOrder1
  Put "1" into field "SSOrder"
  global scstustyle
  if field "Session Number" = 1
  then
    put "x1" into scstustyle
    put "x1" into field "scstustyle"
  else
    if field "Session Number" = 2
    then
      put "x2" into scstustyle
      put "x2" into field "scstustyle"
    else
      if field "Session Number" = 3
      then
        put "x3" into scstustyle
        put "x3" into field "scstustyle"
      else
        if field "Session Number" = 4
```

Script of Card button id = 8 "Finished" from IntroStack (continued)

```
        then
          put "x4" into scstustyle
          put "x4" into field "scstustyle"
        end if
      end if
    end if
  end if
end SSOrder1

on SSOrder2
  Put "2" into field "SSOrder"
  global scstustyle
  if field "Session Number" = 1
  then
    put "x2" into scstustyle
    put "x2" into field "scstustyle"
  else
    if field "Session Number" = 2
    then
      put "x1" into scstustyle
      put "x1" into field "scstustyle"
    else
      if field "Session Number" = 3
      then
        put "x4" into scstustyle
        put "x4" into field "scstustyle"
      else
        if field "Session Number" = 4
        then
          put "x3" into scstustyle
          put "x3" into field "scstustyle"
        end if
      end if
    end if
  end if
end SSOrder2

on SSOrder3
  Put "3" into field "SSOrder"
  global scstustyle
  if field "Session Number" = 1
  then
    put "x3" into scstustyle
    put "x3" into field "scstustyle"
  else
    if field "Session Number" = 2
    then
```

Script of Card button id = 8 "Finished" from IntroStack (continued)

```
      put "x4" into scstustyle
      put "x4" into field "scstustyle"
    else
      if field "Session Number" = 3
      then
        put "x1" into scstustyle
        put "x1" into field "scstustyle"
      else
        if field "Session Number" = 4
        then
          put "x2" into scstustyle
          put "x2" into field "scstustyle"
        end if
      end if
    end if
  end if
end SSOrder3

on SSOrder4
  Put "4" into field "SSOrder"
  global scstustyle
  if field "Session Number" = 1
  then
    put "x4" into scstustyle
    put "x4" into field "scstustyle"
  else
    if field "Session Number" = 2
    then
      put "x3" into scstustyle
      put "x3" into field "scstustyle"
    else
      if field "Session Number" = 3
      then
        put "x2" into scstustyle
        put "x2" into field "scstustyle"
      else
        if field "Session Number" = 4
        then
          put "x1" into scstustyle
          put "x1" into field "scstustyle"
        end if
      end if
    end if
  end if
end SSOrder4
```

Script of Card button id = 8 "Finished" from IntroStack
(continued)

```
on session
  global statsname
  put field "LastName" && field "FirstName" && field
"Initial"¬
  into statsname
  global statssessnum
  put field "Session Number" into statssessnum
  global statsdate
  put the date into statsdate

  global statsInstructor
  put field "Instructor" into statsInstructor
  global statsClass
  put field "Class" into statsClass
  global statsSSOrder
  put field "SSOrder" into statsSSOrder

end session


on infoCheck
  if field 1 ≠ empty and field 2 ≠ empty and field 3 ≠ empty
and¬
  field 4 ≠ empty and field 5 ≠ empty and field 6 ≠ empty
  then
    finished
  else
    beep two
    repeat with x = 1 to 6
      if field x = empty
      then
        select text of field x
      end if
    end repeat
    put "Please fill in all the information, then click on"¬
    && "FINISHED or press RETURN." into bg field
            "instructions"
  end if
end infoCheck
```

Script of background field 1 "LastName" from IntroStack

```
on returnInField
  if field "LastName" is empty
  then
    beep two
```

Script of background field id = 15 "FirstName" from IntroStack

```
    put "Please fill in all the information, then click on"¬
    && "Finished or press RETURN." into bg field
          "instructions"
  else
    okFin
  end if
end returnInField


on returnInField
  if field "FirstName" is empty
  then
    beep two
    put "Please fill in all the information, then click on"¬
    && "Finished or press RETURN." into bg field
         "instructions"
  else
    okFin
  end if
end returnInField
```

Script of background field id = 16 "Initial" from IntroStack

```
on returnInField
  if field "Initial" is empty
  then
    beep two
    put "Please fill in all the information, then click on"¬
    && "Finished or press RETURN." into bg field
         "instructions"
  else
    okFin
  end if
end returnInField
```

Script of background field id = 11 "Time" from IntroStack

```
on returnInField
  send mouseUp to button "OK"
end returnInField
```

Script of background field id 5 = "Instructor" from IntroStack

```
on returnInField
  if field "Instructor" is empty
  then
    beep two
    put "Please fill in all the information, then click on"¬
    && "Finished or press RETURN." into bg field
        "instructions"
  else
    okFin
  end if
end returnInField
```

Script of background field id 6 = "Class" from IntroStack

```
on returnInField
  if field "Class" is empty
  then
    beep two
    put "Please fill in all the information, then click on"¬
    && "Finished or press RETURN." into bg field
        "instructions"
  else
    okFin
  end if
end returnInField
```

Script of background field id 8 = "SessionNumber" from IntroStack

```
on returnInField
  if field "Session Number" is empty
  then
    beep two
    put "Please fill in all the information, then click on"¬
    && "Finished or press RETURN." into bg field
        "instructions"
  else
    okFin
  end if
end returnInField
```

Script of background field id 13 = "last score study" from IntroStack*

```
on returnInField
  send mouseUp to button "OK"
end returnInField
```

Script of background field id 12 = "ScStuStyle" from IntroStack*

```
on returnInField
  send mouseUp to button "OK"
end returnInField
```

Script of background field id 14 = "SSOrder" from IntroStack*

```
on returnInField
  send mouseUp to button "OK"
end returnInField
```

*These fields were hidden from the subjects

Scripts from the Example Set Stacks

Stack Script of "Example Set D"

```
on openStack
  hide button "READY" of card "instructions"
  put "Please hold the line: we will be right with you!"¬
  into field 1 of card "instructions"
  get msg
  if it ≠ "DonC"
  then
    repeat with x = 1 to 9
      repeat with y = 1 to 32
        put "0" into line y of field x of card "stats"
      end repeat
    end repeat
    set userlevel to 3
    disable menu "FILE"
    disable menu "EDIT"
    disable menu "GO"
    hide menubar
    global statsName
    put statsName into cd field "name" of card "stats"
    global statssessnum
    put statssessnum into cd field "Session Number" of card
        "stats"
    global statsDate
    put statsDate into cd field "Date" of card "stats"
    put the long time into cd field "Start Time" of card
        "stats"
    global statsInstructor
    put statsInstructor into cd field "Instructor" of card
        "stats"
    global statsClass
    put statsClass into cd field "Class" of card "stats"
    global statsSSOrder
    put statsSSOrder into cd field "SSOrder" of card "stats"
  else
    set userLevel to 5
  end if
  hmOpenMIDI
  hmPatcher "connect",card field "connections" of card 2
  hide card field "connections" of card 2
  global questnum
  put 0 into questnum
  global scstustyle
  if scstustyle = x1
  then x1
  else
    if scstustyle = x2
    then x2
    else
```

Script from Stack "Example Set D" (continued)

```
      if scstustyle = x3
      then x3
      else
        if scstustyle = x4
        then x4
      end if
    end if
  end if
end openStack

on openCard
  get the short name of this background
  if it = "Statistics"
  then
    get msg
    if msg ≠ DonC
    then
      lock screen
      enable menu "FILE"
      enable menu "EDIT"
      enable menu "GO"
      doMenu "Copy Card"
      go "hmiditeststats"
      set userLevel to 3
      doMenu "Paste Card"
      show menubar
      go home
      unlock screen
    end if
  end if
  global clicks
  put 0 into clicks
  global totalclicks
  put 0 into totalclicks
  global temp
  put the secs into temp
  global temptotal
  put the secs into temptotal
  global questnum
  put 0 into questnum

end openCard

on mouseUp
  beep one
  put "No, that is incorrect.  Try again." into bg field 2
  global clicks
  add 1 to clicks
```

Script from Stack "Example Set D" (continued)

```
   global totalclicks
   add 1 to totalclicks
   if totalclicks ≥ 10
   then
     put "Sorry, but your time is up: here is the error " &&¬
     "as it was played.  "Please click on the NEXT button" &&¬
     "to go on." into field 2
     set the visible of cd button "errornote" to false
     show bg button "Next"
   end if
end mouseUp

on x1
  put "INSTRUCTIONS FOR THIS SESSION:" &&¬
  "You will be shown a score and will hear at the same time
a" &&¬
  "performance of that score which has one error of either
pitch" &&¬
  "or rhythm.  Once the performance is complete, you will be
asked" &&¬
  "to click on:" into line 1 of field 1 of card
"instructions"
  put "          1) the part; then" into line 3 of field 1 of
card "instructions"
  put "          2) the measure; and last" into line 4 of
field 1 of card "instructions"
  put "          3) the note" into line  5 of field 1 of card
"instructions"
  put "which was played incorrectly.  Be sure that the tip of
the " &&¬
  "cursor is on top of the item which you wish to indicate.
" &&¬
  "When you have correctly identified" &&¬
  "the error, you will see the error as it was played.  Then
you may go to the next" &&¬
  "example by clicking on NEXT. " &&¬
  "Do not click on a button until the computer finishes
playing the example.  " &&¬
  "The session begins with one-line scores and progresses"
&&¬
  "to eight-line scores, with a total of 32 examples.  At the
end " &&¬
  "of the session, the program will automatically return to
the " &&¬
  "HOME stack.  When you are ready" &&¬
  "to begin, click on READY."¬
  into line 7 of field 1 of card "instructions"
  put "x1" into cd field "scstustyle" of card "stats"
```

Script from Stack "Example Set D" (continued)

```
  show button "READY" of card "instructions"
end x1

on x2
  put "INSTRUCTIONS FOR THIS SESSION:" &&¬
  "You will be shown a score and given time to study it.
When" &&¬
  "you are ready, click on the PLAY button and you will hear
a" &&¬
  "performance of that score which has one error of either
pitch" &&¬
  "or rhythm.  Once the performance is complete, you will be
asked" &&¬
  "to click on:" into line 1 of field 1 of card
"instructions"
  put "          1) the part; then" into line 3 of field 1 of
card "instructions"
  put "          2) the measure; and last" into line 4 of
field 1 of card "instructions"
  put "          3) the note" into line  5 of field 1 of card
"instructions"
  put "which was played incorrectly.  Be sure that the tip of
the " &&¬
  "cursor is on top of the item which you wish to indicate.
" &&¬
  "When you have correctly identified" &&¬
  "the error, you will see the error as it was played.  Then
you may go to the next" &&¬
  "example by clicking on NEXT. " &&¬
  "Do not click on a button until the computer finishes
playing the example.  " &&¬
  "The session begins with one-line scores and progresses"
&&¬
  "to eight-line scores, with a total of 32 examples.  At the
end " &&¬
  "of the session, the program will automatically return to
the " &&¬
  "HOME stack.  When you are ready" &&¬
  "to begin, click on READY."¬
  into line 7 of field 1 of card "instructions"
  put "x2" into cd field "scstustyle" of card "stats"
  show button "READY" of card "instructions"
end x2

on x3
  put "INSTRUCTIONS FOR THIS SESSION:" &&¬
  "You will be shown a score and hear it played correctly.
When you are ready, " &&¬
```

Script from Stack "Example Set D" (continued)

```
   "click on the PLAY button and you will hear a" &&¬
   "performance of that score which has one error of either
 pitch" &&¬
   "or rhythm.  (Do not click on a button until the computer "
 &&¬
   "has finished playing the example.)  Once the performance
 is complete, " &&¬
   "you will be asked to click on:"¬
   into line 1 of field 1 of card "instructions"
   put "          1) the part; then" into line 3 of field 1 of
 card "instructions"
   put "          2) the measure; and last" into line 4 of
 field 1 of card "instructions"
   put "          3) the note" into line  5 of field 1 of card
 "instructions"
   put "which was played incorrectly.  Be sure that the tip of
 the " &&¬
   "cursor is on top of the item which you wish to indicate.
 " &&¬
   "When you have correctly identified" &&¬
   "the error, you will see the error as it was played.  Then
 you may go to the next" &&¬
   "example by clicking on NEXT. " &&¬
   "The session begins with one-line scores and progresses"
 &&¬
   "to eight-line scores, with a total of 32 examples.  At the
 end " &&¬
   "of the session, the program will automatically return to
 the " &&¬
   "HOME stack.  When you are ready" &&¬
   "to begin, click on READY."¬
   into line 7 of field 1 of card "instructions"
   put "x3" into cd field "scstustyle" of card "stats"
   show button "READY" of card "instructions"
 end x3

 on x4
   put "INSTRUCTIONS FOR THIS SESSION:" &&¬
   "You will be shown a score and given time to play it on"
 &&¬
   "the keyboard.  When you are ready, click on the PLAY" &&¬
   "button and you will hear a" &&¬
   "performance of that score which has one error of either
 pitch" &&¬
   "or rhythm.  Once the performance is complete, you will be
 asked" &&¬
   "to click on:" into line 1 of field 1 of card
 "instructions"
```

Script from Stack "Example Set D" (continued)

```
  put "          1) the part; then" into line 3 of field 1 of
card "instructions"
  put "          2) the measure; and last" into line 4 of
field 1 of card "instructions"
  put "          3) the note" into line  5 of field 1 of card
"instructions"
  put "which was played incorrectly.  Be sure that the tip of
the " &&¬
  "cursor is on top of the item which you wish to indicate.
" &&¬
  "When you have correctly identified" &&¬
  "the error, you will see the error as it was played.  Then
you may go to the next" &&¬
  "example by clicking on NEXT. " &&¬
  "Do not click on a button until the computer finishes
playing the example.  " &&¬
  "The session begins with one-line scores and progresses"
&&¬
  "to eight-line scores, with a total of 32 examples.  At the
end " &&¬
  "of the session, the program will automatically return to
the " &&¬
  "HOME stack.  When you are ready" &&¬
  "to begin, click on READY."¬
  into line 7 of field 1 of card "instructions"
  put "x4" into cd field "scstustyle" of card "stats"
  show button "READY" of card "instructions"
end x4

on tabkey
  answer "That key does not function in this application."
end tabkey
```

Script of card id 6969 = "Instructions" from card 1 of stack "Example Set D"

```
on closeCard
  put the long time into cd field "Start Time" of card
"Stats"
end closeCard
```

Script of card button id 1 = "Ready" from card 1 of stack "Example Set D"

```
on mouseUp
  put empty into field 1
  go next
end mouseUp
```

Script of background id 2787 = "Examples" from stack "Example Set D"

```
on openCard
  global temptotal
  put the secs into temptotal
  global temp
  put the secs into temp
  global questnum
  put "0" into questnum
  global statssessnum
  put "Session" && statssessnum & ":" && the short name of
this stack into bg field 1
  Global scstustyle
  if scstustyle = x1
  then
    hide bg button "play"
    wait 10
    play
  else
    if scstustyle = x3
    then
      wait 15
      get the number of this card
      subtract one from it
      global exampleNum
      put exampleNum & it & "C" into fileName
      if fileName is not empty then
        go to this card
        put empty into sequence
        put hmMIDIfile("read",fileName,"msec") into sequence
        if char 1 to 5 of sequence is "Error"
        then
          answer sequence with "OK"
          put empty into sequence
        else
          hmWriteMIDI sequence
        end if
      end if
      put "Click the PLAY button to hear the error example."
into bg field 2
```

Script of background id 2787 = "Examples" from stack "Example Set D" (continued)

```
      show bg button "play"
    else
      put "Click the PLAY button to hear the error example."
into bg field 2
      show bg button "play"
    end if
  end if
end openCard


on mouseUp
  beep one
  put "No, that is incorrect.  Try again." into bg field 2
  global clicks
  add 1 to clicks
  global totalclicks
  add 1 to totalclicks
  if totalclicks ≥ 10
  then
    put "Sorry, but your time is up: here is the error as it
was played. " &&¬
    "Please click on the NEXT button to go on."¬
    Into field 2
    set the visible of cd button "errornote" to false
    show bg button "Next"

  end if
end mouseUp

on closeCard
  put empty into bg field 2
end closeCard

on play
  global temp
  put the secs - temp into line (number of this card - 1) of
bg field "studytime" of card "stats"
  put the secs into temp
  get the number of this card
  subtract one from it
  global exampleNum
  put exampleNum & it into fileName
  if fileName is not empty then
    go to this card
    put empty into sequence
    put hmMIDIfile("read",fileName,"msec") into sequence
    if char 1 to 5 of sequence is "Error"
```

Script of background id 2787 = "Examples" from stack "Example Set D" (continued)

```
    then
      answer sequence with "OK"
      put empty into sequence
    else
      hmWriteMIDI sequence
      put "In which part did the error occur?" into bg field
2
      global totalclicks
      add 1 to totalclicks
    end if
  end if
end play
```

Script of card id 3926 = "ExampleD1" from card 2 of stack "Example Set D"

```
on openCard
  set the visible of cd button "errornote" to true
  global clicks
  put 0 into clicks
  global totalclicks
  put 0 into totalclicks
  global temp
  put the secs into temp
  global temptotal
  put the secs into temptotal
  global questnum
  put 0 into questnum
  global played
  put 0 into played
  pass openCard
end openCard

on mouseUp
  pass mouseUp
end mouseUp
```

Script of background button id 4 = "Play" from card 2 of stack "Example Set D"

```
on mouseUp
  global scstustyle
  if scstustyle ≠ x1
  then
    global temp
```

Script of background button id 4 = "Play" from card 2 of stack "Example Set D" (continued)

```
    put the secs - temp into line (number of this card - 1)
of ¬
    bg field "studytime" of card "stats"
    put the secs into temp
    get the number of this card
    subtract one from it
    global exampleNum
    put exampleNum & it into fileName
    if fileName is not empty then
      go to this card
      put empty into sequence
      put hmMIDIfile("read",fileName,"msec") into sequence
      if char 1 to 5 of sequence is "Error" then
        answer sequence with "OK"
        put empty into sequence
        exit mouseUp
      end if
      hmWriteMIDI sequence
    end if
    hide me
    put "In which part did the error occur?" into bg field 2
    global played
    add 1 to played
    global totalclicks
    add 1 to totalclicks
  else
    pass mouseup
  end if
end mouseUp
```

Script of card button id 11 = "rightPart" from card 2 of stack "Example Set D"

```
on mouseUp

  global questNum
  get questNum
  if questNum = 0 then
    add 1 to it
    put it into questNum
    global temp
    put the secs - temp into line (number of this card - 1)
of field
1 of card "stats"
    put the secs into temp
    global clicks
```

Script of card button id 11 = "rightPart" from card 2 of stack "Example Set D" (continued)

```
    add 1 to clicks
    put clicks into line (number of this card - 1) of field 5
of card "stats"
    put 0 into clicks
    global totalclicks
    add 1 to totalclicks
    put "Very good! In which measure did the error occur?"
into bg field 2

  else
    pass mouseUp
  end if
end mouseUp
```

Script of card button id 12 = "rightMeasure" from card 2 of stack "Example Set D"

```
on mouseUp

  global questNum
  get questNum
  if questNum = 0 then send mouseUp to cd button "rightPart"
  else
    if questNum = 1 then
      add 1 to it
      put it into questNum
      global temp
      put the secs - temp into line (number of this card - 1)
of field 2 of card "stats"
      put the secs into temp
      global clicks
      add 1 to clicks
      put clicks into line (number of this card - 1) of field
6 of card "stats"
      put 0 into clicks
      global totalclicks
      add 1 to totalclicks
      put "Very good!  Now, which specific note was
incorrect?" into bg field 2

    else
      if questNum = 2 then pass mouseUp
    end if
  end if
end mouseUp
```

Script of card button id 13 = "rightNote" from card 2 of stack "Example Set D"

```
on mouseUp

  global questNum
  get questNum
  if questNum = 0 then send mouseUp to cd button "rightPart"
  else
    if questNum = 1 then send mouseUp to cd button
"rightMeasure"
    else
      if questNum = 2 then
        put "Right! Here is the error as it was played.
Click on the NEXT button to see the next example." into bg
field 2
        set visible of cd button "errornote" to false
        set the visible of bg button "NEXT" to true
        put 0 into questNum
        global temp
        put the secs - temp into line (number of this card -
1) of field 3 of card "stats"
        put the secs into temp
        global clicks
        add 1 to clicks
        put clicks into line (number of this card - 1) of
field 7 of card "stats"
        put "0" into clicks
        global totalclicks
        add 1 to totalclicks
      end if
    end if
  end if
end mouseUp
```

Script of card button id 11 = "NEXT" from card 2 of stack "Example Set D"

```
on mouseUp
  set the visible of cd button "errornote" to true
  set the visible of me to false
  global temptotal
  put the secs - temptotal into line (number of this card -
1) of field 4 of card "stats"
  put the secs into temptotal
  global totalclicks
  add 1 to totalclicks
```

<u>Script of card button id 11 = "NEXT" from card 2 of stack "Example Set D"</u> (continued)

```
  put totalclicks into line (number of this card - 1) of
field 8 of card "stats"
  put 0 into totalclicks
  global clicks
  put 0 into clicks
  global questnum
  put 0 into questnum

  go next
end mouseUp
```

<u>Script of card id 18136 = "stats" from card 34 of stack "Example Set D"</u>

```
on openCard
  get msg
  if msg ≠ "DonC"
  then
    put the long time into cd field "end time"
    pass openCard
  end if
end openCard
```

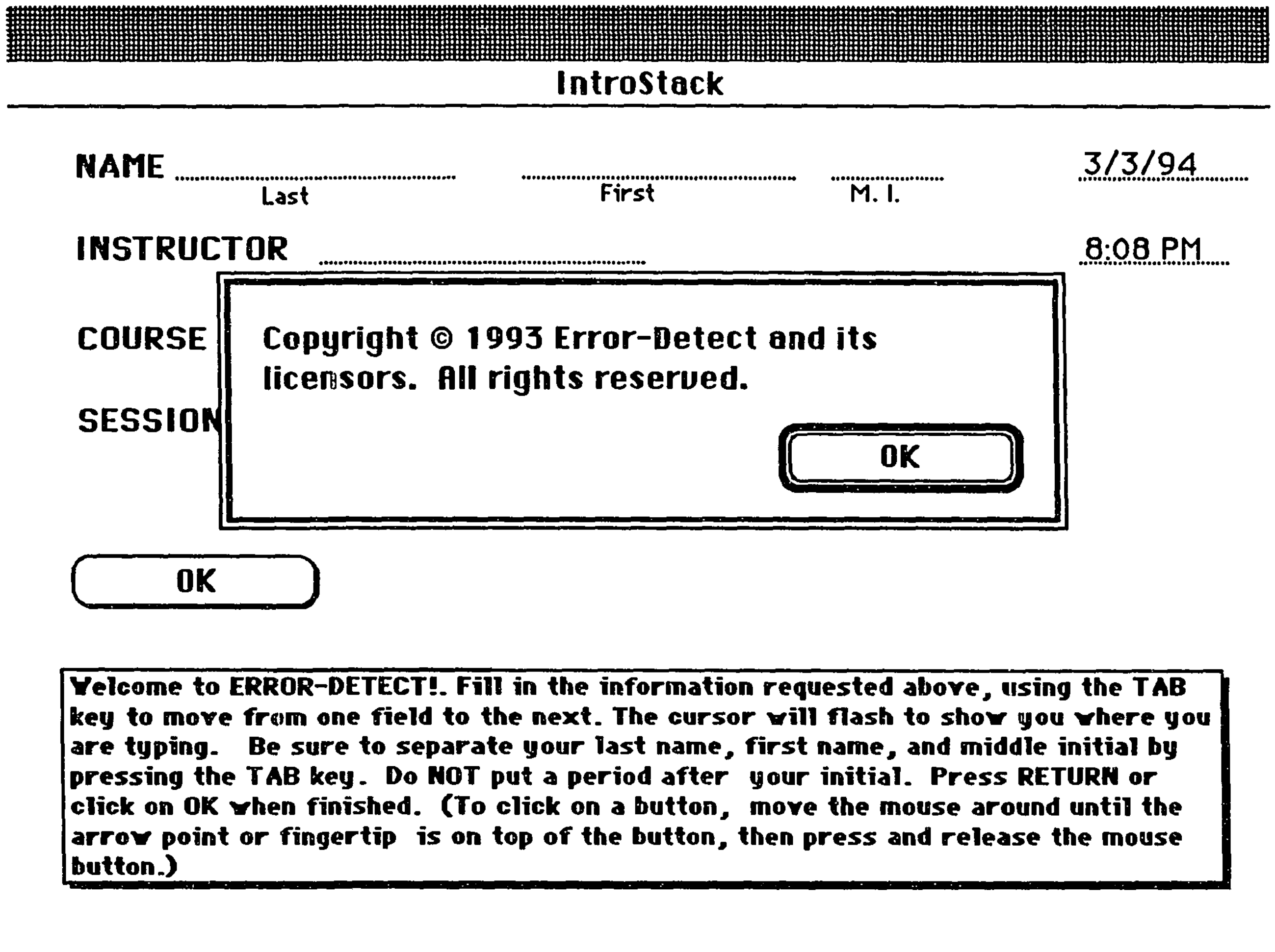
Script from Statistics Stack

Script of stack "hmiditeststats"

```
on openStack
  go first
end openStack
```

Sample Cards

With four exceptions, the cards herein are presented as they would be if all were necessary, that is, if the subject were to make a mistake. First, all four score study style instructions are presented in order, where in actual use only one set of instructions would appear per session. Second, the practice card is shown with the "PLAY" button; in the "no study" condition this button was hidden and disabled. Third, only one excerpt, the practice, is shown before the statistics card. Finally, the buttons which register a subject's correct answers are shown on these cards in outline. These outlines are not visible when the program is used, although the buttons are highlighted when clicked.


IntroStack
NAME
Last
First
M. I.
3/3/94
INSTRUCTOR
8:08 PM
COURSE
SESSION
Copyright © 1993 Error-Detect and its licensors. All rights reserved.
OK
OK
Welcome to ERROR-DETECT!. Fill in the information requested above, using the TAB key to move from one field to the next. The cursor will flash to show you where you are typing. Be sure to separate your last name, first name, and middle initial by pressing the TAB key. Do NOT put a period after your initial. Press RETURN or click on OK when finished. (To click on a button, move the mouse around until the arrow point or fingertip is on top of the button, then press and release the mouse button.)

IntroStack

NAME ______ ______ ______ 3/3/94
Last First M. I.

INSTRUCTOR ______ 9:04 PM

COURSE SECTION/TIME ______

SESSION NUMBER (1,2,3, or 4) ______

OK

Welcome to ERROR-DETECT!. Fill in the information requested above, using the TAB key to move from one field to the next. The cursor will flash to show you where you are typing. Be sure to separate your last name, first name, and middle initial by pressing the TAB key. Do NOT put a period after your initial. Press RETURN or click on OK when finished. (To click on a button, move the mouse around until the arrow point or fingertip is on top of the button, then press and release the mouse button.)

## IntroStack

NAME ______ ______ ______ 3/3/94
Last First M. I.

INSTRUCTOR ______ 9:19 PM

COURSE SECTION/TIME ______

SESSION NUMBER (1,2,3, or 4) ______

Finished

Is all the information correct? If you need to change anything, press the TAB key until either 1) the incorrect information is highlighted, or 2) the cursor flashes on the line you need to fill in (if you left a line blank.) Then type in the correction. When all the information is correct, click on FINISHED. You will then be given INSTRUCTIONS FOR THIS SESSION: read these carefully, as they change for each session.

## IntroStack

**NAME** ............................ ............................ ............ 3/3/94

Last First M. I.

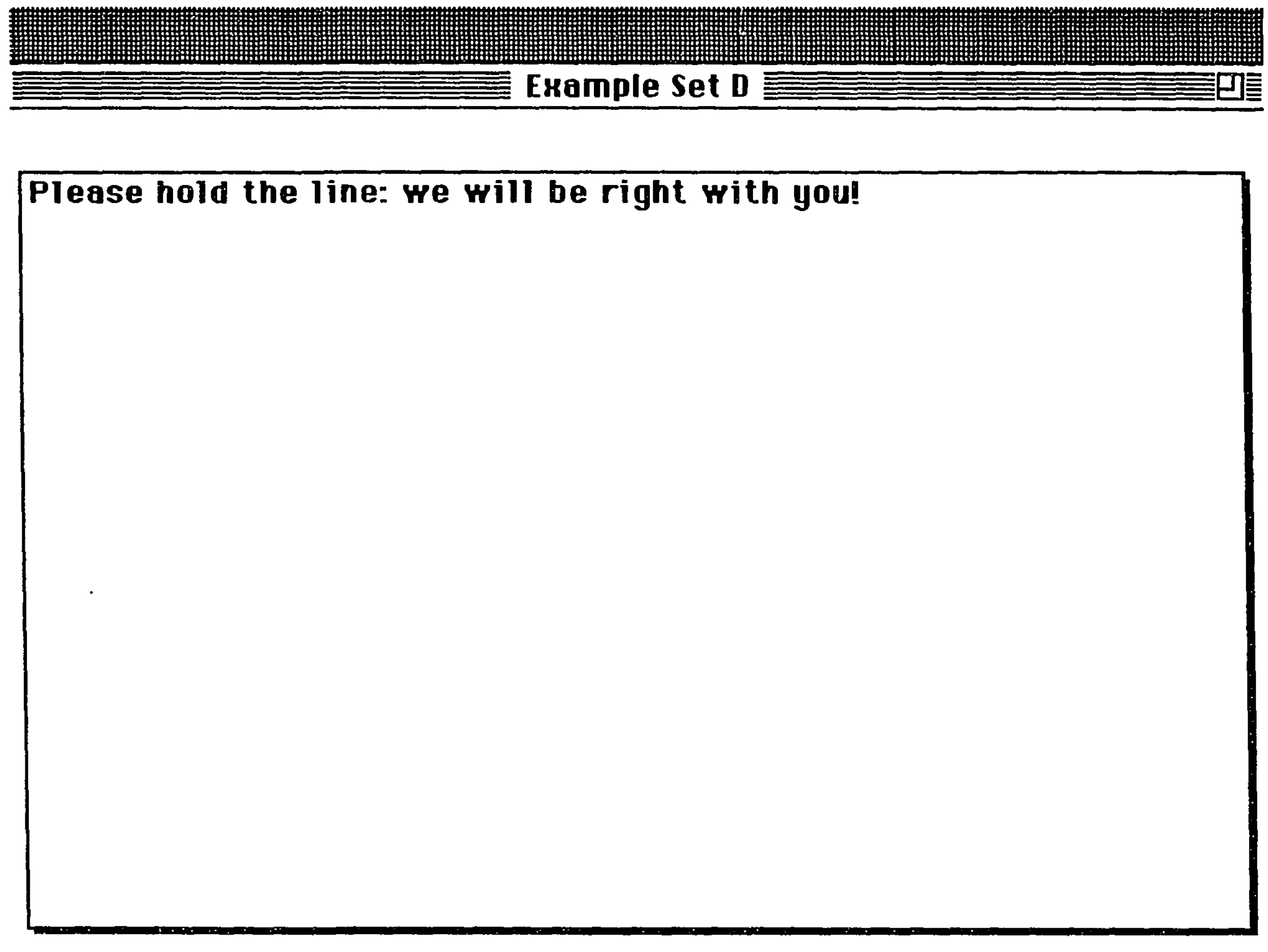
**INSTRUCTOR** ............................................ 9:22 PM

**COURSE SECTION/TIME** ......................

**SESSION NUMBER (1,2,3, or 4)** ......

**Finished**

**Please fill in all the information, then click on FINISHED or press RETURN.**


Example Set D
Please hold the line: we will be right with you!

Example Set D

Ready

INSTRUCTIONS FOR THIS SESSION: You will be shown a score and will hear at the same time a performance of that score which has one error of either pitch or rhythm. Once the performance is complete, you will be asked to click on:

1) the part; then
2) the measure; and last
3) the note

which was played incorrectly. Be sure that the tip of the cursor is on top of the item which you wish to indicate. When you have correctly identified the error, you will see the error as it was played. Then you may go to the next example by clicking on NEXT. Do not click on a button until the computer finishes playing the example. The session begins with one-line scores and progresses to eight-line scores, with a total of 32 examples. At the end of the session, the program will automatically return to the HOME stack. When you are ready to begin, click on READY.

Example Set D

Ready

INSTRUCTIONS FOR THIS SESSION: You will be shown a score and given time to study it. When you are ready, click on the PLAY button and you will hear a performance of that score which has one error of either pitch or rhythm. Once the performance is complete, you will be asked to click on:

1) the part; then
2) the measure; and last
3) the note

which was played incorrectly. Be sure that the tip of the cursor is on top of the item which you wish to indicate. When you have correctly identified the error, you will see the error as it was played. Then you may go to the next example by clicking on NEXT. Do not click on a button until the computer finishes playing the example. The session begins with one-line scores and progresses to eight-line scores, with a total of 32 examples. At the end of the session, the program will automatically return to the HOME stack. When you are ready to begin, click on READY.

Example Set D

Ready

INSTRUCTIONS FOR THIS SESSION: You will be shown a score and hear it played correctly. When you are ready, click on the PLAY button and you will hear a performance of that score which has one error of either pitch or rhythm. (Do not click on a button until the computer has finished playing the example.) Once the performance is complete, you will be asked to click on:

1) the part; then
2) the measure; and last
3) the note

which was played incorrectly. Be sure that the tip of the cursor is on top of the item which you wish to indicate. When you have correctly identified the error, you will see the error as it was played. Then you may go to the next example by clicking on NEXT. The session begins with one-line scores and progresses to eight-line scores, with a total of 32 examples. At the end of the session, the program will automatically return to the HOME stack. When you are ready to begin, click on READY.

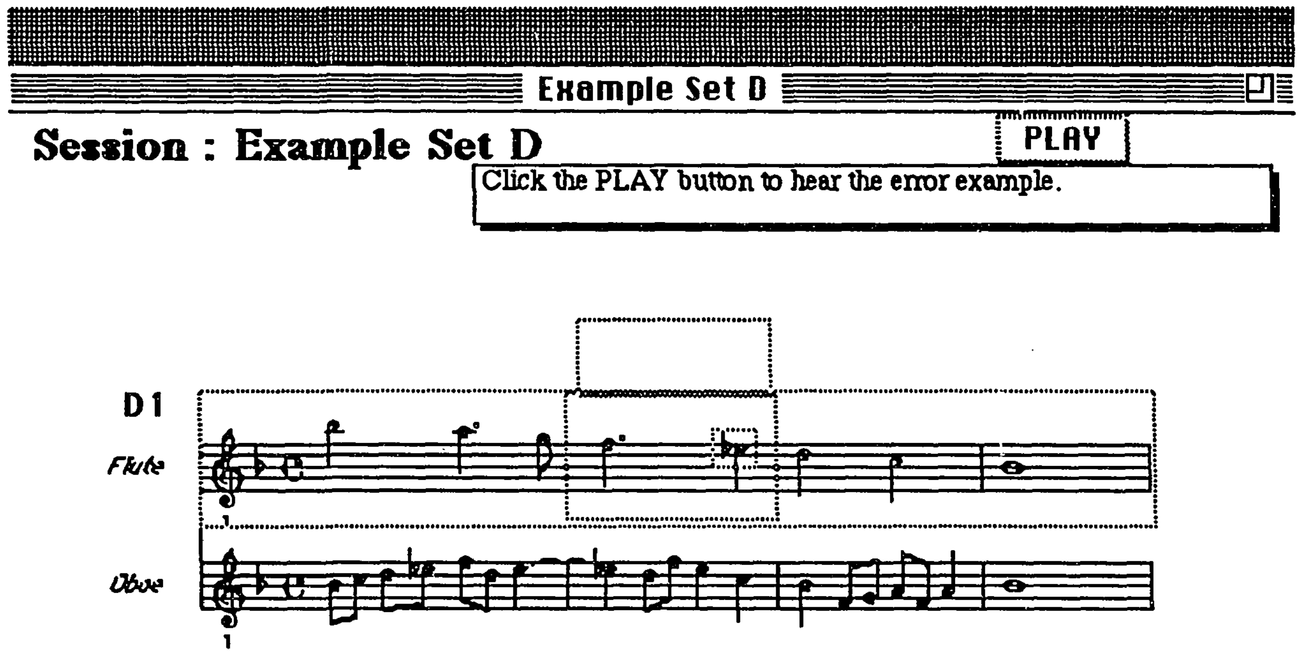
**Example Set D**

Ready

INSTRUCTIONS FOR THIS SESSION: You will be shown a score and given time to play it on the keyboard. When you are ready, click on the PLAY button and you will hear a performance of that score which has one error of either pitch or rhythm. Once the performance is complete, you will be asked to click on:

1) the part; then
2) the measure; and last
3) the note

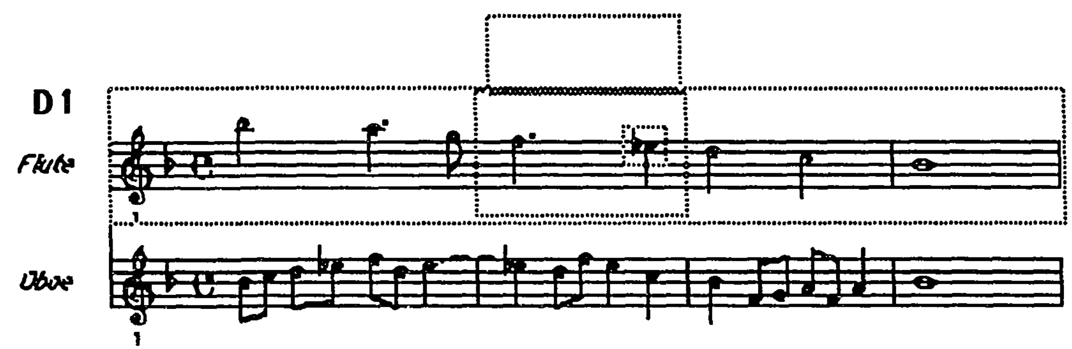
which was played incorrectly. Be sure that the tip of the cursor is on top of the item which you wish to indicate. When you have correctly identified the error, you will see the error as it was played. Then you may go to the next example by clicking on NEXT. Do not click on a button until the computer finishes playing the example. The session begins with one-line scores and progresses to eight-line scores, with a total of 32 examples. At the end of the session, the program will automatically return to the HOME stack. When you are ready to begin, click on READY.


Example Set D
Session : Example Set D
PLAY
Click the PLAY button to hear the error example.
D1
Flute
Oboe

Example Set D

# Session : Example Set D

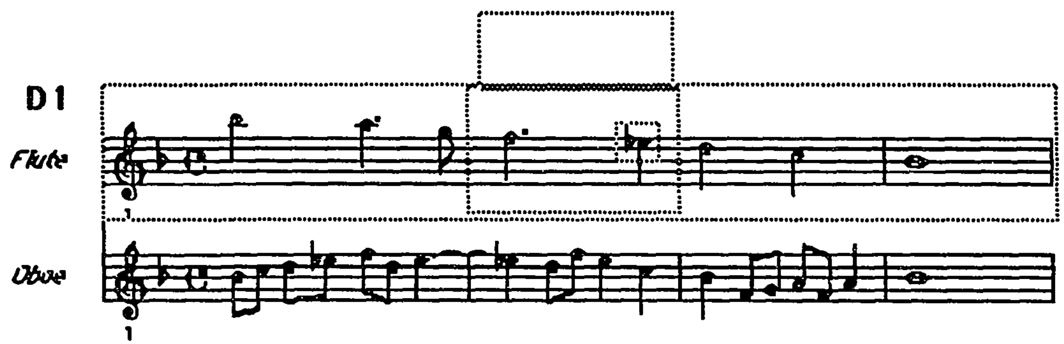
In which part did the error occur?



Example Set D

# Session : Example Set D

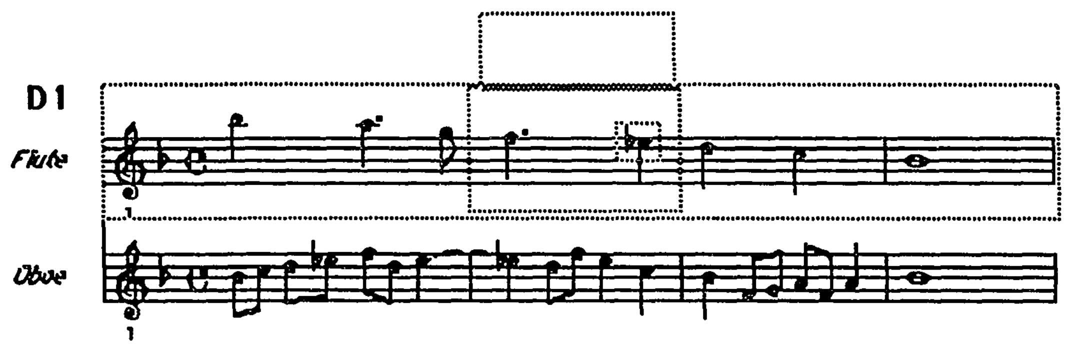
Very good ! In which measure did the error occur?



Example Set D

# Session : Example Set D

Very good! Now, which specific note was incorrect?

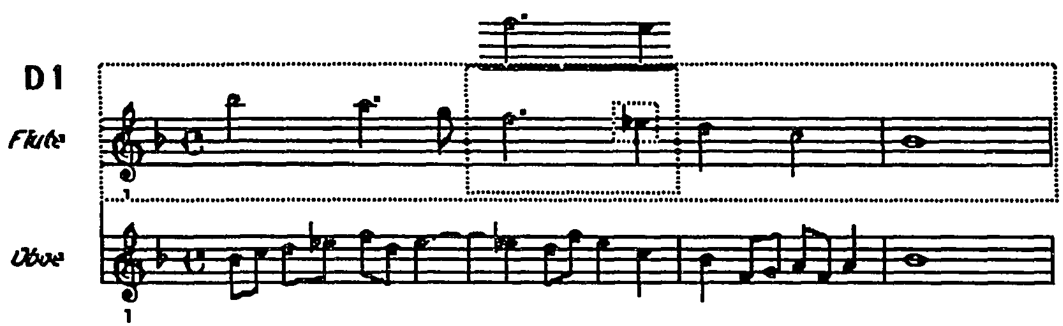


Example Set D

**Session : Example Set D**

NEXT

Right! Here is the error as it was played. Click on the NEXT button to see the next example.



## Example Set D

| | | | |
|---|---|---|---|
| Crowe Don R | Crowe | 1 | x4   3 |
| 3/3/94 | 1 | D | 9:02:14 PM |
| | | | 9:15:36 PM |

| Example | Studied | Time: Part | Time: Measure | Time: Note | Time: Total | Clicks: Part | Clicks: Measure | Clicks: Note | Clicks: Total |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 6 | 16 | 1 | 2 | 0 | 1 | 1 | 1 | 0 |
| 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 7 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 10 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 11 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 12 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 13 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 14 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 15 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 16 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 17 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 18 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 19 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 20 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 21 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 22 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

# Appendix E: Pilot Test Questionnaire

**Response Sheet: ERROR-DETECT**

**NAME:** ________________ **DATE** ___________

**SESSION #** ____ **EXAMPLE SET:** __ (A, B, C, D)

**SCORE STUDY STYLE:** _____ (Determined from the instructions given on the second card you see. Please indicate this as follows: **1** = No score study; **2** = The score only; **3** = Score and a played correct example; **4** = Score and keyboard.)

Should a specific example or section of the program "run aground," or should something seem wrong or go wrong while you are using the program, please note the example and problem on the lines below. When you have completed each session, fill out the questionnaire on the back of this page. Should you wish to make other comments, please do so on the back. All input is appreciated.

Please indicate your agreement or disagreement with the following statements. If you wish to make specific comments, please do so in the space below.

1 = Strongly Disagree
2 = Disagree
3 = Neither Agree nor Disagree
4 = Agree
5 = Strongly Agree

1 2 3 4 5 The computer was easy to use.

1 2 3 4 5 The instructions were easy to follow.

1 2 3 4 5 I had difficulty reading the examples.

1 2 3 4 5 The program ran smoothly, without "glitches."

1 2 3 4 5 It was difficult to hear the music.

1 2 3 4 5 I found the physical set up of the keyboard and computer comfortable.

1 2 3 4 5 I had difficulty using the mouse to point to errors and buttons.

# Appendix F: Program Mailing Enclosures

August 24, 1993

Dr.

Dr. :

Enclosed are new copies of the error detection program disks for my dissertation study, with instructions for their installation on the Macintosh and some suggestions for information to give the students. As you know, there were several setbacks in the last semester which prevented the collection and analysis of data. I am very concerned that everything go smoothly this time, as I must complete the study this fall. I will therefore be calling often to see that all is well. I appreciate your willingness to work with me in this study. We should get some interesting results.

The best procedure will be to have each conducting class student complete one session per week beginning in the third week of classes. This will allow for an extra week before midterms for any required makeup sessions. (**PLEASE NOTE:** It is imperative that I receive data on four sessions for each student.) This timetable will result in my receiving data by October 11: If there is to be a delay, or if any problems arise, please let me know immediately.

Please read through all of the information now and contact me if you have any questions. Again, thank you for your help in this process.

1. I have found that many students need extremely explicit, written instructions as to how to start the computer, how to get into the program, and how to shut down the computer when they are done. Once they have the "IntroStack" up and running, the program itself will give them directions, but until that point they will need some tangible help. At the end of each session, in order to retain the raw data, I have the computer automatically saving this data to another stack and then returning to the Home stack. If at this point they will choose "Shut Down" from the "SPECIAL" menu, and then cut off the machine and the keyboard, everything will be fine. I have made a few suggestions in this regard on the sheet of installation instructions.

2. Each of the colleges and universities which are participating will receive four sets of examples, but the order of these sets differs from school to school. Therefore your students might go through Example Set C in their first session and Example Set A in their last. (If you plan to use the program with more than one class, please contact me immediately: I will need to tell you how to adjust the second class's IntroStack (each class must have its own IntroStack in order to have the prescribed experimental validity.)

3. Each individual student will proceed through the four score study styles (no prestudy, score only, score plus recorded correct example, and score plus keyboard practice) in a different order. The computer automatically assigns such an order to each student in his or her first session, and then keeps track of that student's order in subsequent sessions. For this reason, it is very important that:

    A. The program be kept on only one computer (otherwise, it will not be able to track students);
    B. The students always enter their information at the IntroStack in the exact same way;
    C. No one not in the class utilizes the program until everyone in the class has had at least one session.
    D. Each class has its own IntroStack, as noted above.

    Each student must complete all four sessions (Example Sets): the most efficient way might be to require the completion of one set per week for four weeks.. I must have data on each student as he/she uses each score study style, or my main hypothesis cannot be tested. This is the focus of the study: Does one score study style better facilitate error detection than others? If I do not have data on all styles for all students, I have no study and therefore no dissertation. Please impress upon your students the importance of this aspect of the study. If you are not teaching the classes which will be using the program, please discuss this with your colleague who is. Thank you!

4. Each example will use the sound of the instrument presented in the score. In order for this to occur, the channels on your MIDI keyboard must have sampled voices assigned as follows:

| | |
|---|---|
| Channel I | Flute |
| Channel II | Oboe |
| Channel III | Clarinet |
| Channel IV | Bassoon |
| Channel V | Trumpet |
| Channel VI | Horn |
| Channel VII | Trombone |
| Channel VIII | Tuba or Baritone |

Further, the keyboard must be polyphonic in all these voices, and must be set to receive all 8 channels at the same time.

5. As the student progresses through the examples, which begin with one line of music (after the initial example of each set, which serves as a "trial run") and progress through eight lines of score, the computer keeps track of:

    A. The amount of time taken by the student to identify correctly
        i. the part,
        ii. the measure, and
        iii. the note
    on which the error occurred, and the total time spent per example;
    B. The number of times the student attempted to answer in each case; and
    C. how long the student studied the score before choosing to hear the error example (not applicable in the "no prestudy" condition);

and places all of this raw data, along with identifying information about the student and the session, on a separate card in the HyperCard stack. The sessions will take from 25 to 45 minutes, depending on the student, the score study style, and the computer which is being used (processor speed is a factor.).

6. When the student has completed <u>all 32 examples</u> of the set, the program copies the data card to another stack ("hmiditeststats") and exits to the home card. If this stack is opened, the data on the card which is seen first, and that on every card that is seen, is changed. PLEASE DO NOT OPEN THIS STACK: ONLY KEEP AN UP-TO-DATE COPY OF IT AND SEND ME A COPY OF IT AFTER ALL STUDENTS HAVE COMPLETED ALL FOUR SESSIONS. Also, emphasize to your students that they should complete each session when they start it: the computer will <u>not</u> keep track of how far they got in a certain session before abandoning

it: it will simply throw away the data, and will not realize that they need to do a session over again. There will be no record of their having done that session.

7. Data analysis will begin when all statistics stacks are received. There is no baseline data to which your students' performance can be compared at this time: the performance of all participants in the study will be used to establish means. The analysis of the data will be sent to you as soon as it is available.

8. The program is intended to be a testing tool rather than a teaching tool at this time. For this reason, the examples are presented only once. It is possible that I will work on a follow-up study which will incorporate repeated examples and more student control of the program at a later date.

If anything should go wrong, please contact me immediately. I have the password which will allow access to the program. For purposes of the integrity of the experiment, I prefer not to have it available with the package which I am sending you.

Once more, the most important items are:

1. Every class member <u>must</u> complete <u>all four</u> sessions: this means continuing to work until the program returns them to the HOME stack.

2. The program must be kept on only one computer.

3. The students must always enter their information at the IntroStack in the exact same way.

4. No one not in the class utilizes the program until everyone in the class has had at least one session. (If you wish to test the program yourself, then replace the IntroStack on your computer with the original from these disks before you allow the students to begin their sessions: this will reset the score study order. Such a test would obviously be helpful in making certain that everything is working properly: please feel free to try out the program.)

5. The channels on the keyboard must be correctly assigned, and it must be capable of playing more than one note per sampled voice or channel.

6. Do <u>not</u> open the "hmiditeststats" stack: only keep a running backup of it and send me a copy of it after all students have completed their sessions.

7. Call me <u>immediately</u> if there are any questions or problems: (417) 326-6867.

If all this seems like overkill, please forgive me: the problems which came up last spring have made me nervous. Again, thank you for all your help. Please call me at any time if there are questions or problems.

Yours,

Don R. Crowe

## ERROR-DETECT
## PROCEDURE FOR INSTALLATION

I. Read the accompanying letter.

II. *Install the MIDI Manager by dragging the three program icons MIDI Manager, Apple MIDI Driver, and PatchBay to your System Folder and restart your computer. (This step is not necessary if your computer and keyboard are already working together.) Please note that these programs are copyrighted. I have paid a licensing fee for their use with **ERROR-DETECT**†. If your computer cannot read the High-Density disks, and you cannot copy them on another machine, please call me and I will do so.

III. *Copy (do not simply move) the **Example Set** stacks, the **IntroStack**, the **hmiditeststats** stack and the **hmfinale** folder to your Hypercard folder. **Trash** or **replace** any of the stacks or folders which I sent previously.

IV. For ease of use by the students, create a new button on the HOME card, name it **"Error-Detect,"** and link it to the **IntroStack** (not to the example set stacks!) The student must always begin at the **IntroStack.**

V. For easiest access by students: if the computer will be used only for HyperCard, resize the HyperCard window so that only the HOME Stack icon is showing, and always have the students shut down from that point. In this way, the next student to turn on the computer has only to click twice on the HOME stack icon (the only one visible) and then once on the **"Error-Detect"** button in order to start the program. It will work even better if the resized Hypercard window covers the MacintoshHD icon.

VI. The **"IntroStack"** looks for the **"Example Set [A,B,C,D]"** stacks in a folder called **"HyperCard"** on a drive called **"MacintoshHD"** (your hard drive.) The **"Example Set [A,B,C,D]"** stacks look for the **MIDI files** (the aural examples themselves) in a folder called **"hmfinale"** in the same **"HyperCard"** folder. It will be easiest to simply name the folder and hard drive this way for purposes of this study: if this is not possible, you will have to go into the stacks' scripts and rewrite the affected code. Should this be necessary, call me and I will walk you through the process.

**ERROR-DETECT**
**PROCEDURE FOR INSTALLATION**
**(p. 2)**

VII. This program requires HyperCard 2.0. If you do not have this version of HyperCard, call me immediately.

VIII. I have found that many students need extremely explicit, written instructions on how to turn on the computer, how to turn it off at the end of the session, how to turn on the keyboard, how to plug in headphones, how to adjust the keyboard's volume, how not to reset the keyboard, and even how to read the directions. If you could please have an instruction sheet with this information available to them (preferably taped onto the computer so that they cannot miss it!) they will have a much smoother session.

IX. Possible problems.

1. **"Error: can't open file."** If this error message appears in a session. one or more drive/folder names is incorrect in the script. The simplest mistake is to include a space in a name which should not be there: for example, "Macintosh HD" rather than "MacintoshHD."

2. **There is no sound coming from the headphones/ speakers.** Be sure that the MIDI hardware connections are set up correctly (you really only need MIDI OUT from the computer to the keyboard.) Next check the software connections within the computer.

3. **The voices identified on the scores are not the ones you hear.** Be certain that your voices are assigned to the channels as follows:

| | |
|---|---|
| Channel I | Flute |
| Channel 2 | Oboe |
| Channel 3 | Clarinet |
| Channel 4 | Bassoon |
| Channel 5 | Trumpet |
| Channel 6 | French Horn |
| Channel 7 | Trombone |
| Channel 8 | Tuba (also used for Baritone) |

**ERROR-DETECT**
**PROCEDURE FOR INSTALLATION**
**(p. 3)**

X. *Please feel free to "test drive" the program, in all four modes (just indicate sessions 1 through 4 on the correct line of the IntroStack.) It will make me feel better if you can call me and let me know that it did work for you, under controlled circumstances. Before turning it over to the students, however, you **must** replace the **IntroStack** which you used with the original from the disks: this will ensure that the proper order of score study styles will result.

XI. Keep in touch:

Don Crowe
605 N. Wilson
Bolivar, MO 65613
417-326-6867

†ERROR-DETECT'S LICENSOR MAKES NO WARRANTIES, EXPRESS OR IMPLIED, INCLUDING WITHOUT LIMITATION THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE, REGARDING THE SOFTWARE. ERROR-DETECT'S LICENSOR DOES NOT WARRANT, GUARANTEE OR MAKE ANY REPRESENTATIONS REGARDING TYHE USE OR THE RESULTS OF THE USE OF THE SOFTWARE IN TERMS OF ITS CORRECTNESS, ACCURACY, RELIABILITY, CURRENTNESS OR OTHERWISE. THE ENTIRE RISK AS TO THE RESULTS AND PERFORMANCE OF THE SOFTWARE IS ASSUMED BY YOU. THE EXCLUSION OF IMPLIED WARRANTIES IS NOT PERMITTED BY SOME JURISDICTIONS. THE ABOVE EXCLUSION MAY NOT APPLY TO YOU. IN NO EVENT WILL ERROR-DETECT'S LICENSOR, AND THEIR DIRECTORS, OFFICERS, EMPLOYEES, OR AGENTS (COLLECTIVELY ERROR-DETECT'S LICENSOR) BE LIABLE TO YOU FOR ANY CONSEQUENTIAL, INCIDENTAL OR INDIRECT DAMAGES (INCLUDING DAMAGES FOR LOSS JOF BUSINESS PROFITS, BUSINESS INTERRUPTION, LOSS OF BUSINESS INFORMATION, AND THE LIKE) ARISING OUT OF THE USE OR INABILITY TO USE THE SOFTWARE EVEN IF ERROR-DETECT'S LICENSOR HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. BECAUSE SOME JURISDICTIONS DO NOT ALLOW THE EXCLUSION OR LIMITATION OF LIABILITY FOR CONSEQUENTIAL OR INCIDENTAL DAMAGES, THE ABOVE LIMITATIONS MAY NOT APPLY TO YOU. Error-Detect's Licensor's liability to you for actual damages from any cause whatsoever, and regardless of the form of the action (whether in contract, tort (including negligence), product liability or otherwise), will be limited to $50.

## Introductory Information for the Students

Here is the information which you will need to present to the students in class before they begin working on the program.

1. There will be written instructions on how to start the computer, get into the program, and shut down the computer. Follow them carefully and you should have no difficulty. (Please make sure these instructions are there and highly visible.)

2. You will need to complete four sessions on the computer. Do not leave in the middle of a session: the computer will not keep track of where you were, and all your work will be lost. Each session will take from 25 to 45 minutes: allow 45 minutes for the session itself and you should be safe. The session is over when the program returns you to the HOME stack, which is where you will have begun.

3. At each session, you will be asked to fill in some information on the computer. **Be sure to do this the same way each time**: the program will give you directions.

4. Each session will use a different Example Set and a different approach to score study. The order of presentation varies with the individual student. Read all instructions carefully at each session, as they will change each time you use the program.

5. (You should give them any information about your department/school's procedures for use of the computer lab, its location, whether they will have to sign in/out for you, etc.)

6. Remember, you must **COMPLETE** all four sessions by (fill in the appropriate date: I would like to have all statistics stacks back by October 10.) The first session is due _________. (One session per week for four weeks should work out very well.)

7. (If you think it wise and necessary, make the following promise/threat.)
   If any of you are interested in the workings of the computer, I will put you in touch with the program's author after we have all completed the four sessions. Do not attempt to delve into the program code itself.

Appendix G: Contacts with University I

As shown in the dates listed below, ten phone calls were made to the university from August to December 1993 in an effort to keep track of the progress of the study there. The participant was unavailable for six of these calls and never returned them. On four occassions the investigator was able to speak with the participant and was variously assured that sessions would be beginning within three weeks, that a class of eight was using the program, that the participant would determine the study's progress and call back (but did not), and that the students would complete the sessions by December 19.

From January through February 4, 1994, fifteen phone calls were made to the participant requesting shipment of the data from the fall. For twelve of these calls, the participant was unavailable and messages were left. When the investigator was able to speak with the participant, he was assured that the participant would extract the data that day and send it (January 12), that the data had been sent a day or two before (January 26), that the participant would "overnight" the data that day (February 1),.and that the data had gone out "yesterday" (February 4).

In the second week of February the investigator contacted the participant's superior to request aid in retrieving the data. The original disks which had been supplied to the university were received on February 12, with no data. Attempts were made to enlist the superior's aid in obtaining an explanation, to no avail.

Dates of attempted contact:

August 23

September 3, 7, 13, and 29

October 8 and 22

November 18

December 6 and 14

January 6, 12, 24, 25 (two calls), 26, and 31

February 1 (seven calls), 4

List of References

Aliferis, J. and Stecklein, J. E. (1962). Aliferis-Stecklein Music Achievement Test, College Midpoint Level. Minneapolis: The University of Minnesota.

Atkinson, B. (1989-1990). HyperCard. [Computer program]. Santa Clara, CA: Claris Corporation. (Version 2.0v2)

Bailey, W. (1992). Aural Skills for Conductors. Mountain View, CA: Mayfield Publishing Company

Behmer, C. F. (1984). The effects of a learning program on the ability of undergraduate nusic students to detect errors in performance. Dissertation Abstracts International, 45(10), 3088A. (University Microfilms No. 84-22018)

Blocher, L. R. (1986). An analysis of college band instrumentalists' detection of common performance errors. Dissertation Abstracts International, 35(12), 4225A. (University Microfilms No. 87-08160)

Boyer, C. G. (1974). Developing Score Reading Skills Through Programed Instruction. Dissertation Abstracts International, 35, 3789A. (University Microfilms No. 74-27,774)

Brand, M. and Burnsed, V. (1981). Music Abilities and Experiences as Predictors of Error-Detection Skill. Journal of Research in Music Education, 29(2), 91-96.

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

Byo, J. L. (1983). The Influence of Textural and Timbal Factors on the Ability of Music Majors to Detect Performance Errors. Journal of Research in Music Education, 41(2), 156-167.

Campbell, D. and Stanley, J. (1963). Experimental and Quasi-Experimental Designs for Research. Chicago: Rand McNally.

Collings, D. S. (1974). Development and Evaluation of Techniques To Improve Skill of Student Conductors in Detecting Errors of Pitch in Music Performance While Reading Five-Part and Six-Part Conductor's Scores. Dissertation Abstracts International, 34, 4820-4821A. (University Microfilms No. 73-31988)

Costanza, A. P. (1971). Programed Instruction in Score-Reading Skills. Journal of Research in Music Education, 19(4), 453-459.

Deal, J. J. (1985). Computer-Assisted Instruction in Pitch and Rhythm Error Detection. Journal of Research in Music Education, 33(3), 159-166.

Deal, J. J. (1983). Computer-Assisted Instruction To Teach Pitch and Rhythm Error-Detection Skill to College Music Education Students. Dissertation Abstracts International, 44(13), 3622. (University Microfilms No. ADG84-07733)

DeCarbo, N. J. (1982). The Effects of Conducting Experience and Programmed Materials On Error-Detection of College Conducting Students. Journal of Research in Music Education, 30(3), 187-200.

DeCarbo, N. J. (1981). The Effects of Conducting Experience and Programed Materials on Error Detection Scores of College Conducting Students. Dissertation Abstracts International, 42(4), 1528. (University Microfilms No. ADG81-21086)

Dolbeer, R. L. (1970). Programed Instruction in Detecting Aural-Notational Errors in Musical Performance. Dissertation Abstracts International, 30, 4304A. (University Microfilms No. 70-6766)

Farrand, P. (1989). Finale. [Computer program]. Bloomington, MN: Coda Music Software. (Version 2.6.1)

Gonzo, C. L. (1970). An Analysis of Factors Related to Choral Teachers' Ability to Detect Pitch Errors While Reading the Score. (ERIC Document Reproduction Service No. ED 040 917)

Grunow, R. F. (1980). Investigation of the Relative Effectiveness of Four Modes of Score Preparation on Visual-Aural Discrimination Skills Development. Dissertation Abstracts International, 41, 580A. (University Microfilms No. 80-17275

Hayslett, D. J. (1991). The effect of directed focus on the peripheral hearing of undergraduate instrumental music majors. Dissertation Abstracts International, 52(7), 2454A. (University Microfilms No. 92-00530)

Hopkins, J. E. (1991). The effect of four approaches to score study on student conductors' ability to detect errors in the performance of choral music. Dissertation Abstracts International, 52(11), 3854A. (University Microfilms No. 92-10841)

Jones, D. L. (1990). Design and trial of a computer-assisted system supplying practice in error detection for preservice instrumental music educators. Dissertation Abstracts International, 51(10), 3355A. (University Microfilms No. 91-07199)

Kepner, C. (1986). The effects 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. Dissertation Abstracts International, 47(5), 1643A.

Larson, R. C. (1977). Relationships between melodic error detection, melodic dictation, and melodic sightsinging. Journal of Research in Music Education, 25, 264-271.

Malone, P. R. (1985). Development and Evaluation of an Approach to Increasing Pitch Error Detection Skills in Choral Music Education Students. Dissertation Abstracts International, 46(12), 3644A. (University Microfilms No. ADG 86-02871)

Michels, W. J. (1973). A visual-aural self-instructional program in pitch-error detection for student choral conductors. Dissertation Abstracts International, 33(8), 4459A. (University Microfilms No. 73-2920)

Mount, T. (1982). Pitch and rhythm error identification and its relevance in the use of choral sectional rehearsals. Dissertation Abstracts International, 43(1), 13A.

Pagan, K. A. (1970). An experiment in the measurement of certain aspects of score reading aility. Dissertation Abstracts International, 31(3), 1313A.

Ramsey, D. S. (1979). Programed Instruction Using Full-Score Band Literature to Teach Pitch and Rhythm Error Detection Skill to College Music Education Students. Dissertation Abstracts International, 39, 4978-4979A. (University Microfilms No. 79-02932)

Redmon, N. (1990). HyperMIDI. [Computer program]. Torrance, CA: EarLevel Engineering. (Version 2.0)

Rizzolo, R. (1969). Experimental Tape-Recordings for Teaching Sensitivity to Musical Intonation. Final Report. East Lansing, MI Michigan State University. (ERIC Document Reproduction Service No. ED 027 854)

Shaw, T. (1971). A program to improve choral conducting students' ability to detect rhythmic errors in choral rehearsal. Dissertation Abstracts International, 32(12), 7034A.

Sidnell, R. G. (1968). The Development of Self-Instructional Drill Materials to Facilitate the Growth of Score Reading Skills of Student Conductors. Final Report. (ERIC Document Reproduction Service No. ED 023 344)

Sidnell, R. G. (1971). Self-Instructional Drill Materials for Student Conductors. Journal of Research in Music Education, 19(1), 85-91.

Simard, G. (1982). Correlation Entre le solfège, la dictèe melodique, et la détection d'erreurs. Dissertaton Abstracts International, 43(9), 2921A.

Stuart, M. H. (1978). Teacher Trainees' Development of Error Detection Skills, Two Generic Teaching Skills, and Subsequent Pupil Performance in the Musical Rehearsal. Dissertation Abstracts International, 39(2), 742A. University Microfilms No. 78-12,026)

Stuart, M. H. (1979). The Use of Videotape Recordings to Increase Teacher Trainees' Error Detection Skills. Journal of Research in Music Education, 17(1), 14-19.

Stwolinski, G. de, Faulconer, J. & Schwarzkopf, A. B. (1988). A Comparison of Two Approaches to Learning to Detect Harmonic Alterations. Journal of Research in Music Education, 36(2), 83-94.

Taebel, D. A. (1980). Public school music teachers' perceptions of the effects of certain competencies on pupil learning. Journal of Research in Music Education, 28(3), 185-197.

Tromblee, M. R. (1973). An investigation of the effectiveness of programed drill training in teaching intonation discrimination skills. Dissertation Abstracts International, 33(10), 5770-5771A.

Whittingham, M. (1990-1991). Screenshot. [Computer Program]. Memphis, TN: Baseline Publishing, Incorporated.