

INFORMATION TO USERS

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

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

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

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

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

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

**CONCEPT MAPPING AS AN ASSESSMENT TOOL:
ACCESSING LEARNING IN A CHORAL ENSEMBLE**

A Dissertation
presented to
the Faculty of the Graduate School
University of Missouri-Columbia

In Partial Fulfillment
of the Requirements for the Degree

Doctor of Philosophy

by
DEBORAH LOUISE CARR

Dr. Steven Osterlind, Dissertation Supervisor

MAY 1996

UMI Number: 9713218

UMI Microform 9713218
Copyright 1997, by UMI Company. All rights reserved.

**This microform edition is protected against unauthorized
copying under Title 17, United States Code.**

UMI
300 North Zeeb Road
Ann Arbor, MI 48103

The undersigned, appointed by the Dean of the Graduate School, have examined the dissertation entitled

CONCEPT MAPPING AS AN ASSESSMENT TOOL:
ACCESSING LEARNING IN A CHORAL ENSEMBLE

presented by Deborah Carr

a candidate for the degree of Doctor of Philosophy

and hereby certify that in their opinion it is worthy of acceptance.



Wendy L. Si

David C. Rayl

Richard Lapan

D. J. E. M. [Signature]

DEDICATION

“Now Let Me Fly”

This manuscript is dedicated to Professor Thomas L. Mills (1920 -1995), who always knew that more went on in a rehearsal than the learning of notes. Your love of music, your love of students, and your ability to mentor and support the singers that walked into your rehearsal room will never be forgotten. You allowed our souls to soar on the wings of a song and experience the joy of beauty through the arts. Thank you for all the lessons learned from the piano bench and throughout our shared lifetime. You helped my song take flight.

ACKNOWLEDGMENTS

An effort of this magnitude is never accomplished alone. The opportunity to explore new domains of knowledge is a thrilling journey and one for which I am deeply grateful. My path has crossed with the paths of numerous individuals who have taught, mentored and supported me throughout this academic experience. It is an honor to acknowledge their assistance and express my appreciation for their expertise.

The committee of a doctoral candidate is tremendously important to the learning process. I have been extremely fortunate to have an exceptional committee, made up of faculty that have always supported and encouraged my diverse interests. Dr. Steven Osterlind, my advisor, was ever the optimist, always supportive of my efforts and consistently challenging me with higher expectations. His criticism, always delivered with kindness, helped me tremendously, particularly in the area of literary expression. Dr. Richard Lapan, my Master's advisor, has supported me in the early and latter stages of my academic career and has been an excellent role model and source of advice in how to approach and complete my degree program. Dr. David McIntosh was extremely helpful in the preliminary focus of this work and provided advice and assistance in the area of statistical analysis and written expression. I am very appreciative of his technical expertise and willingness to review the methodology and results. Dr. Wendy Sims has been a brilliant instructor, role model, and supporter of my research efforts. She has provided a key link for me in bridging the disciplines of music and educational psychology. Her enthusiasm and personal interest for my work have been a source of constant support throughout my program. Dr. David Rayl has supported my musical pursuits and my research interests with unlimited amounts of patience and a never-flagging interest. His friendship and encouragement of my growth as a conductor coupled with his willingness to open his rehearsal classroom to research has allowed me to

explore many of my own interests in the area of performing arts.

Other faculty in the College of Education have contributed to my degree program as well. The faculty of the Department of Educational and Counseling Psychology have been tremendously supportive of my efforts. I need to thank Dr. Jim Koller, Dr. Karen Multon, Dr. CarolAnne Kardash and Dr. Michael Patton for their support at various times during my graduate studies. Dr. Leslie Eastman Lukin, one of my earliest mentors and supporters within the department, fostered my development as a researcher and teacher. I also need to thank Dr. John McGowen and Dr. Hope Hills, who encouraged me to pursue doctoral work in the early stages of my Master's program. Dr. John Alspaugh, Department of Leadership and Policy Analysis, has been one of my strongest mentors and supporters over the past four years. As his teaching assistant I was given a tremendous opportunity to expand my working knowledge of statistics.

Special thanks must also be given to Mike Williams for assistance in computer graphing, Beverly Rouse for her professional work in the layout and graphic display of the manuscript and to Marilyn Russell for her splendid assistance as editor and chief literary critic. I must also thank Judy Letourneau, Brenda Baker and Jennie Schultz for their friendship and support throughout my program. Each of these individuals has unique gifts and talents that have helped me through this process.

As a non-traditional student, I have had to reach out to many individuals to help and support me with family responsibilities. However, two very special people have been a source of support that has no equal. Sue Louise Goff and Rebecca Sue Keller have given hours of time to our family during the past six years. Their loving support of my children and their willingness to help with childcare allowed me the opportunity to attend class, to study and to work. Completion of my program of study would not have been possible without their help.

My immediate and extended family has always loved and supported me

throughout my life, and this experience has been no exception. I am most fortunate to come from a family of strong and courageous women who have always been a model of how to live one's life. I am especially grateful for my maternal and paternal grandmothers, Betty Evans Day and Margaret Clarke Carr, who were the epitome of independent woman. My brother, Dr. David Brian Carr, who has also supported my efforts to continue my education through his interest in my work, and through the sharing of his experiences. He, his wife Lisa and their family have supported me for these past few years in so many ways. My Columbia "mother," Jane Grant Bennett, has also provided unwavering support in my role as a mother, as a wife, as a musician, as a student and as a teacher.

To my wonderful parents, John William Carr and Patricia Margaret Carr, my heartfelt thanks for instilling me with a tremendous amount of courage, fostering my spirit and creating an inner strength that has sustained me throughout my lifetime. You have always allowed me to choose my own path in life and supported all my endeavors. How fortunate I am to have had you both in my life. My wonderful husband, Edward Sterling Rollins, has the patience of a saint and a sense of humor that has been an anchor for me. I am so grateful for your love and support. And to my beautiful children, Ryan Christopher Stoll and Emily Elizabeth Stoll, my deepest love and appreciation for who you are and for your support in your own special ways. I hope you will always love to learn, be excited about all there is know and discover and never stop asking questions.

CONCEPT MAPPING AS AN ASSESSMENT TOOL:
ACCESSING LEARNING IN A CHORAL ENSEMBLE

Deborah Carr
Dr. Steven Osterlind, Dissertation Supervisor

ABSTRACT

This study addressed the usefulness of concept maps as a supplemental assessment practice for students enrolled in a performing music ensemble course. Concept maps, visual representations of how an individual organizes knowledge, were generated on 12 pieces performed during the semester by 48 members of an auditioned choral ensemble. A list of terms and a mapping template were provided as no student had generated maps prior to this experience. This university ensemble was comprised of undergraduate non-music majors, undergraduate music majors and graduate music majors. The between-within analyses revealed significant differences for total scores among graduate and undergraduate students and among the pieces studied during the course of the semester, but revealed no significant differences for students who had, prior to the mapping experience, reviewed written commentary regarding the 12 pieces and those who had not. Elements present in the concept maps, including propositional elements, levels of knowledge, crosslinks and misconceptions, were analyzed and compared. Advantages of incorporating this type of assessment practice into a performing ensemble course are addressed.

LIST OF TABLES

Table	Page
1. Performance Pieces Used to Generate Concept Maps.....	30
2. Analysis of Variance for Map Scores by Division, Preparation and Piece	43
3. Rehearsal Time in Minutes, Percentage of Total Time and Number of Days Rehearsed for each Piece	45
4. Correlation Matrix Table - Total Scoring Procedure	46
5. Means and Standard Deviations by Piece - Total Scoring Procedure	47
6. Rehearsal Time and Means by Piece	49
7. Mean Scores for Pieces Across Division	51
8. Analysis of Variance for Map Scores by Division and Piece	52
9. Scheffé Post Hoc Analysis of Division Means for Total Scores	54
10. Contrast Analysis of Bach to Remaining Pieces in the Repertoire.....	55
11. Correlation Matrix Table - Propositional Elements.....	56
12. Analysis of Variance for Propositional Scores by Division, Preparation and Piece.....	58
13. Means and Standard Deviations by Piece - Propositional Elements	60
14. Mean Propositional Element Scores for Piece across Division.....	62
15. Analysis of Variance for Propositional Scores by Division and Piece	63
16. Scheffé Post Hoc Analysis of Division Means for Propositional Elements	65
17. Means and Standard Deviations of Total Scores on Elements by Division	66
18. Analysis of Variance for Propositional Elements	68
19. Scheffé Post Hoc Analysis of Division Means for Propositional Elements	69

20.	Tukey Post Hoc Analysis of Means for Propositional Elements	70
21.	Test of Simple Main Effects for Composer, Date Composed and Type of Composition	72
22.	Level Percentages by Piece for All Students	74
23.	Crosslink and Error Percentages by Piece for All Students.....	75

LIST OF FIGURES

Figure	Page
1. Template Map and List of Concepts Related to the 12 Compositions	33
2. Scoring Procedure for Sample Map.....	35
3. Graph of Map Means by Piece and Division - Total Score Procedure	53
4. Graph of Three-Way Interaction for Propositional Elements	59
5. Graph of Propositional Means by Piece and Division.....	64
6. Graph of Cell Means for Propositional Elements.....	71

TABLE OF CONTENTS

ACKNOWLEDGMENTS	ii
ABSTRACT	v
LIST OF TABLES	vi
LIST OF FIGURES	viii

CHAPTER

1. INTRODUCTION	1
Concept Maps	
Purpose of the Study	
Importance of the Study	
Pre-Service Training Implications	
Limitations	
Delimitations	
Assumptions	
Definition of Terms	
Overview of the Remaining Chapters	
2. REVIEW OF THE LITERATURE	13
Assessment in Music Education	
Assessment Practices in Performing Ensembles	

Concept Map Generation	
Concept Maps as a Means of Learning	
Concept Maps as a Means for Assessment	
3. METHODOLOGY	28
Repertoire	
Subjects	
Data Collection	
Scoring Procedures	
Data Analysis	
4. RESULTS	41
Data Screening Analysis	
Rehearsal Time Analysis	
Score Analysis	
Element Analysis	
5. DISCUSSION AND RECOMMENDATIONS	76
APPENDIX	83
1. Sample Maps and Scoring Procedure	84
REFERENCE LIST	87
VITA	97

CHAPTER 1

INTRODUCTION

The recent shift toward curricula based on cognitive principles has spurred intensive debate over assessment practices. Cognitive psychology is based on the notion that mental processes exist, that people are active information processors, and that mental processes and structures can be revealed through time and accuracy measures (Ashcraft, 1989). These basic assumptions have opened new avenues for researchers in the field of education. One of the central issues in education is the understanding of how information is interpreted. Equally important is how we can measure knowledge gained. Using cognitive principles has led to a “new body of instructional theory,” part of which involves assessment of students (Resnick and Klopfer, 1989). The idea that individuals draw heavily on their prior knowledge or experiences in organizing and acquiring new knowledge has led to an individualistic approach to learning, which has increased the need to personalize methods of assessment (Jonassen, 1992).

Traditionally, standard courses in testing and measurement have been offered for most teacher education programs, with pre-service students adapting such knowledge to their own academic domains. Constructivist views state that the individual constructs meaning within a “specific discourse community” (Roth, 1994). This view believes that learning takes place within a social structure, one in which the individual learns the specific language and methodology of a domain. Therefore, constructivist advocates have pushed for personalized methods of assessment, such as authentic tasks, experiential constructions (focus on process versus product), knowledge constructions, context-dependent evaluation and portfolios (Jonassen, 1992). One means of assessment that has shown to be highly sensitive in several domains, and may prove so in music education, is the concept map.

Assessment practices in the performing arts have, historically, been open to debate and scrutiny. Musical ensemble courses pose a special challenge because they have the overriding goal of performance and display of acquired skills which require cooperation by all members to be successful. For music educators, providing adequate assessment in courses of this type has always been an area of interest and concern (Warnick, 1985).

Teachers of performing ensembles have traditionally emphasized attendance, effort, and affective and attitudinal goals (Lehman, 1992b). McCoy found that high school music directors valued concert attendance and behavior, but that high school principals placed more value on cognitive criteria (McCoy, 1991). Performance is the major goal of most auditioned performing ensembles at the university level, but the hours spent in rehearsals prior to a performance may lead to an increase of knowledge in the areas of pedagogy, music history, music theory and rehearsal techniques in addition to the skills necessary to perform a piece of music. This increase of knowledge may be particularly important at the post-secondary level, where advanced ensembles are comprised of students majoring in music education.

Webster (1992) reviewed the assessment literature on creative thinking in music and noted there is "...a fresh, new trend in assessment as an educational enterprise aimed at broadening the base of what is considered for evaluation." This trend encompasses "the use of portfolios, extended student projects and interviews as means for evaluating student growth in the arts." The use of "varied mechanisms" in student assessment and evaluation will continue to be explored and researched as strides made in the area of cognitive and educational psychology offer more information about how we construct knowledge.

Concept Maps

Concept maps were developed in the 1970s to assist students in representing informative text and other materials used in school settings (Armbruster and Anderson, 1984). Since that time, the use of maps as advance organizers for both teachers and students has been explored and investigated in a variety of disciplines. Concept maps are used to represent knowledge and also incorporate language, becoming a powerful indicator of how an individual integrates concepts within a domain. The underlying principle suggests that meaningful knowledge can be represented by generating the concepts involved within the knowledge domain and then showing the connection of these concepts. The use of concept maps as a viable means of displaying an individual's "schemata" or cognitive organization of knowledge has been used in a variety of domains, including chemistry (Wilson, 1994), biology (Songer and Mintzes, 1994), physics (Roth and Roychoudhury, 1993), social studies (Torney-Purta, 1991) and medical education (Mahler, et al., 1991). Although this technique has been widely explored in a variety of disciplines, the greatest use for concept integration has been in the area of reading comprehension and the sciences.

The extension of concept maps as a measurement tool followed from the use of maps as a means of understanding text. The use of concept maps as a means of assessment was explored by Surber in 1984. He argues that assessment involves many errors, both in terms of measurement of constructs and in diagnosing areas of misunderstanding. Although he allows that an experienced test constructor may be able to generate a test that will show misconceptions (i.e., distractor options in a multiple choice test), the outcome is, in most cases, a score or percent correct. The areas where errors occur are seldom analyzed, even if distractors are well planned. Focusing only on correct responses, then, does not usually lead to a diagnosis of incorrect items. Yet, this diagnosis is necessary to instructors who seek to modify their teaching in ways that will

increase student learning.

Mapping used for testing purposes is one way of detecting errors in declarative or factual knowledge (Surber, 1984). One possible way of providing access to the learning that takes place in the performing ensemble is through the use of such maps. Other disciplines that have used concept mapping as an assessment tool include the sciences, mathematics, economics and history. Research in the area of maps as an assessment tool reveals that one of the advantages of such maps is the opportunity provided students for integration of their learning as well as the opportunity for the instructor to evaluate how information has been interpreted by students over the course of the preparatory sequence. This allows for feedback by the instructor to students, both in terms of proper connections made and the opportunity to correct errors in connections or learning that has taken place (Lambiotte, et al., 1989). Though mapping has been utilized in various capacities in a variety of domains, one area yet to be explored is the viability of this device in the performing arts.

Purpose of the Study

This study addresses assessment in the performing arts and seeks to provide a cognitive assessment technique through the use of concept maps. The general focus of this study seeks to answer the question: Are concept maps a useful addition to the assessment of students in a performing ensemble?

To address this question, all members of a performing ensemble had the opportunity to develop concept maps for each of the pieces of music studied over the course of a semester. Concept maps provide a sensitive measure of how students perceive the connections of related concepts within a domain, and it was hypothesized that maps would discriminate among participants in the ensemble (e.g. maps of students with greater background or experience should be different from those with limited background or experience) and that amount of time spent on rehearsal of the pieces

would be positively correlated with overall piece scores. In addition, the maps would provide feedback to the instructor regarding depth of knowledge for any given piece, as well as feedback regarding any errors in use of terminology or concepts as related to the repertoire studied. Specifically, the research questions addressed in this study were the following:

1. Will concept maps discriminate among groups of students within a performing ensemble?
 - a. Will student scores in a performing ensemble at the university level differ across undergraduate music majors, undergraduate non-music majors and graduate music majors?
 - b. Will preparation (i.e. studying or review of related elements) affect scores on maps?
2. What elements are present in the student-devised maps, and how are these elements related to the individual pieces and to the rehearsal time?
 - a. Are concept maps sensitive to differences among pieces in terms of overall scores?
 - b. Is there a relationship between rehearsal time for a given piece and complexity of maps generated by students?
 - c. What specific elements are evident in student-generated maps, and is there a relationship among groups or among pieces?
 - d. Are errors present in maps generated by students, and is there a relationship among error counts across individual pieces?

Importance of the Study

Traditionally, academic courses start with a set of goals and objectives and end in measurement practices designed to assess what had been learned. In musical performance, however, this model may not work. In fact, evaluation has been debated and discussed at length in the area of music education. Philosophical and practical problems involving assessment are many (Lehman, 1992b). Directors at both secondary public school and college levels utilize subjective grades, given by the instructor for such attributes as participation, preparation for rehearsal, and leadership within the organization. Citizenship qualities, ability to perform literature in small groups (i.e. subsets of the performing group), and worksheets, homework and quizzes based on musical elements are also part of the various evaluative techniques ("Point of View," 1984). Areas such as these are evaluated by the instructor, primarily by regular and systematic observation of the student.

To date, there are only a few studies that have examined the cognitive elements of knowledge acquisition and organization in the performing ensemble. The use of concept maps as a means of exploring knowledge structures is one way to understand how students in this type of course understand, construct and organize the knowledge presented to them during a preparatory sequence.

Awareness of broader application and learning opportunities also enhances the validity of offering performance classes and is a logical extension of the performance aspect. The inclusion of the arts in the public school curriculum is an accepted fact, as is the challenging task of grading and assessing performance in such courses. The same can be said for the post-secondary level. Recent philosophical debate on world view as opposed to a mechanistic view in terms of curriculum design and implementation notes that arts curricula must have a theory of learning to assist in the development of curriculum and instruction. The learner as a constructor of knowledge is part of this new philosophy (Meske, 1987). The traditional challenge of providing clearly stated goals

and objectives is evident in the literature and in curriculum programs for school districts throughout the country. Nevertheless, the validity of offering courses in the arts is accepted as part of a broad-based liberal arts education. Thus, training for teachers to instruct such courses will continue to be a major focus of pre-service training at the post-secondary level. Providing adequate means of assessing skills and knowledge in this area and recognizing current trends in curriculum design and assessment is the basis for this study.

In addition, accountability and validity of academic programs is still a pressing issue in education, as is program evaluation. "The evaluation of music programs is extremely difficult because some of the most important outcomes of music instruction cannot easily be isolated or quantified" (Lehman, 1992b). Expanding the ways of evaluating student achievement may affect overall program evaluation. Therefore, the ability to assess multiple facets of a program's value is highly desirable. The call for new and better assessment tools in the field of music education has been noted (McCoy, 1991).

Pre-Service Training Implications

The performing ensemble at the post-secondary level gives the pre-service music educator not only the opportunity to perform, but also to observe the rehearsal techniques of the conductor, learn various styles of literature, gain knowledge of composers and periods of music, and develop overall strategies for organizational management and leadership. Thus, music education majors learn valuable information in their performance classes that will translate into their own rehearsal settings when they enter the teaching field. Although part of their learning sequence, this information, or presentation of concepts, is not always visible during a performance. However, if this information is learned and retained, it will carry over in pre-service teachers' future work with performing ensembles. Therefore, monitoring acquisition of such information and

providing feedback to students enhances their training, and, in addition, provides students the opportunity to integrate their knowledge. To date, traditional methods of assessment have limited the opportunity to collect and integrate this knowledge. Therefore, better means are needed for evaluating and accessing all knowledge areas within performing ensembles at the post-secondary level.

Numerous researchers have been interested in teacher-training programs for music educators. The development of competencies and skills related to conducting, classroom techniques related to management of performing groups, and teaching of instructional techniques or pedagogy are all areas of interest. It has been noted that a collective look at studies related to the effectiveness of teacher education programs in music education indicate that “at least 80 percent mentioned a lack of practicality in teacher training” (Verrastro & Leglar, 1992). Linking theory to practice is critical to a teacher in the performing arts. The skills required for effective leadership and musicianship are modeled in a pre-service teacher’s performance courses. Capitalizing on every aspect of learning in this environment can enhance the pre-service experience.

Limitations

The scope of this study is limited to exploring the usefulness of concept maps as means of accessing one area of learning in a performing ensemble. Performance literature was chosen as the base of knowledge for the formation of maps. This choice was based on the goals and objectives of the director of the ensemble. In addition, an auditioned post-secondary choral ensemble at a Midwestern university was utilized in this study. As such, generalization of statistical results is limited to those ensembles with the same demographic characteristics as the aforementioned group.

Delimitations

Delimitations of the study include the fact that two elements, study or review of related concepts and overall personal assessment of specific knowledge gains, were gathered through the use of self-report. In addition, individual student preparation for the generation of concept maps was not a controlled variable in the study.

Assumptions

Assumptions underlying this study include the following:

1. Concept maps reflect actual links in the theoretical constructs being assessed. This is based on the cognitive theoretical approach to learning, i.e., organizing and linking information to a pre-existing knowledge base. In cognitive learning theory the classroom rehearsal is viewed as an environmental condition which facilitates learning, with the conductor's explanations and demonstrations providing environmental cues (Schunk, 1991).
2. The individual who constructs the concept map does so in a way that is consistent with the theory.
3. Conductors of performing ensembles provide information and knowledge about rehearsal techniques, performing techniques, and performance literature throughout the course of a preparatory sequence. Voyer and Faulkner (1989),

in their study of organizational cognition in jazz ensembles, refer to the rehearsal setting as a “hothouse” of enactment, both cognitive and social. From a cognitive standpoint, a rehearsal is a setting in which a variety of knowledge is dispensed, received, responded to and evaluated in a very short period of time.

Definition of Terms

Specific terms and concepts pertinent to this study and to the field of performing arts include the following:

- Repertoire:** Musical compositions that comprise the body of literature for a given musical medium (e.g. solo vocal music, choral music, orchestral music).
- Performing Ensemble:** A group of musicians that comprise an organization in which a musical repertoire is selected and rehearsed by a conductor. The most salient goal of such an organization is that of performing the selected music.
- Rehearsal:** A set period of time in which a performing ensemble practices the repertoire for an upcoming performance.
- Preparatory Sequence:** The total amount of rehearsal time an ensemble has prior to a performance.
- Concept Maps:** A graphic representation of how an individual perceives a set of concepts related to a given domain.

Propositional Elements: The basic propositions that are utilized in the generation of a concept map. These propositions will be related to the domain of interest.

Crosslink: A linking of material in a concept map that connects one hierarchical layer to another.

Error: A misconception that is evident within the concept map. This can be the use of incorrect information or the incorrect linking of concepts.

Overview of the Remaining Chapters

Chapter 2 will include the related literature that is pertinent to this study. The first part of this chapter will focus on the assessment literature in the field of music education. The review will discuss past and present assessment practices utilized in the classroom, in program evaluation, in teacher evaluation, and in musical ability. Historically behavioral in theory and focus, assessment practices found in performing ensembles will be reviewed, focusing on the assessment of performance and individual grading systems that have been utilized by instructors/directors. The second part of this chapter will focus on concept maps as a means of organizing learning, as a research tool and as an assessment measure. This will also include research in concept map generation, noting new ways of scoring and interpreting maps, and the difficulties encountered by those who generate such maps.

Chapter 3 will detail the methodology employed in the study. The research problem will be explored through quantitative analysis and will answer questions of

difference through analysis of variance techniques and questions of relationship through correlational analysis. Elements of the concept maps will be explored through factorial analysis of variance and descriptive statistics.

Chapter 4 will discuss the results in three parts. The initial part of this chapter will report statistical results on the relationship between the amount of time spent on each individual piece in the rehearsal classroom and the outcomes of the maps. The second portion will report results on the overall scores on the maps generated by the students. This analysis will report group and piece differences among mean scores on propositional elements and will also examine differences among the propositional elements of the maps. The final section will deal with additional map components present in the maps generated.

The final chapter in this study, Chapter 5, will offer conclusions and discussion of the results of this research. In addition, areas of modification will be addressed, as well as recommendations for further study.

CHAPTER 2

REVIEW OF THE LITERATURE

Research in the area of assessment practices in performing ensembles is extensive, as this has been and continues to be an area of concern in music education. The purpose of this study is to offer a new way of assessing learning outcomes in the performing ensemble through the use of concept maps. This is a new integration, combining knowledge from educational psychology with current practices within the field of music education and relating that knowledge to the performing domain. Therefore, this review will present and examine the literature related to assessment practices in music education and in performing ensembles, the relevant research conducted in the area of concept map generation to facilitate learning and the utilization of concept maps for assessment practices.

Historically, much of the assessment in the field of music has been behavioral in focus. Cognitive assessment in music education is coming to the forefront in terms of research and practice but is still limited in terms of scope. Both behavioral and cognitive practices will be presented.

Assessment Research in Music Education

Curriculum, Program and Teacher Evaluation

Part of the assessment literature in music education focuses on curriculum and teacher and program evaluation. Typically, curriculum research deals with the effectiveness of music curriculum, with evaluation of program needs in terms of number of music specialists needed and with specific tests that accompany texts for evaluation of competencies and skills (Lehman, 1992b). Lehman's summary of program evaluation in music indicates that one of the goals of such evaluation should be "not only to evaluate what students have learned, but also to evaluate the instructional program itself." In other

words, it is necessary to provide information that will assist in improving programs in schools. In addition, he notes that evaluation efforts in the field of music have been hampered by the lack of adequate measurement tools.

Studies have been done to highlight the effectiveness of the training programs that pre-service teachers experience in their undergraduate education. This line of research not only guides programs for music educators, but also for elementary classroom teachers. Evaluation of teacher training programs, viewed collectively over 30 years, points to the “lack of practicality in teacher training” (Verrastro & Leglar, 1992). In particular, two areas of described weakness were conducting and secondary methods courses. One suggestion offered by teachers in the field was the need for “more emphasis on rehearsal techniques in conducting classes.” Verrastro and Leglar also note that one area for further research involves finding ways to assist beginning teachers in the areas of lesson planning and teaching for transfer.

Teacher evaluation for music teachers is an area that is currently in need of research as the trend is toward general teacher competencies being applied across all disciplines or domains. Brophy (1993) calls for a “research-based model of effective teaching” for music education. In reviewing the related literature he notes that there is an attempt to link teaching characteristics to student achievement. This provides a problem in the field of music education as many music teachers believe that their most important results “are affective and therefore cannot be measured.” There is also “the profession’s fundamental lack of agreement on what should be taught.” Brophy proposes that a sound evaluation tool for the assessment of music teachers should include “effective planning for concept learning.”

Thus, the difficulties associated with practicality of teacher training and uncertainty about content may carry over into the classroom. Mapping may prove to be helpful in addressing these issues.

Classroom Assessment in Music Courses

The 1994 project to develop an assessment framework for arts education assessment to be used in the upcoming National Assessment of Education Progress resulted in the identification of a list of arts processes for music. This document indicates that there is a knowledge base for specific content in music and that students should be able to apply knowledge of the context (personal, social, cultural and historical) in which a piece of music was written and knowledge of the form and structure of music. The specific processes are the creating of music, the performing of music and responding to music. Responding to music includes “activities such as labeling, analyzing, classifying and placing a work within a particular context.” The third chapter of this document addresses the “desired attributes of the assessment,” which indicate that “assessment exercises should be as authentic as time and resources permit,” “tasks should be designed to elicit higher-order thinking” and “tasks should be designed to differentiate between students at both the low and high ends of the achievement spectrum.”

The result of this project has been published by the Music Educators National Conference (1996). In Performance Standards for Music (Grades Pre K-12) the following guidelines for assessment are offered:

1. Assessment should be standards-based and should reflect the music skills and knowledge that are most important for students to learn.
2. Assessment should support, enhance, and reinforce learning.
3. Assessment should be reliable.
4. Assessment should be valid.
5. Assessment should be authentic.
6. The process of assessment should be open to review by interested parties.

The Arts Education Research Agenda for the Future (National Endowments for the Arts, 1994) states that traditional methods of assessment—specifically standardized, norm-referenced, multiple-choice tests—may not provide “a complete indication of

achievement in a subject.” This document notes that standardized tests “are not integrated into the learning process and do not draw on or encourage higher order thinking skills necessary to conduct self-assessments.” Alternative methods of assessment are called for, based on the “movement from quantitative to qualitative methods of data gathering and analysis.”

Assessment in the arts is not only related to the measurement of learning outcomes, but is also a source of philosophical debate. MacGregor, et al. (1994) surveyed 527 art, drama and music instructors in Canada and noted that “almost 90% of the teachers rated assessment as desirable, very desirable, or essential.” However, music teachers indicated that the performance aspect of music programs might work “against proper assessment of students,” and that large numbers of students preclude the ability to test individually. In spite of such difficulties, teachers viewed one of the major purposes of assessment to be “an aid to personal development.”

Cognitive Assessment

Research on assessment in music education is beginning to focus on the cognitive aspects of assessment, with some recent work in the area of creative thinking in music. Qualitative analysis has supplied a variety of research techniques (e.g. case studies, protocol analysis) which have led to more sensitive means of accessing student thought. More understanding of how individual students construct knowledge has led to an increase in the need for cognitive assessment tools. Traditional instruments for measurement of skills and knowledge have shifted to methods that examine cognitive processes (Webster, 1992). Webster, in his overview of assessment in this area, notes that “focus on mental abilities” and “stages of thought during the creative act” have been part of the overall assessment focus. This trend is seen in the field of educational assessment as a whole, with a call for portfolio assessment (i.e. examples of student work), projects—both individual and group—and individual interviews.

Gray (1993), in her article on portfolio assessment, presents the assessment program utilized in the ARTS PROPEL project. This assessment method, theoretically based on Gardner's theory of intelligence, focuses on three primary elements in the assessment process—production, perception and reflection. These three areas are interwoven, and through the use of portfolios, student progress and growth can be monitored over time.

This type of assessment has been recommended for a variety of reasons. Brown (1993) argues that portfolio assessment is appropriate for middle school students because it provides a record of musical learning and “is an exhibition of what the student understands and actually knows.”

Portfolio assessment, however, is not without limitations. Wolf (1992) argues that portfolios may not withstand the demand for accountability or hold up with standardized assessments that have been traditionally used to make a variety of decisions in the schools (e.g. placement of students, receipt of services). Still, Wolf maintains that the advantage of such assessment is the ability to view how an individual thinks and to “‘uncover’ rather than to cover the curriculum.” She states that “...we rarely give students the power to make connections, to see long-running issues, or think across works.” This, she claims, is the essence of the arts.

Assessment Practices in Performing Ensembles

Performance Assessment of Ensembles

Traditionally, the assessment of a performing ensemble has been focused on the performance, i.e. the presentation of the music to the public. Ensembles' performances at contests, more structured settings in which ensembles perform for one or more adjudicators, have also been a measure of program success. Lehman (1992a) points out the limitations of such practices in that “the reliability and validity of contest rating have not been adequately investigated and documented.”

Even so, this method of assessment is a regular part of many school music programs. Garmen, et al. (1991) note that “music ensemble performance is complex and multi-dimensional, it does not lend itself readily to precise measurement.” Taking ratings from a county orchestra festival for five years, where each orchestra was rated by three judges, the authors reported interjudge reliability coefficients to range from .54 to .89 with year-to-year differences present in most categories. Higher reliability coefficients were found for overall ratings as compared to individual categories.

Mills (1991) points out the difficulties of summative performance assessment in that this approach tends to compartmentalize performance, rather than creating a more holistic approach which would indicate integration of performance elements. Again, the assessment technique relies on performance and observation of skills. She indicates that a single mark or grade will, of necessity, provide limited information. A wider range of information about achievement is called for.

Individual Assessment in Ensembles - Non-Academic

Philosophical and practical problems involving assessment in the performing ensemble are many (Lehman, 1992a). According to Lehman, many music teachers object to assessment on the grounds that they teach attitudes, values and aesthetics, none of which lend themselves easily to a form of measurement. As a result, subjective grades—given by the instructors for attendance and participation, effort, preparation for rehearsal, and leadership within the organization—have been one method of assigning grades to students. Lehman argues that these methods of assessment are limited and have led to a profusion of high grades, which, in turn, have led educators in other domains to question the validity of assessment in music classes.

Citizenship qualities, ability to perform literature in small groups (i.e. subsets of the performing group), and worksheets, homework, and quizzes based on musical elements are also part of various evaluative techniques (“Point of View,” 1984). Matheny

(1994) presented a grading system for secondary music groups which combined student self-evaluation and director evaluation and included such areas as attendance, contribution to the organization, effort, attitude, musical and technical skill, performance skill and desire to become an excellent musical performer.

Tuley (1985) proposed an assessment format that included five categories for evaluation. These included music skills (i.e. singing, playing), concepts (i.e. melody, rhythm, harmony, form), participation, and conduct. These areas are evaluated primarily by regular and systematic observation and by the student's performance.

McCoy's work in this area shows that there are differences between directors and administrators in their views of assessment. Administrators prefer more emphasis on performance techniques and cognitive criteria and less on attendance and behavior. Therefore, better models for evaluating students in ensembles are needed (McCoy, 1991).

Instructors of performing ensembles are faced with a variety of challenges. Schedules are usually demanding, concert presentations and extracurricular events numerous; performance expectations are usually high and budget money may be low. Robinson (1995) indicates that instructors of performing ensembles find themselves in a position of needing to justify individual grades, many times for high numbers of students, with little or no time available for individual evaluation. At the same time, Robinson notes that "music teachers are also being asked to redefine and justify their programs and sometimes to defend their jobs." Thus, alternate methods of "assessing both individual student achievement and overall program effectiveness" are needed. He advocates the use of criteria-specific rating scales and additive rating scales which can assist the instructor in rating students on performance skills.

Individual Assessment in Ensembles - Academic

Cognitive achievement models have been developed by Lambie (1982) and

Weymuth (1986). Lambie's model suggested a minimum competency program which identified skills and knowledge in behavioral terms with subsequent tests developed for each competency. Weymuth developed the Choral Music Achievement Test which sought to assess students in four categories involving interval identification, rhythmic precision, choral diction and choral vocabulary. This test reported adequate item difficulty and item discrimination results and differences between choral students who had studied piano privately or who participated in an instrumental group (i.e. band or orchestra) and those who did not.

With the exception of these two studies, there is little in the way of available literature relating to the cognitive assessment of individuals in performing ensembles. However, a method that has proven successful in a variety of other domains is the concept map.

Concept Map Generation

Concept mapping, or the use of diagramming informative text, was developed in the mid-1970s. Based on Ausubelian learning theory, the underlying principle suggests that meaningful knowledge can be represented by generating the concepts involved within the knowledge domain and then showing the connection of these concepts. As described by Novak and Gowin (1984), mapping allows the student to take information, organize it and place the information in a meaningful context.

In earlier studies, symbols were used to express the connections between concepts (Armbruster and Anderson, 1984). Armbruster and Anderson also explain that disciplines of study have sets of fundamental concepts that provide the basic framework for the generation of maps in those disciplines. Each concept has a set of indicators or attributes that are related to the concept. By connecting these basic concepts with the attributes, a "text frame" results.

Generation of maps may take on one of two "forms." The first is termed "bottom-

up” or “data driven” processing. This begins with lower level concepts within the domain and then builds the concepts to increasingly difficult levels. The map generation process proceeds from the simple to the complex. The second type, “top-down” or “conceptually driven,” begins with an overall generalization and the individual then generates expectations or hypotheses for the proposed topic. Prior knowledge provides the resulting layers of concepts. The choice of approach is based on the purpose of the mapper.

In their early work with mapping, Novak and Gowin (1984) noted that the propositions or concepts are linked together to form relationships, and this visual “road map” can “provide a schematic summary of what has been learned.” While indicating that linking words were not their original focus, and, in fact, can be supplied by those with knowledge of a given field, the authors note that scientific domains can benefit by including linking words which highlight how the student perceives the relationship between propositions. Their work is based on the idea of individual differences, allowing students to independently show their view of a given domain, rather than traditional assessment through the use of traditional tests.

Novak and Musonda (1991) reported the results of their twelve-year longitudinal study in which science concepts expressed by students in clinical interviews were placed in a concept mapping format by research assistants. They concluded that the development of the concept map was a useful “tool...for the design or improvement of curriculum in all subject matter fields.” The usefulness of mapping has also been shown to be effective in “skill oriented areas such as basketball and nursing.”

Lambiotte, et al. (1989) have published a review of mapping studies. They contend that application of mapping is extensive but that knowledge derivations are lacking in terms of research in this area. Findings from their studies indicate that maps appear to be helpful in the instructional and communicative processes. “Expert” maps (i.e., those generated by the instructor and used as pre-organizers or learning tools) may

be even more effective than learner-generated maps.

To address the issue of knowledge derivations and the implication of map construction as related to the processing of information and individual differences in spatial and verbal abilities, Wiegmann, et al. (1992) developed three experiments that looked at differences in map construction related to spatial configuration, map format and link structure. These differences were also examined in light of the individual's spatial and verbal prowess. The results indicated that format and linking structures are related to individual differences in spatial and verbal abilities and that the interaction of these structures affects encoding and retrieval processes.

Concept Map Generation as a Means of Learning

The use of concepts maps as a viable means of displaying an individual's "schemata" or cognitive organization of knowledge has been utilized in a variety of domains, including chemistry (Wilson, 1994), biology (Songer and Mintzes, 1994), elementary science (Fellows, 1994), social studies (Torney-Purta, 1991), theoretical orientation in teachers (Mergendoller and Sacks, 1994), teamwork (Rentsch, et al., 1994), medical education (Mahler, et al., 1991) and Alzheimer's disease (Chan, et al., 1993a).

Science education has utilized mapping as a learning tool, and research has supported the use of mapping as a facilitative activity to enhance learning and comprehension. Wallace and Mintzes (1990) proposed a study in which conceptual change in a biology class was charted through the use of concept maps. Establishing evidence of concurrent validity, these researchers suggested that mapping is a tool that complements other research methods (e.g. clinical interview, sorting tasks) in determining what students have learned over a course of instruction and is the only tool that also gives information about how students structure and organize their knowledge.

In their study of concept mapping combined with cooperative learning, Roth and Roychoudhury (1993) sought to investigate both the process and the product of this

activity by analyzing maps generated by groups of high school physics students. Their results, acquired through quantitative and qualitative analyses, indicated that when working collectively rather than individually, students spent more time discussing concepts related to material and resolving conflicts in understanding by querying the instructor for “expert advice.” However, misconceptions in understanding that were not viewed as conflicting by the group remained unchecked, and the authors concluded that instruction in determining valid hierarchies was needed, as was the opportunity for students to reflect on their maps. Although this study focused on the use of maps in a cooperative learning situation, the authors pointed out the evaluative strengths of such a procedure for the instructor, noting that during the process and through final evaluation of maps, teachers have the ability to monitor student understanding.

In a subsequent qualitative research piece based on the same sample, Roth investigated student perception regarding the construction of maps. Although most students found mapping to be a positive experience, the author did find “negative case studies” in which students did not find the mapping experience to be worthwhile. Complaints were centered in the “tediousness” of the activity, also defined as “boring.” Another individual indicated that lack of structure in terms of a starting point for hierarchical organization resulted in “mass confusion due to the endless possibilities.” In addition, a student self-described as a “bottom-up” processor of information, found the “top-down” format (i.e. general to specific concepts) extremely confusing.

Okebukola (1990), in his study of pre-degree biology students in Nigeria, found significant differences between students who had utilized concept maps in their study of genetics and ecology versus students who had not. Utilizing standard tests for each area as the dependent variable in the study, Okebukola concluded that in areas such as genetics and ecology, which are by nature complex, students who utilized concept maps had an advantage over students who did not use mapping to form connections among concepts. He noted that a common strategy for the latter group of students was to rely on

rote memorization even though research shows that this type of strategy seldom leads to critical thinking within a domain. Later research indicated that mapping is one way of taking a “conceptually opaque” domain (e.g. biology) and bringing about “conceptual transparency and learner friendliness” (Okebukola, 1993).

The incorporation of mapping into a university course in evolution (Trowbridge and Wandersee, 1994) utilized a “micromapping” technique in which students generated small maps following certain lecture presentations. This study noted that students had difficulties constructing maps, even after a presentation in map construction. Specific difficulties included the generation of linking words, the creation of linear construction as opposed to an hierarchical construction and lack of examples and cross-links. However, students did report spending more time studying by reconstructing lecture material in a mapping exercise. The authors also developed a checklist procedure to assist students in the development of their maps and an assessment procedure which allowed the instructor to score maps quickly and provide frequent feedback to students during the course of the semester.

Concept mapping has also been utilized as a research tool in studies that are designed to assess differences between experts and novices in a given domain. Songer and Mintzes (1994) utilized mapping techniques in conjunction with clinical interviews to develop an open-ended instrument to ascertain differences in conceptual understanding in introductory or novice biology students and advanced or experienced students. Mapping “provided insight into the understandings students held” in a specific scientific concept such as cellular respiration.

The use of mapping as an advance organizer has also been explored in the literature. Eighth graders were divided into two groups in a study by Willerman and Mac Harg (1991), with one group developing a map under teacher supervision at the beginning of a two-week science unit. Students who were given the mapping experience as an advance organizer did significantly better than those students who were presented

with the objectives for the unit. Curriculum development has also used mapping as a means of focusing instructors on important concepts and integrations that should be a part of classroom instruction. Middle school teachers (Starr and Krajcik, 1990) constructed and revised concept maps related to the development of a science curriculum in grades four through eight. The authors point out the constructivist basis of such an activity, with teachers involved in a cooperative effort to determine concepts and focus on integration of material, ultimately leading to a presentation of material in the classroom which leads to “meaningful conceptual knowledge structures.”

Mapping has even been used as a means of examining the differences in texts. Lloyd (1990) successfully mapped three high school biology books which were intended for different audiences and was able to make comparisons of the texts via their respective maps. Elaboration of concepts and the number of concepts provided were compared.

Concept Maps as a Means of Assessment

The use of concept maps as a means of assessment was first explored by Surber in 1984. The extension of maps as a measurement tool followed from the use of maps as a means of understanding text. Surber argues that assessment involves many errors, both in terms of measurement of constructs and in diagnosing areas of misunderstanding. Although he allows that an experienced test constructor may be able to generate a test that will show misconceptions (i.e. distractor options in a multiple choice test), the outcome is, in most cases, a score or percentage correct. This does not usually lead to a diagnosis of incorrect responses (Surber, 1984).

Research indicates that mapping is a skill that may be developed with training. Surber recommends that students be trained to generate maps prior to their use as a measurement device. His studies utilized a 23-page instruction manual and began with two to three 50-minute class sessions for training. Over the course of study, the initial

training time was extended by five 5-minute class sessions devoted to improving mapping skills.

The information evaluated in the generated map consisted of correct relationships, incorrect relationships, and omitted information. By applying this to topic content, Surber was able to generate 11 response categories. Scorers were taught the system with final interrater correlations generated at .71 and .76.

Novak and Gowin (1984) also developed a scoring system for concept maps. His system utilized propositional links, hierarchical display of information and crosslinks (connections of information within a hierarchy). This system has been used extensively in the field of science and has yielded high interrater correlations.

Markham, et al. (1994) explored the usefulness of concept maps in examining differences between post-secondary biology majors and non-biology majors in terms of structural complexity and organizational patterns of information. Their results supported the idea that concept maps are “a theoretically powerful and psychometrically sound tool for assessing conceptual change in experimental and classroom settings.”

The domain of science has the longest history and use of mapping techniques, but other disciplines have sought to utilize this method as an assessment tool. Schick (1991) utilized mapping to test elementary art teachers' understanding of library use. The findings of this study as related to the usefulness of mapping as an assessment tool was unclear. The author indicates that insufficient training time in instructing teachers how to construct maps may have hindered the effectiveness of the maps to reveal learning that had taken place. However, misconceptions about the use of library materials was clearly evident in the maps, and the author suggests that mapping may be more useful as a “qualitative assessment tool” in the field of library science.

New scoring systems designed for map analysis are developing as well. Quantitative categorization is advocated by Schreiber and Abegg (1991). This method allows for scoring in the area of propositional validity and scoring for hierarchical

structure. Wilson (1994) utilized multidimensional scaling (MDS) to allow for comparisons between individual maps and the Pathfinder network-generating algorithm to ascertain differences in the way high and low achievers construct their maps.

It is important to note that mapping, though utilized as an alternative assessment tool (particularly in the sciences), has not produced conclusive findings in the area of validity and reliability. Liu (1994) hypothesizes that this may be due to different methods of scoring. He reports that Novak's traditional scoring method is adapted freely by researchers. Liu proposes the use of item-response theory (IRT) to examine the overall structure of student maps, specifically looking at links, hierarchies and crosslinks. His findings indicate a significant correlation between Novak's total mapping scores and IRT reliability estimates.

CHAPTER 3

METHODOLOGY

The methodology of this study was developed to address the research questions proposed and was divided into three parts. The first part of the analysis explored the relationship between the amount of rehearsal time and the overall scores on the concept maps. The second part of the analysis dealt with specific hypotheses related to the scores generated from the concept maps in relationship to heterogeneity of students within the performing ensemble and differences among the pieces selected as repertoire. Finally, relationships among map elements specifically related to percentages evident across pieces and the amount of time spent rehearsing each piece during the preparatory sequence were examined. Rehearsal time for each piece was recorded during each rehearsal during the preparatory sequence, which comprised ten weeks of a sixteen week semester.

Setting and Repertoire

The instructor of this auditioned university ensemble was an award-winning junior faculty member at a Midwestern university. Repertoire was carefully chosen with vocal training of the students given high priority. A variety of choral literature coupled with instruction about the individual composers, periods of time and circumstances surrounding the composing of the pieces was addressed in class. This ensemble met daily for a 50-minute class period for a 16-week semester. The set of compositions chosen for this study were rehearsed for the first 10 weeks of the semester and performed at the end of this 10-week sequence.

The repertoire chosen by the professor for the Winter 1993 semester (see Table 1) included a set of four pieces by German composers; three classical pieces, two by Russian composers and a third by a French composer; a set of contemporary pieces, three

from American musical theatre; an arrangement of an American folk song; and a Sandburg poem set to music. These twelve pieces were the focus of the concept mapping experience. Three of the four German pieces were difficult and utilized a large portion of the overall rehearsal time. The Bach and the Schütz selections were pieces that exemplified music of the Baroque and Early Baroque periods, with an emphasis on a dance-like quality in the performance of the rhythm. The Bach was the most difficult work performed during the semester, and the conductor indicated that this particular double-choir motet was the most difficult of all the Bach motets. As such, this piece took the most rehearsal time, with the choir dividing into groups (i.e. sectionals) to achieve mastery of the parts. As stated by the conductor during the course of the semester, “This is a piece you can sing every day all semester and never get it all.”

The Distler “Lobe den Herren,” sung in German, also utilized a dance-like interpretation and was primarily rehearsed by a student conductor. The Zimmermann “Mache dich auf,” also sung in German, is a contemporary bitonal piece, based on a Lutheran chorale, which provided unique challenges to the ensemble. Zimmermann was highly influenced by the work of Bach and incorporated jazz idioms into his writing.

Also included in the semester’s repertoire was the “Sanctus and Benedictus” by Charles Gounod, featuring soprano and tenor soloists. This work was also primarily rehearsed by a student conductor. Two Russian compositions, the “Chvalite Gospoda” by Tchaikowsky, sung in Russian, and Rachmaninoff’s “Bogoróditse Dévo,” utilizing a Slavonic text, proved to be emotionally powerful for the students. Rich in tone color, with great depth and fullness in the chordal structure, these pieces were taught with verbal instructions that utilized imagery as a rehearsal strategy.

Table 1

Performance Pieces Used to Generate Concept Maps

Hodie Christus natus est	Heinrich Schütz
Fog	Phyllis Zimmerman
Sunday (Sondheim I)	Stephen Sondheim
Bogoróditse Dévo	Sergei V. Rachmaninoff
Singet dem Herrn	Johann Sebastian Bach
Our Time (Sondheim II)	Stephen Sondheim
Shenandoah	arr. James Erb
Sanctus and Benedictus	Charles Gounod
Lobe den Herren	Hugo Distler
Chvalite Gospoda	Peter I. Tchaikowsky
Mache dich auf	Heinz Werner Zimmermann
No One Is Alone (Sondheim III)	Stephen Sondheim

The conductor included an American folksong, “Shenandoah,” arranged by James Erb, and a contemporary composition by the composer Phyllis Zimmerman.

Zimmerman’s piece, “Fog,” was based on the poem by Carl Sandburg. This piece featured a soprano soloist and utilized non-traditional choral sounds to achieve the text effect.

Finally, the conductor chose three Stephen Sondheim selections to round out the program. All of these selections were from Sondheim musicals. Of the three, “Sunday,” inspired by a Seurat painting, “A Sunday Afternoon on the Island of La Grande Jatte,” was the most difficult. Field observation reports indicated that, similar to the Russian

compositions, the conductor utilized more imagery techniques with the three Sondheim pieces, which tended to draw a more emotional or affective response from the choir. The ensemble had 44 50-minute rehearsals prior to their final concert performance. They performed in a 50-minute choir exchange concert three weeks into the semester, performing the fugal section from the Bach motet, the Rachmaninoff and the arrangement of "Shenendoah." Two days prior to their final performance, the choir went on a two-day tour, performing portions of the repertoire in four separate concert appearances.

Subjects

Students enrolled in the most advanced choral ensemble (with membership determined by audition) at a Midwestern university were invited to participate in this study. This particular ensemble was comprised of 51 singers, 26 women and 25 men. Three of these students were unavailable during the final week of classes. The 48 remaining students included undergraduate non-music majors ($n=10$), undergraduate music majors ($n=22$) and graduate music majors ($n=16$). Undergraduate music majors consisted of 14 education majors, 6 performance majors and 2 majoring in theory and composition. Graduate students consisted of 2 education majors, 11 performance majors, 1 composition major and 2 that did not respond.

At the end of the semester, all students were given the option of participating in the research study or choosing the regularly scheduled written exam given in the course. The only information given to the students about the research option was that they would be doing a paper/pencil task related to the pieces studied during the semester. All students chose the research option and participated in the study. No student had generated concept maps prior to this experience.

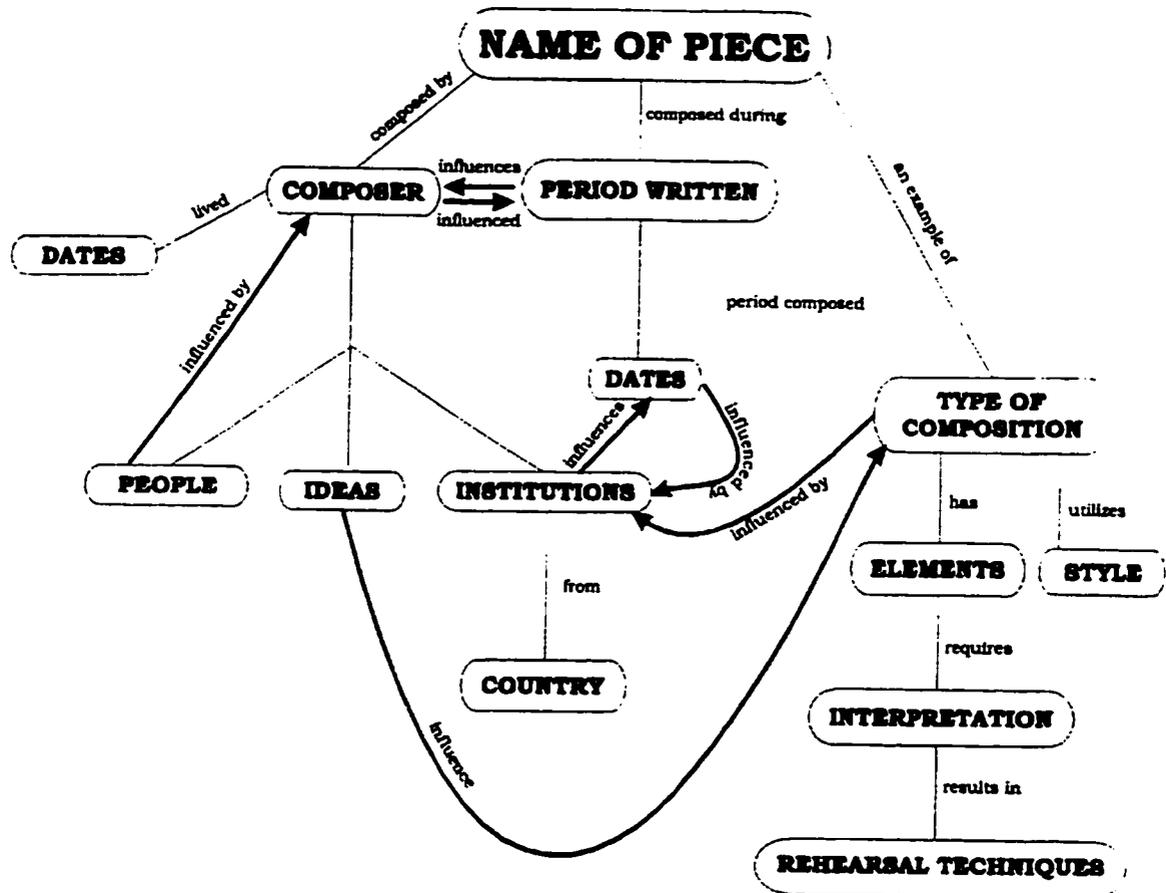
Data Collection

During the last week of the semester, students were asked to fill out basic demographic data and were given a questionnaire. Specifically, these questions asked for their most and least favorite selections of the semester and the three pieces which were the source of greatest “learning.” This was defined as an increase in knowledge of music, performance, literature and/or increase of skills as a vocalist/performer (see Figure 3). Students were also asked to review, prior to their final 50-minute class session, written program notes (contained in the concert program) prepared by the director. Students were asked to bring the demographic sheet and questionnaire to the final class session. During the final class session of the semester, students were asked to indicate by a “yes” or “no” written response whether they had read, studied or reviewed the instructor’s program notes prior to class.

Coded packets were then distributed to all students and information about concept maps was presented. Due to time constraints and the inexperience of the students with concept mapping, the decision was made to provide a list of terms, a specified set of basic propositions and a researcher-generated example of a concept map. The list of terms was generated by the researcher and the director of the ensemble for students to utilize in developing their maps, and this information and template map were placed in the student packets (see Figure 1). These terms were related information and specific concepts associated with the 12 pieces studied and performed. All terms came from the program notes and/or instruction given during the course of the semester. Students were instructed to complete the basic propositions for all maps and to add as much detail as possible. Basic propositions included the composer of the piece, the time period in which the piece was written and the type of composition. Students were allowed the remaining class time to work on maps for the 12 pieces. At the end of the class period, students were instructed to place their demographic sheet and questionnaire into the packet with their completed maps. Demographic material and questionnaires in each packet received

Figure 1

Template Map and List of Concepts Related to the 12 Compositions



German Lutheran
Cello
Folksong
Baroque
Choral Speaking
Erb
Harpsichord
Slurs
Heinz Zimmermann
Composed 1900-1950
Chorale
20th Century
Schütz
Bipartite Prelude
Praetorius
Antiphonal Singing
Tenor Solo
Oscar Hammerstein II
Composed 1950-present
Phyllis Zimmerman
Homophony

Double-Choir
1818-1893
Adjusted Vibrato
Russia
Georges Seurat
Byzantine Liturgy
Soprano Solo
Composed 1750-1800
Spirituals
1873-1943
United States
Composed 1850-1900
Merrily We Roll Along
Bach
Distler
Jazz
Rachmaninoff
Composed 1650-1700
Mussorgsky
Dance-like
Composed 1600-1650

Germany
Gounod
Bi-tonality
Monteverdi
Liturgy of St. John Chrysostom
Tchaikowsky
Gabrieli
1840-1893
Sondheim
Fugue
Eastern Orthodox
Unaccented Syncopation
String Bass
Motet
Eighth-note Pulse
Early Baroque
Sandburg
Musical Theatre
Space before pairs of 16th notes
Episode

Composed 1700-1750
Text stress-strong/weak
Renaissance
France
Appoggiatura
Sunday in the Park With George
Messe solennelle de Ste. Cecil
Ave Maria
Brothers Grimm
Polyphony
Communion Hymn
Babbitt
Short-long
Church Slavonic
Romantic
Legato
The All-Night Vigil
Detached
Dynamic Contrast
Into the Woods

the same code as packet materials. Coding of maps and demographic material was done by the researcher to insure anonymity.

Scoring Procedures

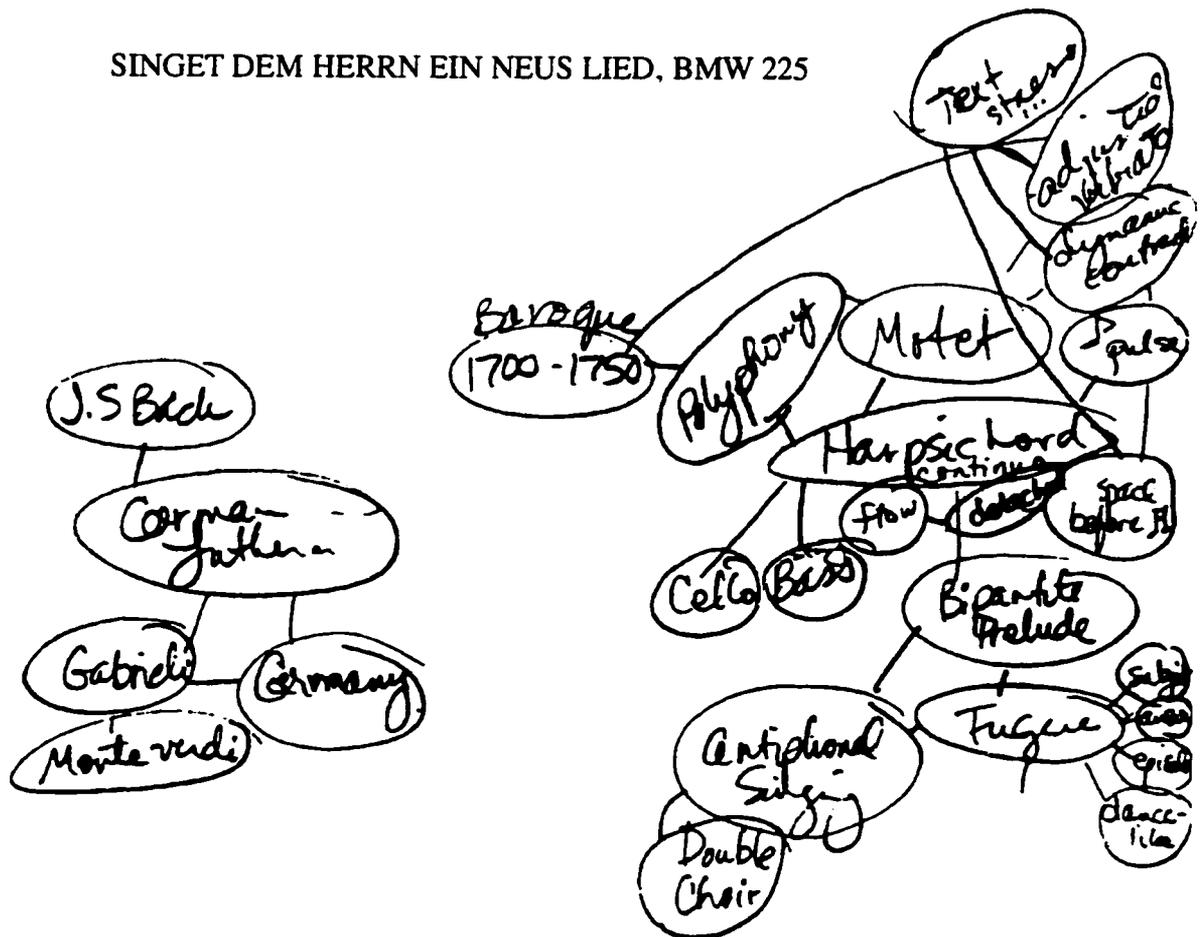
A scoring system, adapted from Novak and Gowin (1984), was developed for the purposes of this study. In this system, the instructor was primarily interested in three items of information for each piece—the composer, the time period in which the piece was written and the type of composition. These elements, referred to as basic propositions, were awarded two points each, for a total of six points. Maps which contained extended knowledge related to the first set of basic propositions were given five additional points. This extension of knowledge was referred to as “Level 1 concepts,” and earned students an additional 5 points for each basic proposition embellished. Further embellishment of Level 1 concepts was considered “Level 2,” and, again, embellishment of Level 1 concepts in Level 2 earned students five points for each basic proposition enhanced. If students continued to display connections correctly related to the second level of concepts, the map was classified as “Level 3,” and the same scoring principles were used. Crosslinks, or connections of material among the three initial elements, indicated a synthesis of information and a link between two basic propositions. As such, these were awarded 10 points. Errors were indicated by an “E,” one “E” for each error present in the map. In terms of points earned, students were not penalized by the presence of errors.

The conductor of the ensemble generated a master set of maps that was utilized in the scoring of the student-generated maps. Points were totaled for each piece; Level 1, 2 and 3 counts were taken, as well as error and crosslink counts. Each subject’s total score,

Figure 2

Scoring Example of a Student Generated Concept Map

SINGET DEM HERRN EIN NEUS LIED, BMW 225



<u>Scoring</u>	
Propositional Elements:	6
Level One:	10
Level Two:	5
Crosslink:	10
Total:	31

Sample Map: Graduate Music Major

number of levels reached and error and crosslink counts were entered into the data set. Both the conductor and the researcher scored the maps to establish interrater reliability. A sample map, indicating scoring procedures, is given in Figure 2. Additional examples of scoring are given in Appendix A.

Data Analysis

A Priori Planned Comparisons

Of particular interest was the comparison of the Bach “Singet dem Herrn” with the remaining repertoire. This piece was designated a priori by the director of the ensemble as the most challenging piece of literature studied over the semester. To determine if this piece did significantly differ from the other 11 pieces studied during the semester, an orthogonal comparison was run, comparing the Bach to the remaining literature selections. This resulted in a more powerful comparison of the data, rather than the traditional post hoc analysis utilizing Tukey’s pair-wise comparison procedures.

Data Screening Procedures

Screening of the data set to check for normality was done through residual analysis, with adjustments made where necessary. Traditionally, between-within analyses were based on the assumption of compound symmetry, requiring homogeneity of variance and homogeneity of covariance. However, more recent research involving Monte Carlo techniques has indicated that this is unnecessarily stringent. A modification of the compound symmetry, sphericity, is considered to be sufficient in satisfying this assumption. Since the sample sizes for the between factor were unequal, this was screened (Stevens, 1992). In addition, screening for outliers, again important due to sample size, was conducted, utilizing the procedures suggested by Tabachnick and Fidell (1989). Missing data points were analyzed for patterns, and decisions regarding estimation or case deletion were made following analysis.

Relationship of Rehearsal Time to Concept Map Scores

The degree of relationship between the amount of rehearsal time and overall mean scores for each piece was calculated. Correlational analysis was utilized to establish the degree of relationship between amount of rehearsal time and overall scores present in the maps. The hypothesis for this analysis was

Ho₁ - There is no statistically significant correlation between the amount of rehearsal time spent on each piece and the overall mean scores for the 12 pieces ($\rho = 0$, $\alpha = .05$).

Primary Analysis

The primary analysis for the overall scores on the maps, a three factor, 3 X 2 X 12, two between and one within design, was utilized to partition the variance components of the overall scores. The between subjects variables were the division represented by the students in the performing ensemble and preparation of the material. The first independent variable was comprised of the three levels previously introduced. The second independent variable, that of preparation, involved two levels, students who studied the program notes and those who did not. The within-subjects factor was the 12 pieces of music studied over the course of the semester and ultimately performed (see Table 1); and the dependent variable was the overall scores of the concept maps generated by the students in the ensemble. This analysis was run through a MANOVA procedure as a between-within design. The reader is referred to Winer (1971) for a breakdown of the variance components. The hypotheses for the three main effects and interactions were

MAIN EFFECTS

- Ho₁ - There is no statistically significant difference among mean scores for the three divisions of students ($\alpha = .05$).
- Ho₂ - There is no statistically significant difference between mean scores for students who studied the program notes and those who did not ($\alpha = .05$).
- Ho₃ - There is no statistically significant difference among mean scores for the 12 pieces of music ($\alpha = .05$).

INTERACTIONS

- Ho₄ - There is no statistically significant interaction between division and preparation (those who studied versus those who did not) in their influence on mean scores ($\alpha = .05$).
- Ho₅ - There is no statistically significant interaction between division and piece in their influence on mean scores ($\alpha = .05$).
- Ho₆ - There is no statistically significant interaction between preparation (those who studied versus those who did not) in their influence on mean scores ($\alpha = .05$).
- Ho₇ - There is no significant interaction among division, preparation and piece in their influence on mean scores ($\alpha = .05$).

Analysis of Concept Map Elements

Propositional Elements - Propositional elements were scored for all pieces and analyzed in two ways. Total scores on propositional elements were analyzed by a three factor, 3 X 2 X 12, two between and one within design. The between subjects variables were the division represented by the students in the performing ensemble and preparation

of the material. The first independent variable was comprised of three levels—graduate music majors, undergraduate music majors, and undergraduate non-music majors. The second independent variable, that of preparation, involved two levels—students who studied the program notes and those who did not. The within-subjects factor was the 12 pieces of music studied over the course of the semester and ultimately performed (see Table 1), and the dependent variable was the score received for the propositional elements. The hypotheses for the three main effects and interactions were

MAIN EFFECTS

Ho₁ - There is no statistically significant difference among mean propositional scores for the three divisions of students ($\alpha = .05$).

Ho₂ - There is no statistically significant difference between mean propositional scores for students who studied the program notes and those who did not ($\alpha = .05$).

Ho₃ - There is no statistically significant difference among mean propositional scores for the 12 pieces of music ($\alpha = .05$).

INTERACTIONS

Ho₄ - There is no statistically significant interaction between division and preparation (those who studied versus those who did not) in their influence on mean propositional scores ($\alpha = .05$).

Ho₅ - There is no statistically significant interaction between division and piece in their influence on mean propositional scores ($\alpha = .05$).

Ho₆ - There is no statistically significant interaction between preparation (those who studied versus those who did not) and piece in their influence on mean propositional scores ($\alpha = .05$).

Ho₇ - There is no significant interaction among division, preparation and piece in their influence on mean propositional scores ($\alpha = .05$).

Additional Elements - Levels, error and crosslink counts were calculated for each student and reported descriptively for each piece, by percentages, with observations of map characteristics noted and discussed.

In addition to the above analyses, a comparison of propositional elements was done. It was hypothesized that students would be able to indicate the composer of a piece easily but would experience greater difficulty identifying the time the piece was composed and the type of composition. Scores for each of the three propositional elements were totaled across all pieces for each student and compared by division. This resulted in a two factor repeated measure analysis. The between factor was the division by the student, comprised of the previously stated levels, and the within factor was the total scores for composer, date written and type of composition. The dependent variable was the total scores for each propositional element. Hypotheses for this analysis were

Ho₁ - There is no statistically significant difference among mean scores for the three divisions of students ($\alpha = .05$).

Ho₂ - There is no statistically significant difference among mean scores for the three propositional elements ($\alpha = .05$).

Ho₃ - There is no statistically significant interaction between division and propositional elements in their influence on mean scores ($\alpha = .05$).

CHAPTER 4

RESULTS

Data Screening Analysis

Of the 48 sets of maps (one map for each of the 12 pieces studied) generated by students, seven sets were incomplete. Specifically, two sets were missing a map for only one piece, two sets were missing two maps, and three additional sets were missing five, nine and ten maps, respectively. The decision was made to delete those cases missing more than two maps, as there was not enough information to estimate viable values for those cases. Of the remaining four sets, the mean score of the individual was utilized in place of the missing value. This procedure is supported by Tabachnik and Fidell (1989) as a conservative method of dealing with missing values. As the missing data points were randomly placed throughout the scores, the possibility of loss in variability was minimal. Early screening tests showed no difference in the major of music students (e.g. performance versus education majors) so all music students were collapsed into a single category and divided by division, graduate and undergraduate. The tests for homogeneity of variance, utilizing the Bartlett-Box procedures as outlined in Tabachnick and Fidell (1989) were satisfactory.

Comparison of both the director's and researcher's scores yielded an interrater reliability coefficient of .91. Using the procedures outlined in Crocker and Algina (1986), this value is in accordance with the reliability ratings reported by Novak and Musonda (1991).

Complete scoring of the maps generated several outliers which exceeded the range of scores depicted in box plot analyses. One of these cases proved to be a scoring error and was corrected. The remaining outliers were present in all three groups, one in the undergraduate non-music majors, two in the undergraduate music majors, and one in

the graduate music majors. As these scores were plausible in terms of range of scores, the decision was made to keep these cases in the data set. Coupled with the aforementioned deletion of three cases, this resulted in group sample sizes of 10, 21 and 14 for a total sample size of 45. However, three students did not respond to the question regarding preparation. For the main analysis, the hypotheses for the three main effects and interactions were

MAIN EFFECTS

Ho₁ - There is no statistically significant difference among mean scores for the three divisions of students ($\alpha = .05$).

Ho₂ - There is no statistically significant difference between mean scores for students who studied the program notes and those who did not ($\alpha = .05$).

Ho₃ - There is no statistically significant difference among mean scores for the 12 pieces of music ($\alpha = .05$).

INTERACTIONS

Ho₄ - There is no statistically significant interaction between division and preparation (those who studied versus those who did not) in their influence on mean scores ($\alpha = .05$).

Ho₅ - There is no statistically significant interaction between division and pieces in their influence on mean scores ($\alpha = .05$).

Ho₆ - There is no statistically significant interaction between preparation (those who studied versus those who did not) and piece in their influence on mean scores ($\alpha = .05$).

Ho₇ - There is no significant interaction among division, preparation and piece in their influence on mean scores ($\alpha = .05$).

Preliminary analysis (see Table 2) revealed no significant differences between the self-reported variable for those who prepared versus those who did not ($n=42$).

Table 2

Analysis of Variance for Map Scores by Division, Preparation and Piece

Source	SS	df	MS	F
Between subjects				
Division	454.17	2	227.08	3.11
Preparation	36.85	1	36.85	.50
Division X Preparation	29.88	2	14.94	.20
Within	2629.30	36	73.04	
Within subjects				
Piece	2913.22	11	264.84	12.62*
Division X Piece	342.05	22	15.55	.74
Preparation X Piece	122.87	11	14.15	.53
Division X Prep X Piece	393.77	22	17.90	.69
Within Cells	8309.31	396	20.98	

* $p < .0001$

Division differences in this analysis for $F(2, 36) = 3.11$, $p < .057$ were also not significant at the .05 level of significance. The within subject effect of piece, however, was highly significant ($F(11, 396) = 12.62$, $p < .0001$).

Since three students did not respond to the question regarding preparation, this variable was dropped from the final analysis, to maximize the available data. Thus, a sample size of N=45 was utilized for all subsequent procedures.

Rehearsal Time for Pieces

The selections chosen for the semester varied in length and difficulty. Therefore, the amount of time spent on each piece varied tremendously. Table 3 presents the amount of rehearsal time spent, in minutes, during the 44 instructional periods for this performing sequence. Time was calculated for all minutes spent on the pieces, including verbal instructions and information related to the pieces, in addition to actual singing time. Given 50 minutes for each instructional period, the entire amount of rehearsal time available to the director was 2,200 minutes. Warm-up exercises, which included vocalizing and specific exercises related to individual pieces, comprised 150 minutes of the total rehearsal time. Time related directly to the rehearsal of the pieces totaled 1,680 minutes. The remaining 370 minutes were related to information necessary to the functioning of the organization (e.g., distribution of syllabus, instructions for the choir exchange, tour and final performance), and a small amount of time was used rehearsing an additional work that was performed in conjunction with another choral organization in the music department.

The results clearly indicate that the Bach motet received the largest portion of total rehearsal time. In fact, the four German compositions comprised nearly 72 percent of the total rehearsal time. The remaining pieces ranged from 26 to 88 minutes of time, encompassed between 1.55 and 5.24 percent of the rehearsal time and were only rehearsed for four to nine rehearsal days, depending on the piece. Not only did the Bach piece receive the largest amount of rehearsal time, but this piece was also rehearsed 34 of

the 44 days in the performing sequence. The other German composers—Schütz, Distler and Zimmermann—were rehearsed 18, 9 and 12 days respectively.

Table 3

Rehearsal Time in Minutes, Percentage of Total Time and Number of Days Rehearsed for each Piece

Piece	Title	Composer	Minutes	Percentage	Days
1	Hodie Christus natus est	Schütz	238	14.12	18
2	Fog	Zimmerman	26	1.55	4
3	Sunday (S1)	Sondheim	45	2.68	4
4	Bogoróditse Dévo	Rachmaninoff	88	5.24	7
5	Singet dem Herrn	Bach	718	42.74	34
6	Our Time (S2)	Sondheim	62	3.69	7
7	Shenandoah	arr. Erb	61	3.63	6
8	Sanctus and Benedictus	Gounod	86	5.12	8
9	Lobe den Herren	Distler	104	6.19	9
10	Chvalite Gospoda	Tschaikowsky	65	3.87	9
11	Mache dich auf	Zimmermann	150	8.93	12
12	No One Is Alone (S3)	Sondheim	37	2.20	4

Division and Piece Analysis - Total Scoring Procedure

Scores for all pieces were totaled and entered into the data set. Intercorrelations for the 12 pieces are presented in Table 4.

Table 4

Intercorrelation Between Pieces - Total Scoring Procedure (n=45)

Piece 1	2	3	4	5	6	7	8	9	10	11	12	
1	—	.24	.16	.02	.27	.15	.16	.26	.31*	.29	.21	.18
2		—	.40*	.28	.29	.49**	.37	.44**	.18	.23	.39**	.29
3			—	.13	.20	.56**	-.01	.32*	-.04	-.10	.23	.13
4				—	.00	.18	.34*	.31*	.18	.06	.21	.02
5					—	.00	-.08	.31*	-.08	.49**	-.04	.17
6						—	.37*	.40**	-.01	-.03	.35*	.11
7							—	.20	.07	.15	.16	.18
8								—	.31*	.29*	.54**	.35*
9									—	.31*	.44**	-.13
10										—	.21	.42**
11											—	.30**
12												—

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

The results indicate several significant relationships between pieces, all of them positive. This could indicate a commonality between elements in some pieces in that the ability to complete one map may have influenced success on another map. Correlations may be restricted due to the small sample size and restriction in range.

To answer the hypotheses of no differences among group members or pieces studied, a 2-factor between-within analysis was conducted. Division and piece were the independent variables, with piece being a within-factor for the 12 pieces studied. Total score attained on each piece was the dependent variable. Piece means and standard deviations are presented in Table 5. This table indicates that the Bach selection had the highest mean score of 17.82, with the lowest score represented by the Distler “Lobe den Herren” at 7.29.

Table 5
Means and Standard Deviations by Piece (n=45)

Piece	Title	Mean	Standard Deviation
1	Hodie Christus natus est	9.36	6.01
2	Fog	7.96	3.74
3	Sunday (S1)	14.44	5.72
4	Bogoróditse Dévo	12.70	5.57
5	Singet dem Herrn	17.82	6.59
6	Our Time (S2)	13.10	4.33
7	Shenandoah	8.58	4.44
8	Sanctus and Benedictus	8.19	4.49
9	Lobe den Herren	7.29	3.36
10	Chvalite Gospoda	10.22	6.33
11	Mache dich auf	10.13	5.44
12	No One Is Alone (S3)	11.89	3.33

In addition, the three selections that received the next highest mean scores were the two Sondheim pieces—"Sunday" at 14.44 and "Our Time" at 13.10—and the Rachmaninoff at 12.70. The next highest mean score was for the third Sondheim selection, "No One Is Alone," at 11.89.

To test the relationship between the number of minutes that each piece was rehearsed during the performing sequence and the overall mean score for each piece, a correlational analysis was conducted. The number of minutes and mean score for each piece is presented in Table 6. The hypothesis for this analysis was

H_{o_1} - There is no statistically significant correlation between the amount of rehearsal time spent on each piece and the overall mean scores for the 12 pieces ($\rho = 0$, $\alpha = .05$).

Table 6

Minutes Spent in Rehearsal and Means by Piece

Piece	Title	Minutes	Mean
1	Hodie Christus natus est	238	9.36
2	Fog	26	7.96
3	Sunday (S1)	45	14.44
4	Bogoróditse Dévo	88	12.70
5	Singet dem Herrn	718	17.82
6	Our Time (S2)	62	13.10
7	Shenandoah	61	8.58
8	Sanctus and Benedictus	86	8.19
9	Lobe den Herren	104	7.29
10	Chvalite Gospoda	65	10.22
11	Mache dich auf	150	10.13
12	No One Is Alone (S3)	37	11.89

The correlation between minutes rehearsed and mean score was .62. This was statistically significant at the .01 level of significance. The coefficient of determination was .40, which indicates that mapping scores and the amount of time spent in rehearsal have 40% shared variance.

Group means and standard deviations by piece are presented in Table 6. Undergraduate non-music majors had the highest mean score for the Bach (14.40) and the lowest scores on the Schütz “Hodie” (6.40) and Zimmermann “Mache dich auf” (6.40). Undergraduate music majors scored highest on the Bach (17.76) and lowest on the Distler

(6.19). Highest overall, the graduate music majors scored highest on the Bach (20.36) and lowest on the arrangement of “Shenendoah” (9.14) and the Distler (9.14). Overall, scores were higher for all pieces for the graduate music majors than for the two undergraduate divisions. Highest scores across all divisions were evident in the Bach.

For the main analysis, the hypotheses for the two main effects and interaction were

MAIN EFFECTS

Ho₁ - There is no statistically significant difference among mean scores for the three divisions of students ($\alpha = .05$).

Ho₂ - There is no statistically significant difference among mean scores for the 12 pieces of music ($\alpha = .05$).

INTERACTION

Ho₃ - There is no statistically significant interaction between division and piece in their influence on mean scores ($\alpha = .05$).

Table 7

Mean Scores for Pieces across Division

Piece	Title	Division			
		Undergraduate Non-Majors (n=10)	Undergraduate Music Majors (n=21)	Graduate Music Majors (n=14)	
1.	Hodie Christus natus est	<u>M</u>	6.40	8.57	12.64
		<u>SD</u>	2.95	4.73	7.90
2.	Fog	<u>M</u>	6.80	7.43	9.57
		<u>SD</u>	2.70	3.60	4.26
3.	Sunday (S1)	<u>M</u>	12.60	14.29	16.00
		<u>SD</u>	6.42	5.48	5.55
4.	Bogoróditse Dévo	<u>M</u>	13.65	12.57	12.21
		<u>SD</u>	8.27	4.11	5.55
5.	Singet dem Herrn	<u>M</u>	14.40	17.76	20.36
		<u>SD</u>	4.65	6.31	7.42
6.	Our Time (S2)	<u>M</u>	11.66	13.38	13.71
		<u>SD</u>	3.90	4.36	4.65
7.	Shenandoah	<u>M</u>	8.70	8.14	9.14
		<u>SD</u>	4.24	4.11	5.25
8.	Sanctus and Benedictus	<u>M</u>	7.20	6.97	10.71
		<u>SD</u>	4.94	4.06	3.99
9.	Lobe Den Herren	<u>M</u>	7.00	6.19	9.14
		<u>SD</u>	2.79	3.17	3.42
10.	Chvalite Gospoda	<u>M</u>	8.71	9.57	12.29
		<u>SD</u>	2.81	7.59	5.85
11.	Mache dich auf, werde licht	<u>M</u>	6.40	10.43	12.36
		<u>SD</u>	3.57	5.34	5.58
12.	No One Is Alone (S3)	<u>M</u>	11.30	11.67	12.64
		<u>SD</u>	3.43	3.37	3.30

The between-within analysis of variance is given in Table 8. The between-subjects effect for group was significant $F(2,42) = 5.16, p < .01$, as was the within effect for Piece $F(11, 462) = 18.29, p < .0001$.

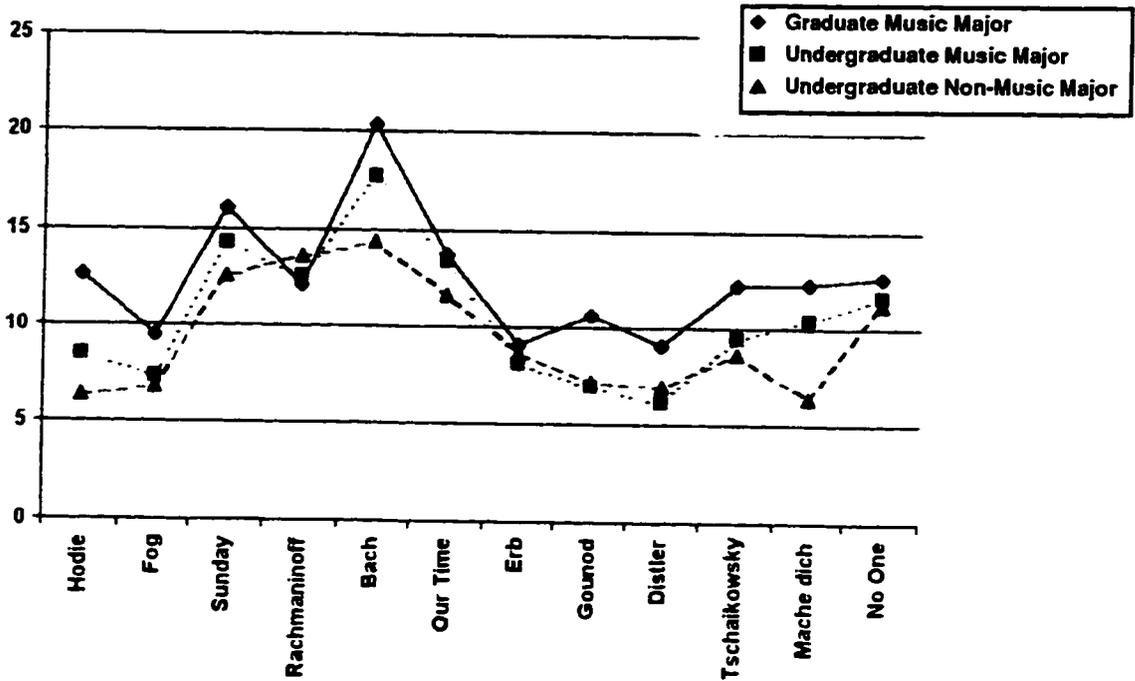
Table 8
Analysis of Variance for Concept Maps

Source	SS	df	MS	F
Between subjects				
Group	701.58	2	350.79	5.16*
Within Cells	2854.71	42	67.97	
Within subjects				
Piece	4155.20	11	377.75	18.29**
Group by Piece	449.10	22	20.41	.99
Within Cells	9541.96	462	20.65	

* $p < .01$, ** $p < .0001$

This indicates that a difference among the groups was present, as well as a highly significant difference among the overall piece means. The graph of the total mean scores by piece and by division is presented in Figure 3. The interaction of group and piece in their influence on mean scores was not significant.

Figure 3
Graph of Map Means by Piece and Division - Total Score Procedure



Scheffé post hoc analysis was applied to the group means and is presented in Table 9. This post hoc analysis did not reveal any significant differences despite an overall omnibus F with $p < .01$. This was most likely due to the small range in values for the three groups.

Table 9

Scheffé Post Hoc Analysis of Group Means

	Undergraduate Non-Major (<u>n</u> =10)	Undergraduate Music Major (<u>n</u> =21)	Graduate Music Major (<u>n</u> =14)
Undergraduate Non-Major		.38	9.56
Undergraduate Music Major			6.15
Graduate Music Major			

* $p < .05$

** $p < .01$

To compensate for the small range in overall mean values, a one-factor ANOVA was run utilizing the sum of total scores for each division. This analysis, yielding the same result as the between factor F test, was significant, $F(2,42) = 5.16, p < .01$. The post hoc analysis revealed a significant difference between the undergraduate non-music majors and the graduate music majors, but no significant differences were found between the two undergraduate groups nor between the undergraduate music majors and the graduate music majors.

Contrast Analysis

One of the specific questions of interest was the relationship of the Bach “Singet dem Herren” to the remainder of the pieces. To reiterate, this particular piece was the longest, most difficult piece of literature that the choir worked on during the preparatory sequence. To answer the question of differences between the Bach and the remaining

eleven pieces, a contrast analysis was done comparing the Bach piece mean with the remaining eleven pieces. The results of this analysis are presented in Table 10.

Table 10

Contrast Analysis of Bach to Remaining Pieces in the Repertoire

	<u>P1</u>	<u>P2</u>	<u>P3</u>	<u>P4</u>	<u>P5</u>	<u>P6</u>	<u>P7</u>	<u>P8</u>	<u>P9</u>	<u>P10</u>	<u>P11</u>	<u>P12</u>
Score	421	358	650	571.5	804	589.6	386	368.36	333	460.1	456	535
Lambda	-1	-1	-1	-1	+11	-1	-1	-1	-1	-1	-1	-1

Critical Sums of Squares = 39.08

Contrast Sums of Squares = 2535.26*

* $p < .05$

The results reveal that of the original 4533.58 attributed to the sums of squares due to variation by piece, 2535.26 can be attributed to the contrast of the Bach compared to the remaining 11 pieces. To be significant at the .05 level of significance, the contrast value must exceed the critical sums of squares (39.08). This indicates that the Bach was significantly different from the combined effect of the additional 11 pieces in the repertoire.

Analysis of Concept Map Elements

Propositional Analysis

The propositional analysis studied the three propositional elements—composer, date piece was composed and type of composition—and made comparisons of these elements by division, by piece and by element. These three pieces of information were individually worth two points, which could result in a maximum score of six points for each piece. Intercorrelations for the 12 pieces are presented in Table 11.

Table 11

Intercorrelation Between Pieces

Piece	1	2	3	4	5	6	7	8	9	10	11	12
1	—	-.12	.25	.06	.09	.38*	.24	.44**	.52**	.40**	.48**	.26
2		—	.15	.25	.26	.41**	.11	.35*	.22	.13	.28	.21
3			—	.08	.27	.62**	-.28	.30*	.10	.31*	.29	.31*
4				—	.25	.15	.20	.37*	.26	.59**	.53**	.34*
5					—	.44**	.06	.37*	-.03	.42**	.26	.29
6						—	-.14	.48**	.19	.29	.45**	.47**
7							—	.34*	.44**	.33*	.18	.19
8								—	.69**	.62**	.59**	.50**
9									—	.41**	.55**	.36*
10										—	.59**	.43**
11											—	.56**
12												—

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

Students were instructed to complete all propositional elements for each piece, and then to continue to expand their maps as time permitted. As such, the overall scoring results must be interpreted cautiously, as not all students were able to embellish all pieces

during their final exam period. To examine the question of possible group, preparation and piece differences among propositional elements, a second repeated measure analysis was conducted. The hypotheses for the three main effects and interactions were

MAIN EFFECTS

Ho₁ - There is no statistically significant difference among mean propositional scores for the three divisions of students ($\alpha = .05$).

Ho₂ - There is no statistically significant difference among mean propositional scores for students who studied the program notes and those who did not ($\alpha = .05$).

Ho₃ - There is no statistically significant difference among mean propositional scores for the 12 pieces of music ($\alpha = .05$).

INTERACTIONS

Ho₄ - There is no statistically significant interaction between division and preparation (those who studied versus those who did not) in their influence on mean propositional scores ($\alpha = .05$).

Ho₅ - There is no statistically significant interaction between division and piece in their influence on mean propositional scores ($\alpha = .05$).

Ho₆ - There is no statistically significant interaction between division and piece in their influence on mean propositional scores ($\alpha = .05$).

Ho₇ - There is no statistically significant interaction among division, preparation and piece in their influence on mean propositional scores ($\alpha = .05$).

The results of this analysis are presented in Table 12, utilizing total points for propositions as the dependent variable.

Table 12

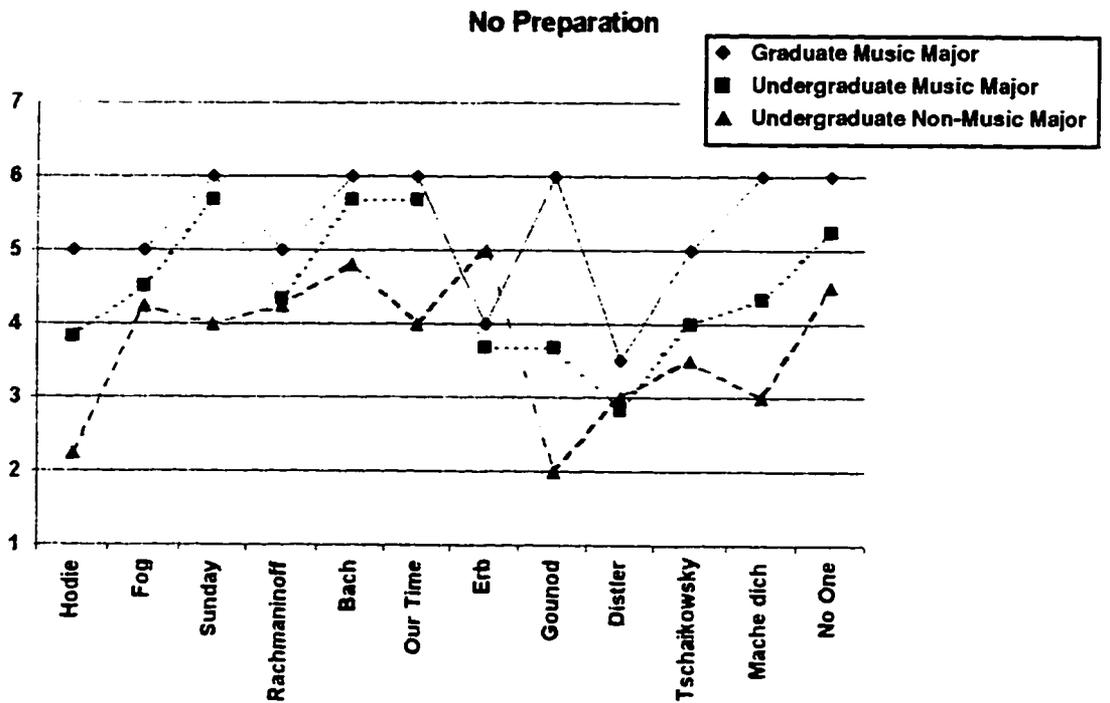
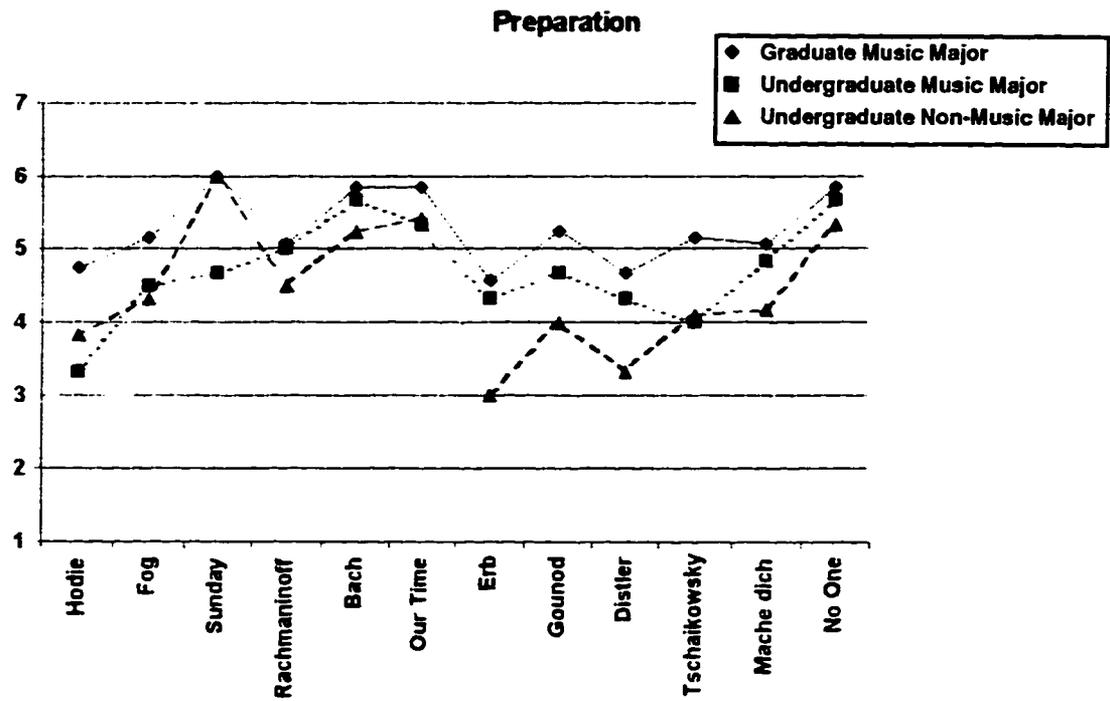
Analysis of Variance for Map Proposition Scores by Division, Preparation and Piece

Source	SS	df	MS	F
Between subjects				
Division	71.18	2	35.59	7.27**
Preparation	9.14	1	9.14	1.87
Division X Preparation	8.01	2	4.00	.82
Within	176.27	36	4.90	
Within subjects				
Piece	135.79	11	12.34	11.52**
Division X Piece	22.11	22	1.00	.94
Preparation X Piece	7.64	11	.69	.65
Division X Prep X Piece	46.49	22	2.11	1.97*
Within Cells	424.42	396	1.07	

* $p < .01$, ** $p < .001$

This preliminary analysis, similar to the results of the total scoring procedure revealed no significant differences between the self-reported variable of those who prepared versus those who did not ($n=42$). Division differences in this analysis for $F(2,36) = 7.27$, $p < .002$ were highly significant. The within subject effect of piece showed similar results with $F(11, 396) = 11.52$, $p < .001$. An interesting finding in this analysis was the significant three-way interaction. The graphing presentation of the interaction is presented in Figure 4.

Figure 4
Graph of Three-Way Interaction for Propositional Elements



This figure clearly shows that relationship among mean scores is different for those students that indicated preparation of some sort as compared to those who reported no preparation. This discrepancy results in a significant three-way interaction. Homogeneity of score is more evident in those students who prepared than those students who reported no preparation. However, given that the main effect of preparation was not significant this variable was dropped and the analysis re-run to capitalize on the available data. Piece means and standard deviations are given in Table 13.

Table 13

Means and Standard Deviations by Piece for Propositional Elements (n=45)

<u>Piece</u>	<u>Title</u>	<u>Mean</u>	<u>Standard Deviation</u>
1	Hodie Christus natus est	3.98	1.59
2	Fog	4.69	1.13
3	Sunday (S1)	5.47	.89
4	Bogoróditse Dévo	4.76	1.35
5	Singet dem Herrn	5.54	.87
6	Our Time (S2)	5.52	.85
7	Shenandoah	4.11	1.57
8	Sanctus and Benedictus	4.29	1.59
9	Lobe Den Herren	3.73	1.42
10	Chvalite Gospoda	4.37	1.38
11	Mache dich auf, werde licht	4.64	1.37
12	No One Is Alone (S3)	5.49	.97

As the table indicates, the highest propositional mean score was found for the Bach (5.54) and the lowest mean score for the Distler (3.73), resulting in a range of 2.81. The Sondheim pieces were closest in scoring to the Bach, with “Sunday,” “Our Time” and “No One is Alone” reporting scores of 5.47, 5.52 and 5.49, respectively.

Table 14 reports the Division means and standard deviations by piece. Again, highest scores are noted for graduate music majors. The hypotheses for the three main effects and interactions were

MAIN EFFECTS

Ho₁ - There is no statistically significant difference among mean propositional scores for the three divisions of students ($\alpha = .05$).

Ho₂ - There is no statistically significant difference among mean propositional scores for the 12 pieces of music ($\alpha = .05$).

INTERACTION

Ho₃ - There is no statistically significant interaction between division and piece in their influence on mean propositional scores ($\alpha = .05$).

Table 14

Mean Element Scores for Pieces across Division

Piece	Title	<u>Division</u>			
		Undergraduate Non-Majors (n=10)	Undergraduate Music Majors (n=21)	Graduate Music Majors (n=14)	
1.	Hodie	<u>M</u>	3.20	3.81	4.76
		<u>SD</u>	1.55	1.33	1.72
2.	Fog	<u>M</u>	4.30	4.57	5.14
		<u>SD</u>	1.25	1.20	.77
3.	Sunday (S1)	<u>M</u>	5.20	5.24	6.00
		<u>SD</u>	1.03	1.00	.00
4.	Bogoróditse Dévo	<u>M</u>	4.40	4.71	5.07
		<u>SD</u>	1.65	1.10	1.49
5.	Singet dem Herrn	<u>M</u>	4.95	5.62	5.86
		<u>SD</u>	1.12	.81	.54
6.	Our Time (S2)	<u>M</u>	4.86	5.62	5.86
		<u>SD</u>	1.00	.81	.54
7.	Shenandoah	<u>M</u>	3.80	4.00	4.50
		<u>SD</u>	1.75	1.55	1.51
8.	Sanctus and Benedictus	<u>M</u>	3.20	4.10	5.36
		<u>SD</u>	1.69	1.41	1.15
9.	Lobe Den Herren	<u>M</u>	3.20	3.48	4.50
		<u>SD</u>	1.03	1.47	1.35
10.	Chvalite Gospoda	<u>M</u>	3.86	4.10	5.14
		<u>SD</u>	.68	1.41	1.46
11.	Mache dich auf	<u>M</u>	3.70	4.71	5.21
		<u>SD</u>	1.77	1.19	.98
12.	No One Is Alone (S3)	<u>M</u>	5.00	5.48	5.86
		<u>SD</u>	1.41	.87	.54

The between-within analysis of propositional elements is reported in Table 15. The between-subject effect of group was significant ($F(2, 42) = 9.13, p = .001$), showing stronger evidence of group differences than the total score analysis. The within effect of piece was also highly significant ($F(11, 462) = 14.89, p = .0001$).

Table 15

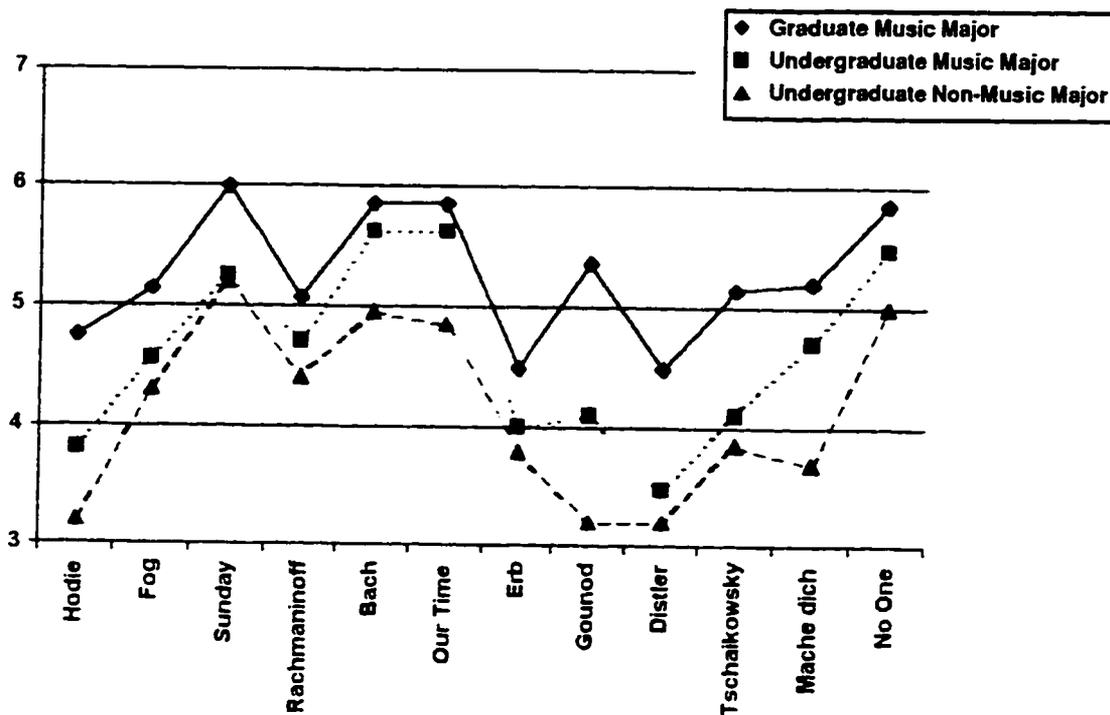
Analysis of Variance for Total Propositional Element Scores by Division and Piece

Source	SS	df	MS	F
Between subjects				
Division	94.61	2	5.18	9.13*
Within Cells	217.66	42	5.15	
Within subjects				
Division	188.14	11	17.10	14.89**
Group by Piece	20.38	22	.93	.81
Within Cells	530.63	462	1.15	

* $p < .01$, ** $p < .001$

The group by piece interaction was not significant in this analysis. Graphic depiction of the propositional mean scores by division and piece are presented in Figure 5.

Figure 5
Graph of Propositional Means by Piece and Division



The Scheffé post hoc analysis, similar to the total score procedure, did not reveal any significant differences in mean comparisons between division scores. The Scheffé analysis is presented in Table 16.

Table 16

Scheffé Post Hoc Analysis of Group Means

	Undergraduate Non-Major ($n=10$)	Undergraduate Music Major ($n=21$)	Graduate Music Major ($n=14$)
Undergraduate Non-Major		.08	1.58
Undergraduate Music Major			1.41
Graduate Music Major			

As in the previous analysis, to compensate for the small range in overall mean values, a one-factor ANOVA was run utilizing the sum of total scores for each division. This analysis, yielding the same result as the between factor F test, was significant, $F(2,42) = 9.13, p < .05$. The post hoc analysis revealed a significant difference between the undergraduate non-music majors and the graduate music majors and the undergraduate music majors and the graduate student majors but no difference between the undergraduate groups.

As the three propositional elements differed in terms of prior knowledge available to students (i.e., composer information was readily available to students throughout the semester, as pieces were frequently identified by the composer) and propositional elements were not deemed to be at the same levels of difficulty (i.e., composer should be easier for students to remember than the date or time period in which a piece was composed), an analysis of elements was deemed appropriate.

To study differences among the number of points earned for the three basic areas,

total points for composer, date piece was written and type of composition were scored across all pieces for all students and comparisons made for groups and elements. A one between and one within factor analysis yielded the following means and standard deviations, presented in Table 17.

Table 17

Means and Standard Deviations of Total Scores on Elements By Division

Proposition	<u>Division</u>		
	Undergraduate Non-Major (<u>n</u> =10)	Undergraduate Music Major (<u>n</u> =21)	Graduate Music Major (<u>n</u> =14)
Composer			
<u>M</u>	20.20	22.38	22.21
<u>SD</u>	4.57	1.96	2.16
Date Composed			
<u>M</u>	10.60	13.71	19.86
<u>SD</u>	3.81	6.58	5.11
Type of Composition			
<u>M</u>	16.30	19.57	20.93
<u>SD</u>	4.50	2.84	2.27

The expected results indicate mean scores were highest for the composer of the piece, with little variation displayed across division. The next highest set of scores was generated for the type of composition, with more discrepancy noted among division, and the lowest scores, as anticipated, were for the date or period in which a piece was composed. This would be expected as information given about the dates and period surrounding a piece were given only one or two times during the course of the preparatory sequence, whereas the type of composition was referred to more frequently during the rehearsal of any given piece. As stated earlier, the composer was frequently utilized in identifying the selection and was utilized as frequently as the actual title of the

piece. Hypotheses for this analysis were

MAIN EFFECTS

Ho₁ - There is no statistically significant difference among mean scores for the three divisions of students ($\alpha = .05$).

Ho₂ - There is no statistically significant difference among mean scores for the three propositional elements ($\alpha = .05$).

INTERACTION

Ho₃ - There is no statistically significant interaction between division and propositional elements in their influence on mean scores ($\alpha = .05$).

The repeated measure analysis of variance found a significant difference among group means $F(2, 42) = 9.92, p < .0001$, among elements $F(2, 84) = 40.37, p < .0001$ and for the group by proposition interaction $F(4, 84) = 4.67, p < .002$. The repeated measure analysis of variance is presented in Table 18.

Table 18

Analysis of Variance for Division by Propositional Elements

Source	SS	df	MS	F
Between Subject				
Division	493.45	2	246.72	9.92**
Within	1044.52	42	24.87	
Within Subject				
Proposition	987.22	2	493.61	40.37**
Division X Prop	228.58	4	57.15	4.67*
Within	1026.96	84	12.23	

* $p < .001$, ** $p < .0001$

To follow up this analysis, post hoc comparisons were run on division, utilizing the Scheffé procedure. These results are presented in Table 19. All pairwise comparisons were significant at the .01 level of significance. This indicates that differences were found between undergraduate non-music majors and undergraduate music majors, undergraduate non-music majors and graduate music majors and for undergraduate music majors and graduate music majors.

Table 19

Scheffé Post Hoc Analysis of Group Means for Propositional Elements

	Undergraduate Non-Major (<u>n</u> =10)	Undergraduate Music Major (<u>n</u> =21)	Graduate Music Major (<u>n</u> =14)
Undergraduate Non-Major		16.59**	59.77**
Undergraduate Music Major			22.55**
Graduate Music Major			

The total scores for the propositional elements were based on a total of 45 scores. The post-hoc analysis for this main effect utilized Tukey's procedure for making pair-wise comparisons. The results of this analysis are presented in Table 20 and indicate that pair-wise comparisons were significant for all pair-wise combinations.

Table 20

Tukey Post Hoc Analysis of Means for Propositional Elements (n=45)

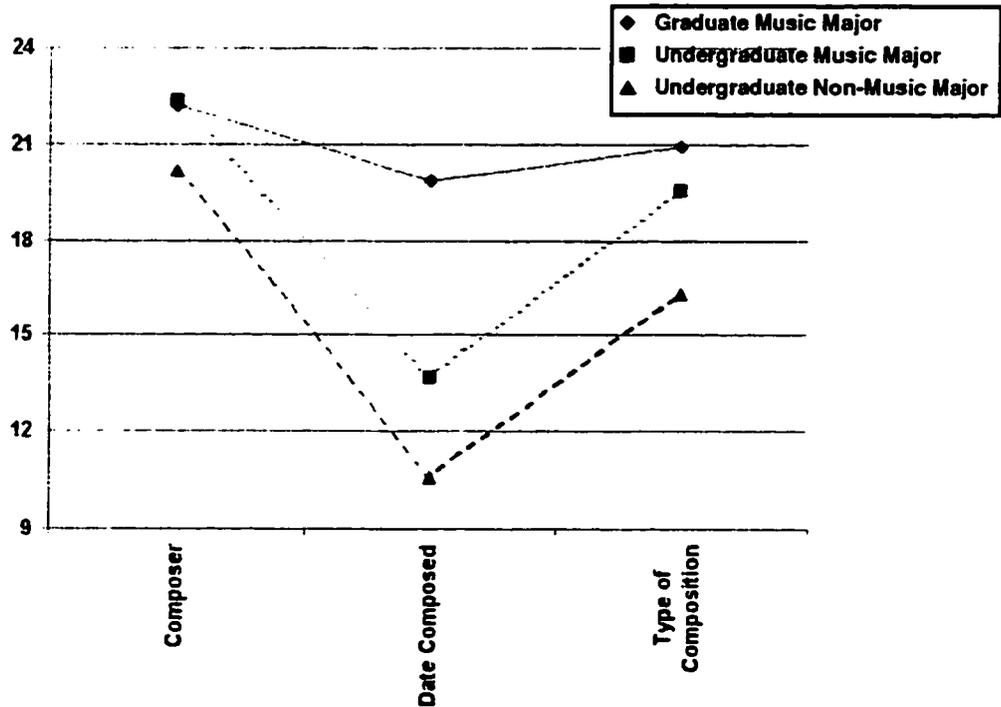
	Composer	Period Piece Was Written	Type of Composition
Composer		13.68**	5.31**
Period			8.36**
Type of Composition			

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

The significant interaction warranted a closer look at the cell means for this analysis. The graph of the cell means is presented in Figure 6.

Figure 6

Graph of Cell Means for Propositional Elements



The test of simple main effects consisted of three one-way ANOVAs comparing the division means across levels of the propositional elements. These analyses revealed no significant differences among mean scores for composer $F(2,42) = 2.26, p > .05$. However, the analyses for the period in which the piece was written and the type of composition were both significant, with $F(2,42) = 10.03, p < .01$ and $F(2,42) = 6.91, p < .05$, respectively. These analyses are presented in Table 21.

Table 21

Test of Simple Main Effects for Composer, Date Composed and Type of Composition

<u>COMPOSER</u>				
Source	SS	df	MS	F
Division	35.00	2	17.50	2.26
Within	324.91	42	7.74	
<u>DATE COMPOSED</u>				
Source	SS	df	MS	F
Division	128.89	2	64.45	6.91*
Within	391.91	42	9.33	
<u>TYPE OF COMPOSITION</u>				
Source	SS	df	MS	F
Division	585.32	2	292.66	10.03**
Within	1225.92	42	29.19	

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

Pairwise comparison of the differences between means for division revealed significant differences between undergraduate non-majors and graduate music majors and undergraduate music majors and graduate music majors, but no difference between undergraduate music and non-music majors. Post-hoc analysis for the type of composition revealed only one significant pair-wise combination between undergraduate non-music majors and graduate music majors.

Levels, Crosslink and Error Patterns

To understand components evident in maps, the entire set of maps ($n=48$) was utilized to compute percentages of levels, crosslinks and error patterns that were present. This descriptive data revealed that students were able to supply a great deal of information beyond the initial propositional elements.

Table 22

Level Percentages by Piece for All Students

Piece	Title	Composer	Level One	Level Two	Level Three
1	Hodie Christus natus est	Schütz	54.60	10.40	0.00
2	Fog	Zimmerman	52.10	0.00	0.00
3	Sunday (S1)	Sondheim	85.40	16.70	0.00
4	Bogoróditse Dévo	Rachmaninoff	85.40	8.30	0.00
5	Singet dem Herrn	Bach	93.80	20.80	2.10
6	Our Time (S2)	Sondheim	83.30	8.40	0.00
7	Shenandoah	arr. Erb	64.60	2.10	0.00
8	Sanctus and Benedictus	Gounod	54.20	2.10	0.00
9	Lobe Den Herren	Distler	66.70	2.10	0.00
10	Chvalite Gospoda	Tschaikowsky	70.80	8.40	0.00
11	Mache dich auf	Zimmermann	68.80	10.40	0.00
12	No One Is Alone (S3)	Sondheim	89.60	2.10	0.00

Mapping levels indicated that most students were able to generate concepts at Level 1 for the Bach (93.8%), and 20.8% of all students were able to go into Level 2 on that particular piece. Level 3 was only evident on the Bach (2.1%).

Crosslink and error patterns were also evident in the maps. These percentages are given in Table 23. The linking of information across elements was limited. This may have been due to the lack of time that students had to complete maps. Highest crosslink rates were once again found with the Bach motet.

Table 23

Crosslink and Error Percentages by Piece for All Students

Piece	Title	Composer	Crosslinks	Errors
1	Hodie	Schütz	2.10	56.30
2	Fog	Zimmerman	0.00	33.40
3	Sunday (S1)	Sondheim	4.20	16.70
4	Bogoróditse Dévo	Rachmaninoff	2.10	35.40
5	Singet dem Herrn	Bach	12.50	33.40
6	Our Time (S2)	Sondheim	0.00	8.40
7	Shenandoah	arr. Erb	0.00	16.70
8	Sanctus and Benedictus	Gounod	0.00	14.60
9	Lobe Den Herren	Distler	0.00	35.50
10	Chvalite Gospoda	Tschaikowsky	2.10	39.60
11	Mache dich auf	Zimmermann	0.00	14.60
12	No One Is Alone (S3)	Sondheim	0.00	4.20

Errors were evident in all maps, with the highest error rates found in the Schütz piece, followed by the Distler, the Tschaikowsky, the Rachmaninoff and the Zimmerman “Fog.” Individual analysis of these errors revealed that most misconceptions were related to dates, the period of time in which a piece was composed and, in the Schütz, confusion over the type of piece and stylistic elements.

CHAPTER 5

DISCUSSION AND RECOMMENDATIONS

The research question in this study addressed the issue of the usefulness of concept maps as an assessment practice designed to give evidence of cognitive learning following a preparatory sequence in a performing ensemble. Specifically, the research questions addressed in this study were the following:

1. Will concept maps discriminate among groups of students within a performing ensemble?
 - a. Will student scores in a performing ensemble at the university level differ across undergraduate music majors, undergraduate non-music majors and graduate music majors?
 - b. Will preparation (i.e. studying or review of related elements) affect scores on maps?
2. What elements are present in the student-devised maps, and how are these elements related to the individual pieces and to the rehearsal time?
 - a. Are concept maps sensitive to differences among pieces in terms of overall scores?
 - b. Is there a relationship between rehearsal time for a given piece and complexity of maps generated by students?
 - c. What specific elements are evident in student-generated maps, and is there a relationship among groups or among pieces?
 - d. Are errors present in maps generated by students, and is there a relationship among groups or among pieces?

The mapping procedure was able to discriminate among students in the ensemble. For the total mapping procedure, statistical differences were found between undergraduate non-music majors and graduate music majors. This was an expected result. However, there were no statistically significant differences found between the two undergraduate groups (music and non-music majors) and the undergraduate music majors and the graduate music majors. This was not expected but could be explained in at least two ways. The differences between undergraduate non-music majors and undergraduate music majors may not be extensive due to the fact that all undergraduates who successfully audition for an ensemble of this caliber will have an extensive background in music regardless of whether music is chosen as the primary choice of study. The second possible explanation for these results may be the limitation of time, which did not allow for the display of differences to appear in the maps. In other words, time may have prevented some students from completing all maps utilizing all available knowledge.

Preparation did not appear to affect the scores for the total scoring procedure. However, this variable was based on self-report and there was no way to discern how much or how little students prepared for the mapping exercises. It is important to note that all terminology presented to students had been discussed at some point during the preparatory sequence. Basic elements were present in most maps, but depth of structure, which could be related to prior study, might show significance if students were allowed more time to construct maps.

Analysis of map elements showed extreme sensitivity to differences among pieces in terms of overall scores. The Bach motet showed the greatest amount of information gained and certainly commanded the majority of the rehearsal time during the preparatory sequence. Both students and director indicated that the Bach motet was the selection in which they learned the most over the course of the semester.

Time spent on pieces in rehearsal was significantly and positively correlated with the overall scores on pieces. An interesting element of this analysis was the fact that the

Sondheim pieces generated the second highest set of scores overall. These pieces were linked to stories and pictures presented in class and with a rehearsal strategy utilizing verbal imagery. This music was also familiar to the students, drawing from American musical theater, and the text, in English, seemed to draw an affective or emotional response from the choir. The relationship between literary connections, use of imagery and salience of the piece to the ensemble and how this relates to overall scores on maps was not explored in this study but would be an interesting exploration for future research.

The second analysis, run on the propositional elements, dealt with information that all students had the opportunity to complete. The differences among the three divisions were highly significant with significant pair-wise comparisons between graduate students and both groups of undergraduates but, again, no differences between the undergraduate groupings. Again, for this basic set of information, differences between the two undergraduate divisions may have been due to time factors or the high caliber of students in the course.

Differences among all three propositional elements were evident in the results of this analysis. Most students were successful in identifying the composer of the piece, which was expected. Results of the date or period in which each piece was written showed the lowest mean scores, also expected, with significant differences between graduate students and both divisions of undergraduates, but no differences between the two groups of undergraduates. Students were more successful in identifying the type of composition, but differences here were only evident between graduate students and undergraduate non-music majors. These results support the preceding analyses.

The proportion of levels present in the maps clearly favored the Bach motet, with 93.8% of students generating additional information at Level 1 and 20.8% at Level 2. Level 3 was only evident on the Bach. This may be attributed to the complexity of this work and the time spent in rehearsing this piece. Additional information was present, overall, in greater quantity for the four German selections, again supporting the overall

score analyses and the time spent on the pieces during the rehearsal sequence. Crosslinks were not common but, again, they were most evident in the Bach motet. Again, this may have been due to time limitations. Errors were found in all pieces, mostly related to the date or period in which the pieces were composed. This was evident in the lower scores on this particular propositional element.

The results of this study indicate that this technique was useful in that it provided a sensitive measure of cognitive information for the ensemble members and provided useful feedback to the director in terms of misconceptions in understanding of terminology, facts and concepts presented during the course of the semester. Mapping gave ensemble members an opportunity to integrate prior knowledge and relate it to the repertoire studied during the semester. Qualitative information gained from informal interviews with members indicated that they felt the technique was useful in organizing and in understanding the breadth and depth of information they had learned. This mapping experience also promoted an awareness of the cognitive aspects of the class. One member commented that, before engaging in the mapping experience, she was not aware of how much she had actually learned in addition to the performance aspects of the music.

The sensitivity of the maps was clearly evident in that graduate students showed a broader knowledge base in all areas, which was expected. While overall scores did not reveal statistically significant differences among the two undergraduate groups, propositional analysis revealed that undergraduate music majors did show greater strengths in the area of dates and type of composition as compared to undergraduate non-majors. The former observation was not unexpected, in that undergraduate non-music majors who are selected for an auditioned ensemble clearly display a high degree of musical skill, which would indicate a strong background in the area of music. As such, the homogeneity of the undergraduate students would reduce the power of finding differences.

Future Research

This study had several limitations that can be addressed in future research efforts. The first of these is related to the construction of the maps. Students only received a cursory introduction to mapping. Even with the example provided, few students, similar to results reported by Schick (1991) and Trowbridge and Wandersee (1994), utilized linking words, established a clear hierarchical construction or established crosslinks. Students generally favored linear construction, which has been shown to be a common presentation in the maps of novices. Novak and Gowin (1984) indicated that those familiar with the domain can easily interpret maps without linking words, but the use of linking words does provide a measure of how the student connects the propositional elements and concepts. This would be most critical for the third propositional element in this study related to the type of piece and the elements and rehearsal strategies utilized. The first two propositions dealt with information related to the individual composers and the dates the pieces were composed. The clarity of nodes connected in these areas was greater than in the third propositional area. If one considers that mapping is a skill that needs to be developed over time (Surber, 1984), then continued instruction in the construction of maps, followed by increased opportunity to practice this skill, should result in better constructed maps.

The second issue, related to the first, is the element of time. With only one class period available for students to complete maps, time must be considered as a factor in the overall mapping scores. The propositional analyses could be considered more valid indicators of the overall group in that students were asked to complete all propositional elements, but time may have prevented many students from completing maps, thus lowering scores and misrepresenting the overall amount of knowledge for any individual student. Future work utilizing this technique should allow ample time for students to complete maps. This could be accomplished by allowing students to work on maps

outside of class time (i.e., individual projects) or focusing on only a portion of the repertoire, thus limiting the number of maps that students construct at the end of a semester.

Additionally, the scoring system needs to be refined. While this study provided a basic system of scoring, refinements in terms of establishing linking words and awarding points only for valid linkages between concepts would allow for even greater sensitivity in interpretation and understanding. The evolution of more complete scoring systems is clearly evident in the literature of the sciences, which have utilized and researched this technique for the last two decades. As more studies investigate the use of this technique in the arts, the scoring system should evolve in this domain as well.

This study did not provide a control for prior knowledge. A pre-post mapping study would allow for control of this variable and provide a more powerful test of actual learning that was gained from the chosen repertoire. Such a study would enable students to integrate prior knowledge of a given composer, understanding of interpretation for a given period and general learning strategies related to given periods, prior to the introduction of the actual music. This type of experience would be extremely useful in focusing student awareness of how prior knowledge is linked to a new set of repertoire and might result in increased efficiency of rehearsal time.

However, the use of this technique is clearly a valuable asset to the assessment practices of instructors who teach performance based classes. Lehman (1992), Brophy (1993), Gray (1993), and McCoy (1991) advocate the inclusion of assessment measures that are cognitive in focus and will support and reinforce learning in the music classroom. Authentic assessment and the need for cognitive criteria are clearly needed in the arts. Concept mapping provides a sensitive measure of student thought and organization of concepts. This device provides both summative and formative assessment in the classroom.

The uses of this technique are clearly adaptable to other areas of music education,

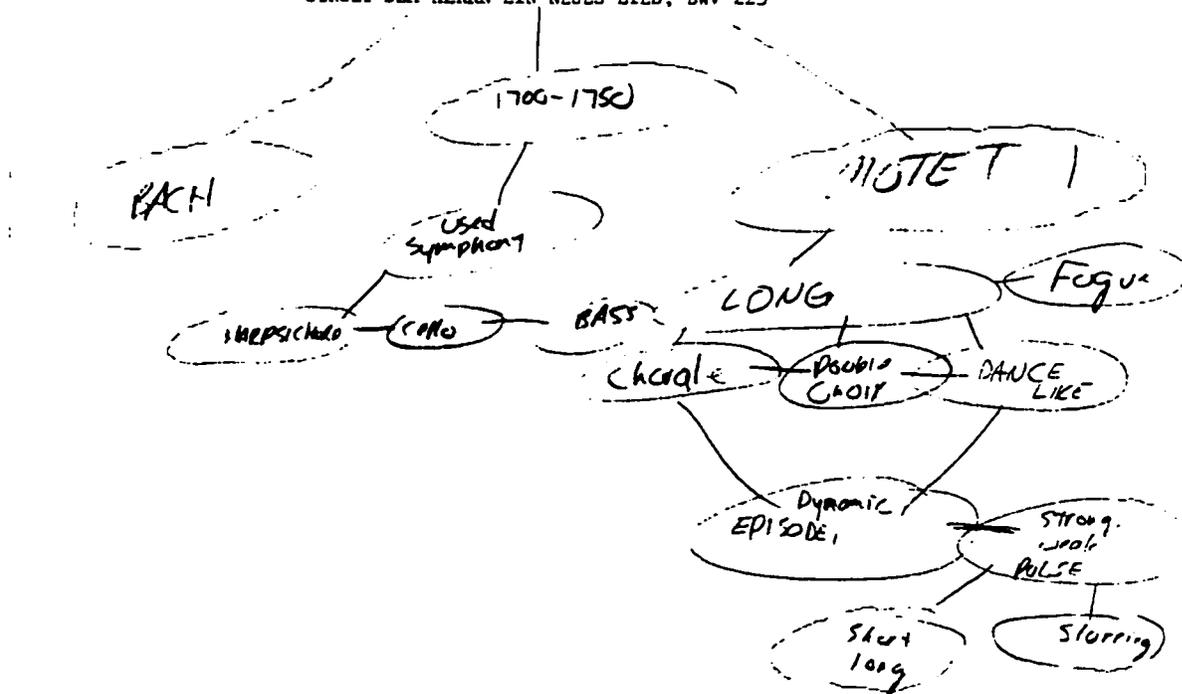
both in the overall learning process and as a method of assessment. This type of technique would be extremely useful for graduate students majoring in conducting. By changing the propositional elements, conducting students could incorporate their knowledge of a given composer, general rehearsal strategies and specific strategies aimed at valid interpretation of a given work. Incorporating this technique prior to the introduction of a piece of music would provide both novice and expert conductors a visual template of how to interpret, plan and proceed with the learning sequence. This could be used by students to prepare for their own conducting experiences or as a means of assessment following a course in conducting.

Conclusion

The usefulness of the concept map has an extensive history in the literature of several educational domains, especially science and reading. The purpose of this study was to explore the usefulness of this technique, specifically as an assessment tool, in the performing arts. This study was successful in pointing out the possibilities of this technique, but clearly revision and refinement are necessary. The emphasis on accountability in our teaching, connected to the accountability of arts programs, and the new focus on cognitive principles of instruction demand that we continue to look for methods of assessment that provide an individualized look at the student. Mapping is an effective tool that allows the student to access cognitive components and cognitive structure of a given domain, which in turn enhances the instructor's ability to modify instruction and correct misconceptions. The results of this study certainly warrant continued investigation of this technique for the performing arts.

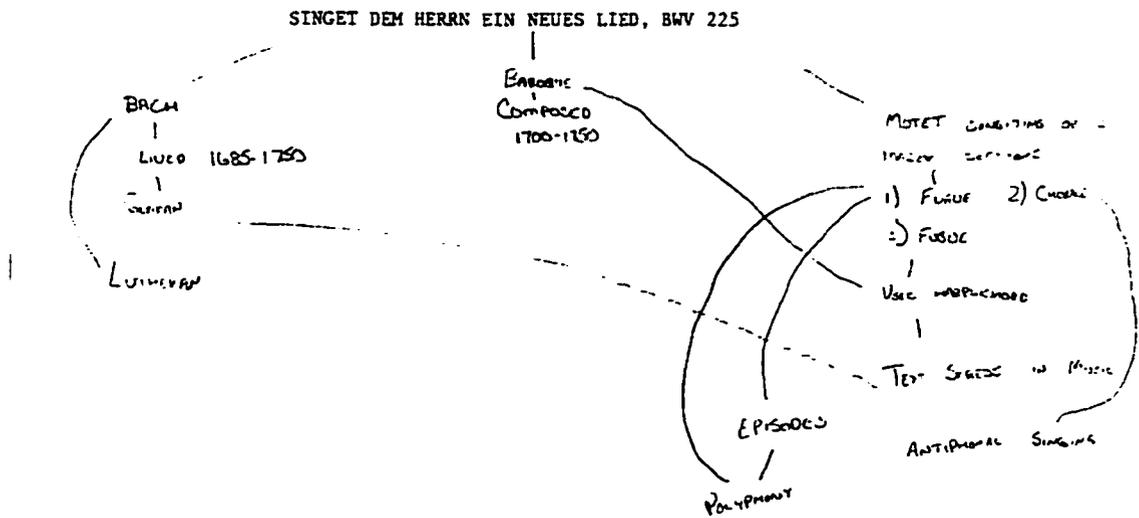
Appendix

SINGET DEM HERRN EIN NEUES LIED, BWV 225



<u>Scoring</u>	
Propositional Elements:	6
Level One:	10
Level Two:	0
Crosslink:	0
Total:	16

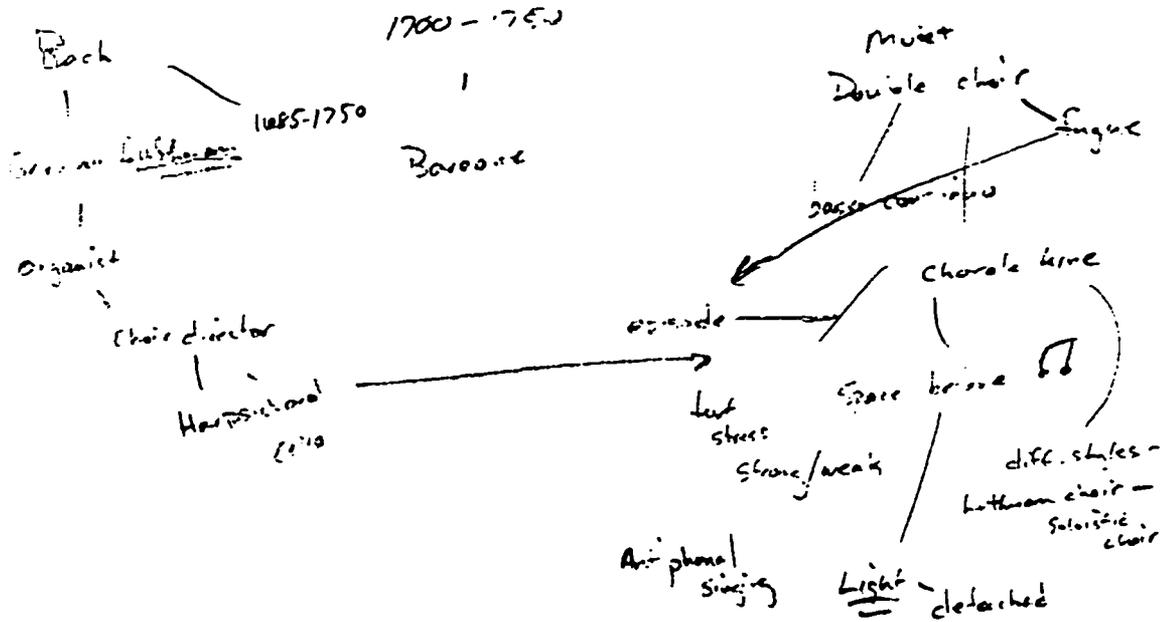
Sample Map: Undergraduate Non-Music Major



<u>Scoring</u>	
Propositional Elements:	6
Level One:	10
Level Two:	5
Crosslink:	10
Total:	31

Sample Map: Undergraduate Music Major

SINGET DEM HERRN EIN NEUES LIED, BWV 225



<u>Scoring</u>	
Propositional Elements:	6
Level One:	10
Level Two:	10
Crosslink:	0
Total:	26

Sample Map: Graduate Music Major

REFERENCES

- Armbruster, B. B. & Anderson, T. H. (1984). Mapping: Representing Informative Text Diagrammatically. In C. D. Holley & D. R. Dansereau (Eds.), Spatial learning strategies: Techniques, applications, and related issues (pp. 189-209). Orlando: Academic Press.
- Ashcraft, M. H. (1989). Human memory and cognition. Glenview, IL: Scott, Foresman.
- Brophy, T. (1993). Evaluation of music educators: Toward defining an appropriate instrument. (ERIC Document Reproduction Service No. ED 375 029)
- Brown, C. F. (1993). Capture the portfolio: Episodes of student musical learning. Paper presented at the Annual Conference of the National Middle School Association, Portland, OR. (ERIC Document Reproduction Service No. ED 375 030)
- Chan, A. S., Butters, N., Salmon, D. P. & McGuire, K. A. (1993a). Dimensionality and clustering in the semantic network of patients with Alzheimer's disease. Psychology and Aging, 8 (3), 411-419.
- Chan, A. S., Butters, N., Paulsen, J. S., Salmon, D. P., Swenson, M. R. & Maloney, L. T. (1993b). An assessment of the semantic network in patients with Alzheimer's disease. Journal of Cognitive Neuroscience, 5 (2), 254-261.

College Board (1994). Arts education assessment framework (Report No. RN-92167001). Washington, D.C.: National Assessment Governing Board. (ERIC Document Reproduction Service No. ED 374 073)

Crocker, L. & Algina, J. (1986). Introduction to classical and modern test theory. Fort Worth, TX: Holt, Rinehart and Winston, Inc.

Davis, K. L. (1983) Mastery learning and music performing groups. (EDRS Report No. 241 400). Southwest Missouri State University.

Fellows, N. J. (1994). A window into thinking: Using student writing to understand conceptual change in science learning. Journal of Research in Science Teaching, 31 (9), 985-1001.

Gagné, E. D., Yekovich, C. W. and Yekovich, F. R. (1993). The cognitive psychology of school learning (2nd ed.). New York: HarperCollins College Publishers.

Garman, B. R. and others (1992). Orchestra festival evaluations: Interjudge agreement and relationships between performance categories and final ratings. Research Perspectives in Music Education, 2, 19-24. (ERIC Document Reproduction Service No. ED 360 204)

Gray, L. (1993). Portfolio assessment: The ARTS PROPEL project. General Music Today, 6 (3), 9-14.

Jonassen, D. H. (1992). Evaluating constructivistic learning. In T. M. Duffy & D. H. Jonassen (Eds.), Constructivism and the technology of instruction: A conversation (pp. 137 - 148). Hillsdale: Lawrence Erlbaum Associates.

Kirk, R. E. (1982). Experimental design: Procedures for the behavioral sciences (2nd ed.). Pacific Grove: Brooks/Cole Publishing Company.

Lambiotte, J. G., Danseareau, D. R., Cross, D. R., & Reynolds, S. B. (1989). Multirelational semantic maps. Educational Psychology Review, 1 (4), 331-367.

Lamble, W. H. (1982). The development and evaluation of a model minimum competency testing program in high school choral music (Doctoral dissertation, University of Iowa, 1982). Dissertation Abstracts International, 43, DA8222250.

Larkin, J. H. (1994). Forward. In K. McGill (Ed.), Classroom Lessons: Integrating cognitive theory and classroom practice. Cambridge: MIT.

Lehman, P. R. (1992a). Assessing Learning in the Music Classroom. NASSP BULLETIN, 76 (544), 56-62.

Lehman, P. R. (1992b). Curriculum and program evaluation. In R. Colwell (Ed.), Handbook of research on music teaching and learning (pp. 281-294). New York: Schirmer Books.

Levinson, J. (1987). Evaluating musical performance. Journal of Aesthetic Education, 21 (1), 75-88,

Lincoln, Y. S. and Guba, E.G. (1985). Naturalistic inquiry. Newbury Park: Sage Publications.

Liu, X. (1994). The validity and reliability of concept mapping as an alternative science assessment when item response theory is used for scoring. Paper presented at the meeting of the American Educational Research Association, New Orleans, LA.

Lloyd, C. V. (1990). The elaboration of concepts in three biology textbooks: Facilitating student learning. Journal of Research in Science Teaching, 27 (10), 1019-1032.

MacGregor, R. N. (1994). Assessment in the arts: A Cross-Canada study (Tech. Rep. No. 143). Vancouver, Canada: British Columbia University (ERIC document Reproduction Service No. ED 382 560)

Mahler, S., Hoz, R., Fishel, D. & Tov-ly, E. (1991). Didactic use of concept mapping in higher education: Application in medical education. Instructional Science, 20 (1), 25-47.

Markham, K. M., Mintzes, J. J. & Jones, M. G. (1994). The concept map as a research and evaluation tool: Further evidence of validity. Journal of Research in Science Teaching, 31 (1), 91-101.

Matheny, J. (1994). A grading system for secondary music groups. Music Educators Journal, 80 (4), 37-39.

McCoy, C. W. (1991). Grading students in performing groups: A comparison of principals' recommendations with directors' practices. Journal of Research in Music Education, 39 (5), 181-190.

Mergendoller, J. R. & Sacks, C. H. (1994). Concerning the relationship between teachers' theoretical orientations toward reading and their concept maps. Teaching and Teacher Education, 10 (6), 589-599.

Meske, E. B. (1987). Music in the schools. Journal of Aesthetic Education, 21 (4), 41-48.

Mitchell, R. (1992, May). Assessment of learning: Prolegomena to a theory of arts education. Paper presented at the meeting of Arts in American Schools, Annapolis, MD.

Mills, J. (1991). Assessing musical performance musically. Educational Studies, 17 (2), 173-181.

Music Educators National Conference (1996). Performance standards for music. Reston, VA: MENC.

National Endowment for the Arts (1994). Arts education research agenda for the future. Office of Educational Research and Improvement. Washington, DC. (ERIC Document Reproduction Service No. ED 367 587)

Niebur, L. (1994). Assessment as a class activity. Music Educators Journal, 80 (5), 23-25.

Novak, J. D. & Gowin, D. B. (1984). Learning how to learn. Cambridge: Cambridge University Press.

Novak, J. D. (1990). Concept maps and vee diagrams: Two metacognitive tools to facilitate meaningful learning. Instructional Science 19, 1-25.

Novak, J. D. & Musonda, D. (1991). A twelve-year longitudinal study of science concept learning. American Educational Research Journal, 28 (1), 117-153.

Okebukola, P. A. (1990). Attaining meaningful learning of concepts in genetics and ecology: An examination of the potency of the concept-mapping technique. Journal of Research in Science Teaching 27 (5), 493-504.

Okebukola, P. A. (1993). Making college science transparent through the use of concept maps. Paper presented at the annual meeting of the National Association of Research in Science Teaching, Atlanta, GA. (ERIC Document Reproduction Service No. ED 360 167)

Point of View. (1984, March). Music Educators Journal, 70 (7), 59-62.

Rentsch, J. R., Heffner, T. S. & Duffy, L. T. (1994). What you know is what you get from experience: Team experience related to teamwork schemas. Group and Organization Management, 19 (4), 450-474.

Resnick, L. B. & Klopfer, L. E. (1989). Toward the thinking curriculum: An overview. In L. B. Resnick & L. E. Klopfer (Eds.), Toward the thinking curriculum: Current cognitive research (pp. iv-vii). Association for Supervision and Curriculum Development.

Robinson, M. (1995). Alternative assessment techniques for teachers. Music Educators Journal, 81 (5), 28-34.

Roth, W.-M. (1994). Student views of collaborative concept mapping: An emancipatory research project. Science Education, 78 (1), 1-34.

Roth, W.-M. & Roychoudhury, A. (1993). The concept map as a tool for the collaborative construction of knowledge: A microanalysis of high school physics students. Journal of Research in Science Teaching, 30 (5), 503-534.

Schreiber, D. A. & Abegg, G. L. (1991). Scoring student-generated concept maps in introductory college chemistry. Paper presented at the annual meeting of the National Association for Research in Science Teaching, Lake Geneva, WI. (ERIC Document Reproduction Service No. ED 347 055)

Schunk, D. H. (1991). Learning theories. New York: Macmillan.

Schick, C. J. (1992). The use of concept mapping to evaluate the effects of both bibliographic and discipline-based art education instruction on the library skills of elementary art teachers. (Master's thesis, Kent State University, 1992). (ERIC Document Reproduction Service No. ED 352 981)

Songer, C. J. & Mintzes, J. J. (1994). Understanding cellular respiration: An analysis of conceptual change in college biology. Journal of Research in Science Teaching, 31 (6), 621-637.

Starr, M. L. & Krajcik, J. S. (1990). Concept maps as a heuristic for science curriculum development: Toward improvement in process and product. Journal of Research in Science Teaching 27 (10), 987-1000.

Stevens, J. (1992). Applied multivariate statistics for the social sciences (2nd ed.). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Surber, J. R. (1984). Text maps as tests. In C. D. Holley & D.F. Dansereau (Eds.), Spacial learning strategies: Techniques, applications, and related issues (pp. 213-233). Orlando: Academic Press.

Tabachnick, B.G. & Fidell, L. S. (1989). Using multivariate statistics (2nd ed.). New York: HarperCollins.

Torney-Purta, J. (1991). Schema theory and cognitive psychology: Implications for social studies. Theory and Research in Social Education, 19 (2), 189-210.

Trowbridge, J. E. & Wandersee, J. H. (1994). Identifying critical junctures in learning in a college course on evolution. Journal of Research in Science Teaching, 31 (5), 459-473.

Tuley, R. J. (1985). Framework for performance evaluation. Music Educators Journal, 71 (7), 32-33.

Tunks, T. (1987). Evaluation in music education: The value of measurement/the measurement of value. In D. Alexander (Ed.), Research, learning theory, and the ideas of Robert G. Sidnell (pp. 53-59). Urbana, IL: Council for Research in Music Education.

Verrastro, R. E. & Leglar, M. (1992). Music teacher education. In R. Colwell (Ed.), Handbook of research on music teaching and learning (pp. 676-696).

Voyer, J. J. & Faulkner, R. R. (1989). Organizational Cognition in a jazz ensemble. Empirical Studies of the Arts, 7 (1), 57-77.

Wallace, J. D. & Mintzes, J. J. (1990). The concept map as a research tool: Exploring conceptual change in biology. Journal of Research in Science Teaching, 27 (10), 1033-1052.

Webster, P. R. (1992). Research on creative thinking in music: The assessment literature. In R. Colwell (Ed.), Handbook of research on music teaching and learning (pp. 266-280).

Warnick, E. M. (1985). Overcoming measurement and evaluation phobia. Music Educators Journal, 71 (8), 32-40.

Weymuth, R. W. (1986). The development and evaluation of a cognitive music achievement test to evaluate Missouri high school choral music students (Doctoral Dissertation, University of Miami, 1986). Dissertation Abstracts International, 47, DA8706495.

Wiegmann, D. A., Dansereau, D. F., McCagg, E. E., Rewey, K. L. & Pitre, U. (1992). Effects of knowledge map characteristics on information processing. Contemporary Educational Psychology, 17, 136-155.

Willerman, M. & Mac Harg, R. A. (1991). The concept map as an advance organizer. Journal of Research in Science Teaching, 28 (8), 705-711.

Wilson, J. M. (1994). Network representations of knowledge about chemical equilibrium: Variations with achievement. Journal of Research in Science Teaching, 31 (10), 1133-1147.

Winer, B. J. (1971). Statistical principles in experimental design. New York: McGraw-Hill.

Wolf, D. P. (1992, May). The case for portfolios: The work of art, ensemble performance, and hearing the conversations of a culture. Paper presented at the meeting of Arts in American Schools, Annapolis, MA.

Zoller, U. (1990). Students' misunderstandings and misconceptions in college freshman chemistry (general and organic). Journal of Research in Science Teaching, 27 (10), 1053-1065.

VITA

Deborah Carr was born June 9, 1955 in Kalamazoo, Michigan. After attending public school in Michigan and Missouri, she received the following degrees: B.S. in Education from the University of Missouri-Columbia (1977); M.Ed. in School Psychology from the University of Missouri-Columbia (1992); Ph.D. in Educational Psychology from the University of Missouri-Columbia (1996).

WORK EXPERIENCE:

- 1996- Instructor, Educational Statistics I, University of Missouri-Columbia
- 1995-1996 Instructor, Child Development, Teaching Assistant, Educational Statistics II, University of Missouri-Columbia
- 1994 Instructor, Educational Psychology, Continuing Education, University of Missouri-Columbia
- 1993-1994 Instructor, Teaching Assistant, Educational Statistics II, University of Missouri-Columbia
- 1993-1995 President - University Singers Tom Mills Foundation Board of Directors
- 1992-1994 Assistant, Psychological Practice of Scott and Lapan, Columbia, Missouri
- 1992-1993 Teaching Assistant, Educational Statistics I and Psychological Testing, University of Missouri-Columbia

- 1991-1993 Organist, St. Mary's Episcopal Church, Columbia, Missouri
- 1992 Psychological Testing, Juvenile Justice Center, Columbia, Missouri
- 1991-1992 Instructor, Child Development, University of Missouri-Columbia
- 1990-1992 Secretary, Area VIII Board, American Guild of English Handbell Ringers, Inc.
- 1989-1995 Conductor and Musical Director, Columbia Handbell Ensemble, Inc., Columbia, Missouri
- 1988-1991 Assistant Music Coordinator, Community United Methodist Church, Columbia, Missouri
- 1988-1990 Missouri State Chair, American Guild of English Handbell Ringers, Inc., Area VIII
- 1984-1987 Handbell and Youth Choir Director, Missouri United Methodist Church, Columbia, Missouri
- 1982-1984 Accompanist, Monroe Little Theatre, Masterworks Chorus, Monroe, Louisiana
- 1978-1982 Vocal Music Teacher, Normal Public Schools, Normal, Illinois
- 1977-1978 Vocal Music Teacher, Fayette R-III Public Schools, Fayette, Missouri

SPECIAL EVENTS:

- Conductor - Missouri Baptist Handbell Festival, Tan-Tar-A, Lake of the Ozarks, MO - March 1996
- Performance - Columbia Handbell Ensemble, National AGEHR Director's Conference, St. Louis, MO - July 1995
- Conductor - National AGEHR Eastern Conferences, Tampa, FL & Niagara Fall, NY - June 1994

- Conductor - Young Ringers Festival, Missouri AGEHR, St. Charles, MO -
February 1993
- Registrar - Area VIII Conference, AGEHR, St. Joseph, MO - July 1992
Performance - Columbia Handbell Ensemble, Area VIII
Conference, AGEHR, St. Joseph, MO - July 1992
- Conductor - Community Handbell Festival, Frankfort, IL - November 1991
- Class Instructor - National AGEHR Western Conference, Ames, IA - June 1991
- Conductor - National Young Ringers Festival, AGEHR, Ames, IA - June
1991
- Adjudicator - Missouri Baptist Handbell Festival, Columbia,
MO - March 1991
- Class Instructor - Area VIII Conference, AGEHR, Cape Girardeau, MO - June
1990
- Guest Conductor - Marshall Spring Handbell Workshop, Marshall, MO - March
1990

HONORS:

Matthews Scholarship 1977, Music Department, University of Missouri-
Columbia

Ruth Norris Scholarship 1993, College of Education, University of Missouri-
Columbia

Donald K. Anderson Graduate Student Teaching Award 1995, College of
Education, University of Missouri-Columbia

PUBLICATIONS:

Program development, survival and growth: A community ensemble perspective.
Overtones, 41 (6) 22-25.