THE INFLUENCE OF EXPERTISE AND CONTEXT ON THE CATEGORIZATION OF MUSIC: A COGNITIVE SCIENCE STUDY

by

JAMES S. IMHOFF

A DISSERTATION

Presented to the School of Music
and the Graduate School of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Musical Arts

March 1995

"The Influence of Expertise and Context on the Categorization of Music: A Cognitive Science Study," a dissertation prepared by James S. Imhoff in partial fulfillment of the requirements for the Doctor of Musical Arts degree in the School of Music. This dissertation has been approved by and accepted by:

[redacted]

Dr. David P. Doerksen, Chair of the Examining Committee

January 18, 1995

Date

Committee in charge:

Dr. David P. Doerksen, Chair
Dr. Richard Clark
Dr. Randall Moore
Dr. Marian Smith
Dr. Tom Givon

Accepted by:

Vice Provost and Dean of the Graduate School

An Abstract of the Dissertation of

James S. Imhoff for the degree of Doctor of Musical Arts

in the School of Music to be taken March 1995

Title: THE INFLUENCE OF EXPERTISE AND CONTEXT ON THE CATEGORIZATION OF MUSIC: A COGNITIVE SCIENCE STUDY

Approved: [redacted]

Dr. David P. Doerksen, Chair

Recent developments in Cognitive Science have demonstrated that, contrary to traditional thinking, categories are not rigid, feature-defined phenomena. Rather, they are influenced by human experience and by the context in which the categorization takes place. The labels people use to describe music reflect the way they categorize it. In this study, 32 music experts and 32 novices labeled short selections of recorded music. In each group, 16 subjects heard all Western Art selections (Context 1), and 16 heard a mixture of Rock, Blues, Jazz, and Western Art music (Context 2). All subjects used style terms (Classical, Renaissance, Baroque) as labels significantly more often than genre, instrument, or national origin. The results indicate that experts used more specific labels than novices, but context did not have a significant influence on the kinds of labels

used by either group. This implies that musical categories are more stable than suggested by the current Cognitive Science literature.

CURRICULUM VITA

NAME OF AUTHOR: James S. Imhoff

PLACE OF BIRTH: Riverside, New Jersey

DATE OF BIRTH: January 10, 1947

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon
Western Oregon State College
University of Delaware

DEGREES AWARDED:

Doctor of Musical Arts, 1995, University of Oregon

Master of Music Education, 1989, Western Oregon State College

Master of Education in Special Education, 1972, University of Delaware

Bachelor of Arts in Music Education, 1969, University of Delaware

PROFESSIONAL EXPERIENCE:

Assistant Professor, Music Education and Choral Conducting, Crane School of Music, SUNY Potsdam, 1993-

Graduate Teaching Fellow, School of Music, University of Oregon, Eugene, 1989-1993

Choral Director and Music Teacher, North Albany Middle School, Albany, Oregon, 1979-1989

Special Education Teacher, New Castle County School District, Wilmington, Delaware, 1972-1979

High School and Middle School Choral Director, Tower Hill School, Wilmington, Delaware, 1970-1971

Elementary Music Teacher, Hart Elementary, Warminster, Pennsylvania, 1969-1970

Teaching Fellowship, University of Oregon, Eugene, 1989-1993

Federal Fellowship Grant in Special Education, University of Delaware, Newark, 1970-1971

ACKNOWLEDGMENTS

The author expresses sincere appreciation to the Committee members and the many people who offered support in pursuing this research and in preparing this document. David Doerksen, as chair, was a source of constant encouragement, inspiration, and practical editorial advice. Richard Clark provided a questioning mind and the curiosity of a musician not immersed in cognitive science literature but nevertheless insisting on clarity and accuracy. Marian Smith kept the study focused on genuine musical questions and challenged me to rethink any intellectual or esthetic shortcuts I had made for the sake of experimental design.

Tom Givon helped me focus on the conceptual basis for this study, and kept me from drifting away from this base. We spent all too few hours fiddling and picking old time music, and I still have the dollar tossed by a satisfied listener as we played in front of the Fall Creek Bakery.

Randy Moore, although he joined the committee late in the process, was of great help throughout on questions of design and analysis and provided me with the facilities to

run the study. Two people were not available to serve on the final committee, but were vital to the early stages of the study. Randi L'Hommedieu established rigorous scholarly standards for writing and searching the literature. Marjorie Taylor, Professor of Psychology, applied Occam's Razor to my muddled ideas and supplied the elegance and clarity of experimental design.

This document embodies the culmination of a dream that began many years ago. It was inspired and nurtured by two teachers who will always have a tremendous influence on my musical endeavors. Joseph Huszti and Calvin Bourgeault started me down a path--a tao--that was much more than a career or a profession. I did not so much follow it as I kept stumbling back on to it. The highest tribute I can pay them is to pass along to the next generation the inspiration they gave me.

Any pursuit such as this takes place over many years, mingled with the distractions, joys, and pains of everyday life. I have been blessed during those years to have the company and support of my wife, Faith. Through her, I learned of so many things that were every bit as important as my work, and every bit as much fun.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

What kind of music did Mozart write?

This question, many would argue, cannot be answered in a few words. Words cannot capture so broad a sweep of creative genius, and so rich and diverse a body of brilliant music. The answer to this question must embrace many different forms and genre, such as operas, symphonies, concertos, piano sonatas, church music, and chamber music. It must also embrace the composer's output at various stages of his development, from piano pieces written as a child to the mature and elaborate operas. Surely, no single word or short phrase can represent all of this great music.

Music teachers, however, ask and answer just such questions every day. While listening, performing, and responding to music, teachers and students must at some point stop the music and talk. Too much talk, of course, detracts from the essentially musical experience; but this only heightens the need for clarity and efficiency of language. If teachers must assign labels to music, and if their students are to learn these labels and associate

them with the appropriate music, then everyone involved will benefit from an understanding of what labels convey.

In spite of any arguments to the contrary, such questions are asked and answered all the time. Friends discussing favorite recordings, radio stations promoting their specialized format, and graduate level musicology seminars all use verbal labels to describe music. There is a tacit understanding that, with one or two words, one individual can communicate to another information that consists of instruments, lyrics, rhythms, melodies, harmonies, and numerous other musical details.

So it appears that the question can be answered. Most people, however, would agree that it can be answered many different ways. Some individuals might answer the question in general terms, such as "Classical" or "symphonic," some more specific, such as "18th Century, First Viennese School." Others might reflect personal preference, such as "beautiful" or "boring." The answer, in fact, can be used as an indicator of knowledge about Mozart. The person who answers *"singspiel"* probably knows more than the person who answers "Classical." The kind of answer a person gives to such a question reflects his or her experiences with the music of that composer. It reflects, at least to some degree, that person's level of

expertise. The teacher cannot expect the student to adopt spontaneously the same vocabulary without at least some experience, training, and discourse.

Furthermore, the circumstances in which the question arises might make a considerable difference in the answers given. One answer might seem appropriate in a college music appreciation course for non-musicians. Another might be expected in an advanced graduate course on the Viennese Classicist School. Still another kind of response might come out of a casual discussion between friends arguing the relative merits of Mozart, Elvis, Coltrane, and B.B. King. Students talking to each other outside class will probably use different terms than those they use on a music test.

It might seem that, with so many possible ways of answering this question, and so many potential influences on the answer, every response will be different. In this extreme scenario, there would be as many responses given as there were examples played and persons asked. A far more realistic scenario would have at least some people using the same label (or the same kind of label) some of the time. This study is concerned with such patterns, and with the kinds of answers or labels people will use most often.

Statement of the Problem

In practical terms, the research problem here concerns communication between music teachers (experts) and music students (novices). Such communication requires a vocabulary of terms and labels, and a repertory of examples. Suppose a teacher wants the class to remember, understand, or appreciate some aspect of Baroque music. He or she might expect the class to recognize the term Baroque (as opposed to the generic "Classical"), to associate it with certain musical characteristics, and to perceive these characteristics in a recorded example.

Recent cognitive studies in categorization have demonstrated that this set of expectations is more complex than commonly supposed. For example, students and teachers might use the same label, such as Baroque or Classical, but with very different meanings. They might listen to the same music but perceive very different characteristics. Students might even recite the supposedly correct list of characteristics--that is, the definition--but use a very different set of characteristics when they actually listen to the music.

Furthermore, there is evidence to suggest that individuals might label or categorize a given item different ways in different circumstances. The context in

which the item is presented can have an impact on the way a person describes it. A listener might label a passage from a Bach organ fugue "Organ music" if he or she hears it in the midst of a cappella choral music, but label the same piece "Classical" or "Baroque" in the midst of Rock, Blues, and Jazz recordings.

The complexity of this situation and the possibilities for miscommunication call for a careful look at how category labels are used by different people in different settings. A number of studies investigate these issues, and they have established the theoretical and methodological approaches employed here. Specifically, this study examines the kinds of labels experts and novices commonly use when they describe music, and whether they use the same kinds of labels in different settings.

Upper division and graduate music majors represent the expert population. Upper division and graduate non-music majors represent the novice population. Subjects listened to a tape of 40 musical examples, each 20 seconds long. During or immediately after each passage, they responded in writing to the question "What kind of music is this?" There were two different settings or tapes in which subjects heard the music. One was an "all Classical" setting (Context 1 or CTXT1), the other included Rock, Blues, and Jazz music (Context 2 or CTXT2).

The labels used by subjects were examined for several different characteristics including the kind of labels (style, genre, national origin), how specific or generic the labels were, and how often subjects used the same labels.

Research Questions

Research Question #1A

Will subjects use style terms (STs) more often than genre, performing media, national origin, or other kinds of labels?

Research Question #1B

Will Novices (representing the general public) use the generic, basic level term "Classical" significantly more often than other style terms?

Research Question #2A

Will Novices use the generic, basic level term (BLT) "Classical" significantly more often than Experts?

Research Question #2B

Will musical context influence the use of basic level terms? That is, will any difference between the two groups be greater with one tape than with the other?

Research Question #3A

Will Experts use the same label--a shared or equivalent term (SET)-- for a given example significantly more often than Novices?

Research Question #3B

Will musical context influence the use of shared equivalent terms? That is, will any difference between the two groups be greater with one tape than with the other?

Justification of the Study

Music educators at all levels present examples of music to their students, point out salient features of the music, and use appropriate labels. Students are expected to associate these labels with the music they hear and with similar examples they might hear later. "Much of our ability to perceive music and the commonalities of musical

situations is the result of categorizing and storing away what we hear (Edwards, 1988, p. 129).

When a student consistently uses the correct label for an example, the teacher probably assumes that the student understands the definition, the important characteristics, and some subtle aspects of the music. By the same token, if the student persists in using a label that the teacher considers incorrect, the assumption is that such understanding is lacking.

This process of applying categorical labels, according to recent developments in cognitive science, is more complex than traditionally assumed. It is subject to a number of influences including the different levels of expertise of the teacher and the students, the context in which the examples are presented, and the categorical associations already learned from previous experience. This study empirically examines this process, and the results will offer useful information for music educators who are engaged in the process on a daily basis.

Cognitive Studies and Music

Cognitive science has pushed back the boundaries of knowledge regarding human perception, thought, and learning. Unfortunately, only a limited portion of this exciting research has been applied to music education.

The growing body of music cognition literature is based on a rather reductionist approach, where the musical stimuli are usually very short patterns of computer generated tones (Dowling & Harwood, 1986; Deutsch, 1982). The perception of isolated pitches, the grouping of a few notes in a rhythmic context, and the expectancy of the next chord are dominant themes in music cognition research. It is generally assumed that higher order concepts such as style and genre cannot be controlled in the scientific manner in which cognition is usually studied.

When a broad concept such as style is considered, the goal is often to reduce it to its defining characteristics. Cutietta (1993), for instance, argues that styles are differentiated by motion, energy, flow, fabric, and color. Eastlund (1992) uses multidimensional scaling to extract "historical period, complexity, and tempo" (p. 19) as features that differentiate styles. There is evidence, however, that the cognitive process of categorization is more complex than this implies.

The way people label everyday objects (Rosch & Lloyd, 1978), the way they decide that two cases are similar or different (Medin & Wattenmaker, 1989), and the way they organize such information in memory (Anderson, 1990) have been the focus of numerous studies. The research proposed

here will apply to music the principles and practices established by categorization research. The musical examples will be brief excerpts, but they will be actual recordings of real compositions, not computer-generated tone patterns.

Contributions of this Study

This dissertation will provide useful information for the development of new approaches to the teaching of musical concepts. In the past, educators have adopted a paradigm promoted by Bruner (1973) known as the Concept Formation Model. This model makes certain assumptions that have not been supported by recent research. These assumptions include a view of categories as discrete entities, clearly defined by specific features; learning the category simply involves identifying and recognizing the features (Merrill & Tennyson, 1977).

Cognitive science has given us a more complex view of category structure (Gardner, 1985). A category, in this new view, is a flexible, dynamic construct that is subject to influences such as prior knowledge, intuitive judgments of similarity, and context. Furthermore, when a large number of people are asked to categorize natural items (trees, furniture) patterns emerge from their responses. Some category labels are used more often and more readily

than others, and some members of a category are more central and typical than others.

For example, Rosch (1978) found that people would most frequently respond with "robin" when shown a picture of that bird; they were less likely to say "bird," "animal," or "red-breasted robin." In addition, when these various labels were flashed on a screen with the appropriate picture, subjects responded more slowly to the very specific label (red-breasted) and to the very general label (animal). They responded most rapidly to what Rosch called the "basic level" term, in this case "robin."

Tanaka and Taylor (1991) found that these response patterns were different for bird experts, who responded rapidly to the more specific labels. A bird expert, for example, could press the "yes" button as rapidly for sparrow as for bird. A novice took slightly longer to respond to the more specific label. The present study will apply some of the theory and methodology from this literature to the categorization of music.

Unique Features of the Study

In addition to the justifications above, this study uses several elements from the literature in new ways. There are a number of studies that examine expert/novice differences, categorization in different contexts, and

musical styles as categories. The present study, however, pulls these elements together in a new light and employs methodologies not previously applied to music.

Musical Style Categories as Cognitive Percepts

Cognitive literature regarding music is, for the most part, concerned with very localized events such as pitch perception and rhythmic groupings (Serafine, 1988; Dowling & Harwood, 1986; Howell, Cross & West, 1985). Less work has been done with more global issues such as style. Dowling and Harwood briefly mention style, but describe it as "invariants across sets of pieces" (1986, p. 160). This certainly sounds like rigid, feature-based categorization.

A few studies have considered style as a variable (Gardner, 1973; Eastlund, 1990; Brittin, 1991), but these studies again assume discrete categories: A given piece represents a given style for both researcher and subject. There do not appear to be any studies that treat perception of style as a cognitive categorization phenomenon. Gardner's 1973 study constitutes a breakthrough in its admission that style is too complex to identify in terms of specific features. Even so, it does not refer to any of the category literature that was available even at that time.

Cutietta (1993) points out that categorization is the most fundamental process in music perception. He argues against the reductionist approach of teaching separate elements of music (rhythm, melody, harmony). Rather, teachers should be aware of the rapid and intuitive categorization taking place as their students are listening to examples or thinking about music. Cutietta writes that junior high students spontaneously categorize music as opera, church, classical, and rock and roll. Such an approach allows the subjects to express their own category schemas rather than respond to those of the researcher. In this sense, Cutietta's position is consistent with the assumptions of this study. It is an important breakthrough in the study of musical categorization in that it establishes style as a percept, rather than as a composite of separate elements.

Cutietta's research questions, however, are quite different from those of the present study. Having rejected melody, rhythm, and harmony as defining attributes, he replaces them with his own proposed defining attributes: motion, energy, flow, fabric, and color. In his study, he asks subjects to rate music on these features, in effect priming them to attend to these features. It must be noted that there was great consistency in student ratings of the proposed attributes,

and they might indeed prove to be the basis for much of the musical categorization that takes place. They are almost certainly an improvement over the traditional abstractions of melody, harmony, and rhythm.

The present study uses an open-ended task, and subjects will be allowed to respond with their own spontaneous, intuitive labels. That is, they will be free to form their own categories and use their own bases for labeling the musical examples, rather than respond to those of the researcher. At no point will one feature be cued or primed over another, and subjects may attend to those which they consider salient. The purpose here is not to find a more realistic basis for categorization, but to observe the influence of expertise and context on certain aspects of categorization.

Free Response Format

In the present study, each subject chooses his or her own basis for labeling the music. In a related pilot study, for example, graduate music students used style periods, performing groups, composer names, genres, forms, and other bases to label musical examples. In the present study, the responses of music experts are compared to those of the general public, but the question is still open ended: "What kind of music is this?" The responses

not only reveal knowledge and vocabulary, but also aspects of the music on which attention is focused. A response such as "piano music" implies a different focus from a response of "cheerful" or "early Chopin."

At no time, however, is there an attempt to extract a specific feature or list of features that might influence such response patterns. This makes the study fundamentally different from Cutietta (1993) and Eastlund (1992), where the focus was on defining features. Here, the focus is on whether the label used is a style term, how generic or specific it is, and how often individuals within a group use the same label.

Expertise and Categorization

The influence of two factors in particular, context and expertise, will be examined in this process. Numerous studies (Tanaka & Taylor, 1991; Murphy & Wright, 1984) point to differences in the way experts and novices categorize material within a given field. These authors suggest that experts categorize at more specific levels. Barsalou (1989) suggests that categorization will vary from one context to the next, but Brooks (1989) argues that some expertise will lead to more consistent results.

This study presents recorded musical examples to experts and to members of the general population. The

hypotheses predict that a frequently used basis for their response will be some aspect of style. Furthermore, the hypotheses suggest that music experts will categorize examples from their area of expertise at a more specific level and with greater consistency, even in changing contexts.

Many of the previous expert/novice categorization studies used laboratory controlled artificial categories, such as color and shape sets (Murphy & Wright, 1984). So-called expertise was established by allowing some subjects to practice the procedure more than others. Murphy and Wright argue that real-world expertise carries with it not only practice, but theoretical information about the domain of expertise. They used actual clinical psychologists at various levels of experience in their study. Tanaka and Taylor (1991), similarly, drew dog and bird experts from members of local dog and bird watching organizations. Expertise was established not by short term artificial practice, but by years of experience and by the personal recommendations of other members.

Growth in musical expertise is associated with lengthy study, practice, and in many cases with specialization. Nevertheless, it is easy to find a group of people that share some expertise in one field of musical style. Most college music programs emphasize

Western Art traditions fostered in Europe from the fall of the Roman Empire to the twentieth century. While different schools might offer programs in Jazz studies, ethnomusicology, or even popular music, the common denominator is that body of literature generally referred to by the public as "Classical."

Thus there is a group of students and faculty one would easily classify as experts in the field of Western art music. Members of the general public might very well be experts in one style or another, but this expertise is, presumably, randomly distributed. If the target examples are Western art music, then, advanced music students may clearly be considered experts.

Context and Categorization

Barsalou (1989) suggests that categories are subject to substantial influence from the context in which cases are presented: That is to say, the subjects will classify the same items quite differently from one time to the next. Givon (1989) also argues strongly that all categorization takes place in some context, and is duly influenced. This implies that different contexts will yield different category schemes. Thus a piece by Bach might be given one label on a classical music radio station, but a very different label if used on an MTV

video. In fact, it is conceivable that a host of environmental and experiential factors will lead to different labels on different occasions.

Brooks (1989), on the other hand, argues that with some practice, subjects might be more consistent. While he is referring to Barsalou's informal *ad hoc* category tasks (such as "ways to hide out from the mob," and "things to take out of the house in a fire"), his point may be extrapolated: With extensive practice, subjects could become quite consistent. Eventually, those subjects who have gained some expertise might, as a group, be very consistent, even under changing circumstances.

A comparison of experts and novices would almost surely show differences in their categorization practices. But when the additional factor of changing context is introduced, these differences might change in degree and character. That is, the novice responses might change radically under radically different circumstances, while the expert responses, according to Brooks, might remain comparatively stable.

In this study, musical context is defined as the range and variety of styles that surround the target examples. These targets consist of twelve Renaissance, Baroque, and Romantic selections from the Norton Anthology of Western Music (1988). In one context (CTXT1), they are

preceded and mixed in with examples from the same source and in similar styles. In the other context (CTXT2), they are preceded by and mixed in with Jazz, Rock, and Blues examples.

Theoretical Assumptions

This study requires several assumptions regarding categorization, some of which might be contrary to popularly held notions. Chief among these assumptions is the degree to which category labels are stable, unique to an example, and unique to a subject. At one extreme is the position that a given item will be categorized the same way by every subject; the rigid Platonic feature-defined model of categories suggests something along these lines (Givon, 1989). At the other extreme, every individual will categorize a given item according to his or her own special ideas about the item. This would emphasize the influence of context (Barsalou, 1989) and individual intuitive theories (Medin & Wattenmaker, 1989).

Neither extreme has been supported in the literature, and various issues have been raised regarding the rigidity versus the flexibility of categories. Medin and Wattenmaker (1989) question the basis for making judgments of similarity. Barsalou (1983) points out that some categories are traditional and learned, while others are

evoked for very specific purposes (such as "things to take out of the house in a fire"). Murphy and Wright (1984) examine differences between experienced professionals and beginners.

There are consistent patterns in the ways people categorize everyday objects. For example, when shown a particular type of chair, such as a high-backed rocker or a captain's chair, most people call it a chair. They do not call it by its more specialized name, nor do they refer to it as a piece of furniture or a thing (Rosch, 1978). Similarly, a Vermont sugar maple or Douglas fir are more likely to be called a tree than their more specific label or than the more generic label plant. Rosch called this most common and useful kind of label the basic level. Moreover, she found that subjects could respond to these labels more rapidly than to specific or generic labels. The basic level term (BLT) for musical style is one of the measures used in this study. The present study seeks to establish a BLT by observing which term--if any-- is used most frequently.

Categorization and Music

There is a limited body of research on the categorization of music. This literature, unfortunately, does not make full use of or reference to another even

greater body of research from the cognitive sciences. These new and sometimes surprising results must be reckoned with if musical research is to stay current and meaningful. They are an important factor in the theoretical assumptions of this study.

The present study differs from previous musical category studies in several important ways. First, it uses the factors of expertise and context as variables, with the prediction of interaction between them. Second, it avoids the issue of defining features. The research questions have to do with the kinds of labels subjects use, not what specific features they attend to. The results might indeed suggest certain features, but to focus on them would run counter to the literature on which this study is based.

Philosophical Considerations

It must be stressed that this study is intended to provide useful information for music teachers. While it employs the terminology and methodology of cognitive science, it is not necessarily intended to break new ground in these fields. Rather, it should take the ground provided by nearly three decades of empirical research and build upon it.

In this light, it is fitting to keep in mind that all of the questions considered above are still being asked. The cognitive literature is infected with the viruses of oversimplification and misunderstanding. For example, a number of reports state such misconceptions as "concepts are coded into memory as prototypes" (Jonassen, 1988, p. 2). Rosch herself insisted that there was no such thing as a prototype; rather, one should speak of the "prototypicality" of some exemplars.

Psychology and linguistics give us insights into the mind, and how the mind deals with information. But none of this information is direct or literal: Objective data can be little more than a basis for inference about actual mental process. The issue of links between the real world and our conception of it is the concern of such monumental philosophers as Plato (1991), Aristotle, (1991), Kant (1965), Wittgenstein (1958), and Lao Tse (1972).

Verbal Knowledge and Musical Knowledge

Even if the focus is concentrated on the meaning of such words as music, kind, and style, epistemological arguments arise that cannot be fully answered here. In fact, "Our facility as language users prevents us from recognizing the complexity of the representations and processes that underlie our mastery of word meanings . . .

There may therefore be some truth in the . . . claim that a complete theory of meaning is impossible" (Garnham, 1985, p. 132). In view of these seemingly unanswerable questions, can a study such as this be useful, or even justifiable?

Three sources provide an affirmative answer. First, Booth and Cutietta (1991) suggest that memories about music are associated with verbal category labels. "The results also suggest that, in subject's minds, these verbal stylistic labels may override musical attributes" (p. 130). In other words, subjects perceived two examples as being in the same category even though certain musical elements were quite different.

Second, the First National Assessment of Musical Progress (1974) and the Second Assessment of Music (1980) use verbal responses--indeed, style category labels--as a measure of musical knowledge. In a format similar to that of the present study, the Second Assessment asks students "Are there any kinds of music that you like to listen to?" The range and specificity of the responses are an important indicator of musical knowledge.

Third, renowned authorities in the field remind their colleagues of a fact so obvious it is often overlooked: "Apart from making music or listening to it, talking about it becomes the next best means for acquiring a musical

education. Talking about music in conjunction with making it or listening to it provides the most potentially powerful educative process" (Tait & Haack, 1984--Crane Symposium p. 54) Bennett Reimer puts it even more succinctly: "In order to teach music, you have to talk about it" (Reimer, 1994).

Clearly, when teachers and students talk to each other (about music or any subject) they use words, and these words do have meanings, legion and ephemeral as they might be. The critical points for teachers are that labels do not necessarily represent a discretely defined class of music, and that understanding is not necessarily demonstrated by the use of a correct label. In fact, a teacher and a student using the same descriptive label might mean very different things:

> In principle, the choice of what to include and what to exclude in descriptions depends on one's pragmatic framing, i.e. on one's judgments of saliency, relevance, importance. None of the judgments are "objective"; nor can they be arrived at deductively or inductively; they are a matter of point of view; of context. (Givon, 1989, p. 89)

Definitions of Terms

This study is an attempt to treat several broad concepts as scientific variables. These include musical style, expertise in music history, and musical context.

It is essential to clarify how these variables will be defined and measured. For example, if one subject labels an example "Classical," another labels it "piano music," and a third labels it "early 19th century chamber music," how will these answers be classified and tabulated?

Style Terms

The details of measurement and statistical analysis will be discussed in Chapters 3 and 4, but a brief explanation is in order to establish the terminology and direction of the study. "Classical," in the case above, would be classified as a style term (ST). It is an idiom that can be associated with a specific historical period (17th-19th centuries) within a particular geographical region (Europe). It is not confined to these, but is associated with them; it can be performed by different people on different instruments, and has its own collection of substyles. In this sense, Classical is a generic rather than specific style label.

"Piano music," on the other hand, is not a style term, but a performing medium. Ludwig van Beethoven, Scott Joplin, and Jelly Roll Morton all wrote piano music in very different idioms or styles. Similarly, a choir (even *a cappella*) might sing gospel, doo-wop, or 16th century polyphony, so "choir music" would <u>not</u> be a style

term. References to instruments or voices are classified in this study as "performing media" terms. Other response types include national origin (French, German), and genre (opera, symphony, song).

The third label, "early 19th century chamber music," refers to a far more specific style period, and specifies a genre (chamber) as well. This response, like "Classical," is a style term (ST), but suggests attention at a more detailed level. It draws finer distinctions between the example in question and another example. In the data analysis, these answers will be tabulated separately.

For purposes of reliability and validity, the researcher and two faculty members with music history backgrounds independently examined the total response pool. Each term was judged as a Style Term, Genre, Performing Media, National Origin, or Other.

Basic Level Terms

It has been noted that examples might be labeled at different levels of specificity. A piece of furniture on which people sit is usually referred to as a "chair." In normal conversation, a person would never say "Sit over there on that piece of furniture," even though this higher level term would be correct English. On occasion, a

person might use a more specific label such as "armchair" or "high-backed rocker." The label used most easily and most often is referred to as the basic level term (BLT). Numerous studies have shown that people can answer yes or no to BLTs more rapidly than to more specific or more general labels (Rosch, 1978).

This phenomenon, logically, should apply to music as well as other domains. For instance, one subject might call a particular passage "New Orleans ragtime" music while another calls it jazz. "New Orleans ragtime" is more specific than "Jazz;" it is indeed a subcategory of Jazz. Studies have shown that experts tend to use labels at more specific levels than novices (Tanaka & Taylor 1991). Thus jazz might be considered a basic level term (BLT), whereas New Orleans ragtime is a more subordinate level term.

The target examples for this study are from the Norton Anthology of Western Music, a collection intended for use in college music history and appreciation courses. They represent a broad spectrum of styles, genres, and performing media, but for the most part would probably be covered under the generic term Classical. Novice subjects are unlikely to have the same specialized vocabulary as the experts. They nevertheless have many options available, such as commonly used genre terms (opera,

symphony), identifying instruments, or guessing at the nationality. One of the hypotheses on which this study is based states that people will use style related terms more often than other kinds of labels, and that the term used most often by the novices will be "classical."

Shared Equivalent Terms

Finally, this study will examine category stability. Stability here refers to the extent to which subjects use the same label for a given example. Some of the literature suggests that the categories of experts will be more stable than those of the novices (Brooks, 1989). Murphy and Wright (1984) found that experts as a group used fewer category labels than novices. In other words, the experts used the same labels more often. The expectation here is that the experts as a group will agree with each other more often than novices, even when the context changes.

To capture this effect, a count will be made of any terms used for a given musical example by more than one subject. These will be called shared equivalent terms (SETs) because some terms might be equivalent, such as "chorus" and "choir," or "religious music" and "sacred music." Such terms, while there might be subtle

distinctions, would be counted along with exact duplicates.

Design and Methodology

Subjects listened to a tape of forty excerpts, twenty seconds each, of varied musical selections. After each excerpt, they wrote the answer to the simple question: What kind of music is this? In studies by Rosch (1978) and Tanaka & Taylor, 1991, subjects were presented with pictures or visual cues and asked to respond. The brief excerpts are the musical equivalent. The music will play for about 20 seconds; there will be a five-second period between examples during which the number of the next example is given. The pilot study demonstrated that this was enough time to listen, decide, and write, but still avoid a lengthy analysis and consideration of secondary or alternative labels. First impressions were desired; in fact, only the first word written was considered as data.

Examples were presented in two different contexts. Context 1 (CTXT1) was exclusively tracks from the Norton Anthology of Western Music (NAWM) (1988), a CD collection of art music from antiquity to the twentieth century. CTXT2 was twelve of the same NAWM tracks--the target cases--mixed in with selections from rock, blues, and

jazz. The data consist of responses to the twelve examples that occur in both contexts.

One part of the subject sample represents members of the general public; while any given individual might be quite knowledgeable about a particular favorite style or genre, this expertise should be randomly distributed throughout the population. In effect, this class of subjects can be considered novices. These subjects were upper division or graduate non-music majors at the University of Oregon.

The other part of the sample includes advanced music students. They are considered experts in the field of serious music from the Western art tradition. Again, individuals might very well be experts on country music or jazz, but as a group the common knowledge base is most likely to be that which is part of their advanced study.

Experimental Design

Independent Variables

The independent variables are expertise and context. Expertise has two levels, expert (advanced music student) and novice (non-music major or faculty). Context also has two levels. In context 1 (CTXT1) all of the examples are from the Norton Anthology CDs, and exemplify Renaissance,

Baroque, and Romantic styles. Classical (18th Century) style is not included because use of the generic label "Classical music" was predicted to serve as a basic level term (BLT). The intent was to avoid having to determine whether a subject was using the term in its generic ("Classical music") sense, or in its specific ("18th Century Classicism") sense.

Dependent Variables

The dependent variable is the response pool of 768 words; only the first word written by each subject in each cell was used. In one sense, the response pool is the one and only dependent variable. But the three research questions focus on different aspects of the response pool. Moreover, each of the three questions focuses on a different portion of the response pool.

Research question 1 concerns all 768 words, regardless of group or context, asking whether a significant portion of them will be style terms. Question 2 concerns only those words deemed style terms, asking how specific or generic they are within each cell. Question 3 concerns all the words--style or otherwise--used by members of each cell, asking how often members agreed with their peers. Thus it will be more convenient

to treat the three different concerns as three different variables.

In this view, the dependent variables are the kinds of labels used by subjects to describe 12 target examples from the Norton Anthology of Western Music. These labels were first classified as style, genre, performing media, national origin, and other (based on the pilot study and on the actual response pool of this study). Question 1 asks whether the style terms (STs) will predominate.

Question 2 asks whether expertise and context will have an influence on the use of a generic label, the basic level term (BLT). Although this is a separate question from the first, it draws on question 1 in that a basic level term for style would appear as the style term most frequently used by the novices (because they represent the general population). If there are no effects, this BLT will appear just as often in the expert response pool, and in both contexts.

Next, shared equivalent terms (SETs) were tallied within each of the four cells (experts and novices in both contexts). Question 3 asks whether experts will use more SETs than novices, and whether context will influence that difference. SETs were any term or its exact equivalent used by more than one subject to label a given example. Thus if two novices in CTXT2 called example number 3

"opera," it was considered a SET. If one subject used the term "choir" and another used the term "choral," it was also considered a SET because the two words have almost exactly the same meaning.

Analysis

Research questions 1A and 1B are a matter of frequency count. Specifically, does the Novice group use one kind of label--Style Terms--significantly more often than other kinds (#1A); and, among these STs, is one term ("Classical") used more than any other? A Chi Square tested for significance in these questions.

Research questions 2 and 3, which concern the influences of expertise and context, call for a 2-by-2 Analysis of Variance (ANOVA). First and foremost, ANOVA captures any interaction between expertise and context, and that is a central issue in this study. Second, it compares the variance within each group, and an important assumption is that the experts will vary in their responses less than the novices. Finally, ANOVA compares group means; in this case, those will be the mean number of times each subject in each cell gives uses a BLT or an SET.

Variables and Measurement

Style Terms (STs) are any labels or responses that specify a historical period, such as the Middle Ages, 16th Century, or Late Baroque. Although the target cases are all Norton Anthology Renaissance, Baroque, and Romantic examples, if a subject uses a stylistic idiom label such as Folk or Jazz, these would also be considered STs.

The Basic Level Term (BLT) for style represents a generic as opposed to a more specific term. For example, record stores typically have a section labeled "Classical," which might include music from the Renaissance to the 20th Century. Classical would be a generic or basic level term for all these style period idioms. It should be noted here that a considerable body of research exists on the question of category levels. This research is examined in Chapter 2.

Shared or Equivalent Terms (SETs) are any labels used by different subjects for the same example. This would include obvious duplicates and equivalent terms such as 16th Century and Renaissance, or chorus and choir. A panel of three experts judged the equivalence of any terms that were not obvious duplicates.

Context (CTXT) refers to the setting in which categorization takes place. Half of the subjects will

hear an all-Norton tape; that is, they along with 12 target examples they will hear other Renaissance, Baroque, and Romantic music (CTXT1). The other half will hear the same Norton target examples, but mixed in with examples of Rock, Blues, and Jazz (CTXT2). Studies will be examined that suggest the different contexts will lead to different kinds of labels being used.

Expertise refers to the group to which a subject belongs; there are two such groups: Experts are upper division and graduate music students, novices are equivalent non-majors. Expertise and context are nominal variables, both with two levels (CTXT1, CTXT2; expert, novice). This calls for a Two-way ANOVA, with possible interaction between the two factors.

Conclusions

Musical categories have characteristics similar to those in other fields. They are not the rigid, exclusive structures often found in textbooks and on tests. They are flexible and dynamic, but within the constraints of actual features and consistent judgments of similarity. Barsalou (1989) argues that category structures depend very heavily on context, almost to the point of total instability. Other research (Brooks, 1989; Diekhoff, 1985) suggests that practice, understanding, and expertise

will lend stability to category structures. This study measures the effects of context on categorizing musical styles, and it measures this influence at different levels of expertise.

What kind of music did Mozart write?

The answer can be found in textbooks at various levels. A teacher can tell a student the answer to this question, and demand it back on a test. But the accepted label might not mean the same thing to both people, and they might actually perceive different characteristics in the music. Categorization always occurs in some context, and the same context might create different cognitive references for the teacher and student. If we give serious consideration to this phenomenon, our teaching will more effective, and the music we play will be better understood.

CHAPTER II

LITERATURE REVIEW

The literature germane to this study falls under three general headings. The first consists of research into categorization as a cognitive and linguistic process. This body of literature grew out of early studies concerned with the way people assign labels, classify objects, and perceive similarities and differences. It established a view of categories fundamentally different from that established by Plato.

The second consists of expert-novice studies in categorization. A number of studies have focused on the way naive beginners and skilled specialists categorize items in a given domain. These studies have focused on how fast subjects respond, how specific their labels are, and how they treat atypical examples.

The third consists of studies that use music as the domain in which categories are formed. These will be examined for insights and results that might have an impact on the present study, but, for the most part, they will be critiqued in terms of the evidence from the two

previous sections. A superficial survey of music categorization literature revealed some deficiencies in connection with cognitive studies in other domains. In general, the music studies did not draw on the cognitive material, and the design and methodology was sometimes inconsistent with that of cognitive research and its findings.

The Classical View of Categories

Over two thousand years ago, Plato (Ed. and trans. 1991) established a view of categories that still greatly influences category thinking today. A category was defined by a set of characteristics: All of the members had those characteristics, and any item that did not was a non-member.

In Phaedo, Plato (speaking as Socrates) points out that tallness and shortness are mutually exclusive; "nor can any other opposite . . . simultaneously become or be its own opposite" (1991, p.188). He draws similar conclusions regarding hot and cold, even and odd. While this might appear a simple explanation of the meaning of opposites, it sets up a rather rigid model of categories in that a given object in one category cannot be in the other. It further implies that a person labeled as tall

cannot be also labeled short; this person must be in one category or the other.

Plato also discusses the concept of a characteristic that runs through a category (such as "large things") giving it definition. "When some plurality of things seems to you to be large, there perhaps seems to be . . . one characteristic that is the same when you look over them all" (1991, p. 264). This has been taken by thinkers though the ages to be the very basis for the existence of a category: one or more defining characteristics shared by all the members. "Both rationalists and empiricists, from Plato and Aristotle down, have subscribed to the view that mental categories are discrete and absolute" (Givon, 1989, p. 36).

Modern Applications of the Classical View

This established an all-or-none, rule-based approach that still appears in educational literature (Merrill & Tennyson, 1977) and in concept formation studies (Bruner, 1973). It assumes that categories function as little more than a bundle of specific and identifiable features.

In this view, a given musical example will have certain discernible attributes that are perceived by the listener. When listeners are asked what kind of music a particular example is, their response would be based on

those characteristics as well as their own musical knowledge. If the response involves style (as this study hypothesizes it will), then the style label used will be a function of such attributes as well. A review of literature will give evidence that the situation is not so clear-cut, and that musical style, along with other category labels, is not so easily defined.

Cognitive Categories: A Modern Alternative

Modern research in category theory began with Eleanor Rosch. She was the first to vigorously challenge the Platonic model of discrete, feature-defined categories. According to Gardner, she is "possibly the cognitivist whose critique did most to undermine the classical view" (1985, p. 342). Her work and that of her colleagues established the issues, methodologies, and theoretical basis for nearly all subsequent category research. Rosch (1973) pointed out that in the laboratory, researchers could arbitrarily manipulate features. Large blue circles were just as probable as small red squares, and a subject entered with little notion about what to expect. In the real world, however, one did not often encounter animals that had wings, feathers, and three legs. Real world categories were not arbitrary like those of Bruner: Wings

and feathers were always matched with two legs. Outside the laboratory, people formed categories that were rich in inference and based on fairly consistent prior experience: They had never seen an animal with feathers and three legs; they knew from personal experience that some features were more likely to appear in conjunction with others.

Thus, with Rosch, began a new body of category literature which focused on the conceptions and subjective experiences of the observer rather than objective features of the observed. Attention shifted away from objects in the environment and toward the judgments and experience of the observer. Category features were not abandoned, but put in the context of human cognition.

Category Levels

Rosch found patterns in the response times of people to computerized categorization tasks. When a general cue (tree, chair) was flashed on a screen, it was easier for subjects to decide whether the exemplar that followed was in or out (YES/NO) of a category. If the cue was too specific (black locust tree, captain's chair) it took longer for the YES-NO decision. Moreover, when subjects were shown common objects such as plants, animals, and furniture, they would most often respond at a certain

level. Rarely would the more general term (animal, piece of furniture) be given; usually a more specific response (dog, table) was given.

The basic level was that at which most subjects responded: "maple tree" for example, rather than "Norway maple" or "plant." This level elicited the fastest response times (RT) in a high-speed decision task. It was also the term that children learned first (Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976). The authors pointed to the cognitive economy of this phenomenon: The basic level term was most useful in immediately discriminating and classifying the environment. In terms of recognizing objects that had similar properties, it was more useful to say "a chair" than "a piece of furniture."

Rosch and her associates (1976) recognized that categories reflected real-world structures. They did not immediately reject the objective feature lists posited by Plato, but they did suggest that human cognition was also an important factor. Category levels were not just an effect of an objective environment, but of human thought and behavior.

Because of their hierarchical nature, these levels were referred to as the vertical dimension in category structure. Rosch called broad, general labels (plant; animal) superordinate; very specific labels (black locust,

red breasted robin) she called subordinate; and the more central, accessible, and frequently used labels she called basic (1978). "Object names at the basic level of abstraction should be the names by which objects are most generally designated by adult speakers of the language" (1976, p. 422).

Categorization of Natural Objects

Rosch and her associates, in 12 separate experiments, established methods that became the mainstay of category research, such as feature lists and response time. The first 4 experiments served to support the concept of basic level objects in a linguistic sense. It examined the words people used most often to describe or label familiar objects. The second set of 8 explored the cognitive, perceptual, and linguistic implications. Here, subjects formed images, matched pictures, and named objects. All subsequent category literature refers to this study, so it deserves some scrutiny here.

In experiment 1, subjects were given an object name and asked to list attributes of that object. The objects used came from taxonomies such as "musical instrument, guitar, folk/classical guitar;" and "tool, hammer, ball-peen/claw hammer" (Rosch, 1976, p. 388). The attributes listed were then judged for truth by a separate set of

subjects as a reliability check. The number of attributes listed at the basic level was higher than at the superordinate ($t = 6.43$, $p < .001$). Additional attributes were listed at the more specific subordinate level, as might be expected: Having described an apple, a subject would need additional terms to describe a particular kind of apple. But the size of increase at this level was not as great as the increase going from superordinate to basic ($t = 6.43$, $p < .001$). In other words, more new descriptive terms were introduced at the basic level than at other levels in the vertical structure.

The basic level was thus established as the one at which most information was immediately available. The categories were general objects such as fruits and tools, and the subjects were students in psychology seminars. There was no consideration of special knowledge, only general experience and common language. The subjects were more or less equal, and expertise was not an issue in these early studies. Experiments 2, 3, and 4 involved similar procedures with physical activities, and outlines of common shapes. In each case, there was a basic level at which the maximum amount of information seemed to be available.

The second part of the study explored the psychological effects of category levels. For example, subjects were asked to visualize an image of furniture. It was found that they would do so by imaging a specific, basic level exemplar such as a chair. There seemed to be no image that represented the higher order category term. Another experiment had subjects deciding if two pictures were the same, similar, or different. Before seeing the pictures, some subjects were given a superordinate level cue (tool, vegetable) and others were given a basic level cue (hammer, carrot). RT with a basic level cue was faster ($p < .05$).

Still another experiment compared a sorting and classifying task by preschoolers, elementary school students, fifth graders, and adults. The basic level effect was dramatic. "At all age levels, basic level sorts were virtually perfect" (1976, p. 417).

Basic Level as Frequency

In Rosch et al. experiment number 10, a picture naming task, "subjects overwhelmingly used the basic level name in this free-naming situation" (p. 423). This supported Rosch's assumption (and that of the present study) that "Object names at the basic level of abstraction should be the names by which objects are most

generally designated by adult speakers of the language" (p. 422).

Again and again, it appeared that category level and vertical structure was a given constraint on human information processing. Rosch et al. summarized the results of their landmark study:

> The categorizations that humans make of the concrete world are not arbitrary, but rather are highly determined. They are determined in the first place, because the perceived world is not an unstructured total set of . . . attributes. . . . categories are determined, in the second place, because, in so far as categorization occurs to reduce the infinite differences between stimuli to . . . useful proportions, the basic category cuts should be those which yield the most information for the least cognitive load. (1976, p. 428)

It might appear circuitous to say that subjects used the BLT most often, and then define the BLT as the term used most often. The present study, however, differs from those of Rosch in two ways. First, Rosch also tested for the basic level by response time (RT). This methodology calls for measurements in milliseconds and is inappropriate for someone listening to several seconds of music. The frequent use of these terms was an additional aspect she noted. Second, the present study predicts that the BLT will be a particular label: "Classical." That is, it predicts that this very label will be used most frequently by the novices.

Category Boundaries and Linear Separability

As pointed out above, Plato (1991) established a model of category structure that included strict boundaries between the categories. This boundary was determined by the presence or absence of critical, defining features. Eleanor Rosch (1978) challenged this conceptualization, arguing that categories had a horizontal as well as a vertical structure.

It creates some intuitive problems to state that there are no boundaries to a category. The term category is used regularly to describe a sharply defined group of ideas, objects, or people. Research designs, for example, call for mutual exclusivity in their nominal categories. Children learn their first words as mutually exclusive category names (Markman, 1989). A young child is puzzled when a "doll" is referred to as a "toy." Medin and Wattenmaker (1989) refer to this quality as linear separability, because it implies that a line can be drawn to separate one category from another.

Contrary to this notion of discrete divisions, Rosch and her associates (1976) found a graded structure in the categories they investigated. The graded structure of categories refers to the phenomenon of some members being more typical of a category than others. A robin was found

to be a prime example or prototype of the category birds. Penguins, chickens, and ostriches were certainly birds, but subjects rated them as less typical. These atypical exemplars occupied a peripheral position in the category, while robins, sparrows, and canaries occupied a prototypical or central position.

Thus the horizontal dimension of a category resembles a bell-shaped curve, with maximum typicality in the center, and diminishing typicality as one moves away from the center. Eventually, some of the more atypical exemplars in one category became potential exemplars in a related category. In other words, there is no linear separability, but a gradual shift from one category to the next.

Rosch and her colleagues (1976) suggested that only some categories are continuous while others are discrete. But subsequent studies have found graded effects in many different kinds of categories. Barsalou (1983) found them in what he termed ad hoc goal driven categories such as "things to take out of the house in a fire" and "ways to avoid getting killed by the Mafia." Subjects compared such options as "change your identity and move to the mountains of South America" and "become a drunk in Detroit" (p. 215). Barsalou used a Kendall's coefficient to test reliability among different subjects ranking the

choices in typicality. He found an average agreement of .87, indicating that subjects agreed on their ratings of the different options. Even these ad hoc categories had graded characteristics.

The Graded Structure of Categories

Armstrong, Gleitman, and Gleitman (1983) found that people rate the number 9 as more typically odd than 259, even though such a comparison contradicts the definitional basis for the category of odd numbers. The graded structure of categories seems to be an effect of human cognition rather than a real world phenomenon. As Plato would have pointed out, all odd numbers share a common trait (indivisibility by 2) and thus all are equally odd. In the perception of humans, however, some numbers seem to be more odd than others, or at least better examples of oddness.

Graded structures do not mandate overlapping categories. Odd numbers are clearly linearly separable from even numbers, and there can be no member in both categories, no matter how atypical it might be judged. But if a number is rated as "less odd," there is some implication that it might be at the outer edge of the category, making it at least a little closer to the other category, "even." In this light, 259 might be construed

as "more even" than the prototypically odd 9. The two categories, though discrete by definition, seem to have a somewhat graduated boundary.

Murphy and Wright (1984) found actual overlap in category features used by clinical psychologists. For example, children classified as "depressed" shared the feature "feels sad" with children classified as "aggressive" (p. 152). In a sense, a child exhibiting shared features might be marginally categorized as either depressed or aggressive, depending on other factors. The line separating the two becomes blurred.

Barsalou characterized the graded structure as a "continuum of category representativeness, beginning with the most typical members of a category and continuing through its atypical members to those nonmembers least similar to category members" (1989, p. 102). He goes on to describe category continuity as an effect of long term memory storage of concepts. "Knowledge in long-term memory from which concepts are constructed for a particular category may generally not have clear boundaries. Instead, knowledge for a particular category may share much structure with knowledge for other categories" (p. 121).

While subjects are able to classify items in different categories, the imaginary lines between those

categories become blurred when the memories of many individuals are called into action. Barsalou points out that this is an observed, behavioral phenomenon that appears when people actively categorize things. He denies any claim of it being a "cognitive structure," and attributes it to an effect of the memory base of different individuals. Categories might have linear separability, he suggests, but human memory does not.

Prototypes

Prototypes are a somewhat controversial concept in cognitive research. Rosch, for example, insisted that a prototype was not a "mental trace" or a single exemplar (1978). She suggested that it was better to write of the "prototypicality" of some exemplars. Hintzman (1986) argues that prototypes do not exist in memory, but are formed in recall. Other writers, ignoring Rosch, went so far as to claim that concepts were stored in memory as prototypes (Jonassen, 1988).

Nevertheless, prototypes are a useful construct when considering graded and continuous categories. Whether they are convenient abstractions or real-world prime examples, they represent a central tendency in each category. That is, they might be thought of as a somewhat

"ill defined center" that complements the "ill defined boundaries."

Clearly, there are some problems with the issue of discrete versus continuous categories. Lakoff (1989) agrees that to a mathematician, odd numbers are a discrete, separable category with clear limits. But the study that showed prototypical effects in odd numbers (Armstrong, Gleitman, & Gleitman, 1983) did not distinguish between numbers as mathematical entities and numbers as a linguistic phenomenon. That is, to most people, numbers are words that stand for the mathematical entities. As such, they are subject to the same linguistic effects as Rosch's basic objects and natural categories. These effects include fuzzy boundaries and continuity; they do not include linear separability. It remains to be seen if recorded musical excerpts are subject to the same effects.

Hintzman (1986) has pointed out that a schema-like effect or prototype can be produced by a system using nothing but discrete memory traces of prior cases. His MINERVA II artificial intelligence model stores memories of specialized shapes; the computer can then produce a composite image that appears to be an "averaged" prototype of the whole set of shapes. The prototype is never

actually stored in memory, according to Hintzman, but is a secondary effect of retrieval.

This effect, however, does nothing to discredit Rosch. It must be remembered that, in her later writings, she cautioned against thinking of prototypes as actual images or neurological traces (1978). Some real-world cases seem to exhibit a higher degree of prototypicality than others, but it is wrong, according to Rosch, to speak of any single case as an actual prototype.

How Humans Categorize

Hintzman's MINERVA II draws on a memory base of all available cases. A less structured human equivalent might be the use of a few specific familiar cases as a guide to classifying new exemplars. Rather than comparing the new case to a cognitive model, people sometimes simply compare it with a well-known exemplar. This model, as proposed by Brooks (1989), is somewhat closer to the classical category view, where new cases are compared to a feature list. The difference, however, is that no actual list need be articulated, and no specific features need be consciously identified.

Expertise and Categorization

In their classic study, Rosch and her associates (1976) encountered one subject, an airplane mechanic, who produced a lengthy list of attributes for the category airplane. They characterized the effects of his expertise as the ability to "make use of attributes that are ignored by the average person" (p. 430).

There were, however, other, more subtle differences, between the airplane expert's categories and those of the other subjects. For example, most subjects used a top and side view when imagining parts of an airplane, but the mechanic used the underside and engine. He was nevertheless able, they wrote, to "take the role of the average person and list attributes common to all airplanes, and could imagine an average airplane shape from the outside" (p. 430). He could, depending on the context, consciously function as a novice or as an expert. This might be analogous to a teacher (expert) discussing music with a music class (novices).

This study, it should be remembered, represented a turning point in category theory. Rather than Platonic feature lists, Rosch and her associates argued that "the basic category cuts should be those which yield the most information for the least cognitive load" (1976, p. 428).

It could be argued that, with greater in-depth understanding of a subject, it is easier to make those cuts at more specific levels. In other words, an expert would automatically label or categorize the same items at a more specific level than a novice. With a change in context, however, the expert might revert to the basic level as did Rosch's airplane mechanic.

Expert-Novice Categorization Research

Expertise is indeed a well-researched topic in cognitive science, but is usually associated with decision-making and problem-solving, not categorization (Anderson, 1990). While there is a wealth of empirical literature on categorization, only a few studies consider expertise as a factor. Those few do suggest noticeable differences in categorization by experts when contrasted with non-experts or novices.

Chi, Feltovich, and Glaser (1981) found that experts in physics approached problems differently from novice students. This was a problem-solving study, but categorization of the problems was stressed. When asked to organize the various problems into related categories, experts put most problems in three large inclusive categories. Novices, on the other hand, had only one large category and many smaller, more exclusive

categories. This might imply that the expert categories were wider in breadth and more inclusive. Novices apparently saw each problem as unique.

Furthermore, novices used superficial, literal details as a basis for categorization, such as the presence of an inclined plane or pulley. Experts, in contrast, used general principles such as Newton's Third Law as a basis for classifying. Attending to the surface information, novices stressed the differences between problems rather than the underlying similarities. Experts saw more similarities overall, and therefore generated fewer categories, each of which included more cases. This, incidentally, should lead to more agreement among experts, more disagreement among novices.

The authors interviewed their subjects and drew protocols of their procedures. That is, they wrote out step-by-step statements that followed the thoughts of the subjects. Analysis of the protocols suggested a structured knowledge set or <u>schema</u> on which the subjects drew to approach the problem.

Both experts and novices included the general principle of <u>conservation of energy</u>, but at different levels. When the researchers graphically diagrammed the schemas, conservation of energy appears in the lower middle section of the novice diagram, and at the very top

of the expert diagram. The novices paid initial attention to specific details, while the experts gave initial attention to overall principles.

Schemas and Expertise

Schemas[1], according to Anderson (1990) "facilitate making inferences about the concepts. If we know something is a house, we can use the schema definition to infer that it is probably made of wood or brick, and that it has walls, windows, and the like" (p. 135). A schema is somewhat like the default settings on a computer, and like these settings, can be easily changed if a situation calls for different settings. Chi, Feltovich, and Glaser (1981) found that physics experts were able to construct useful, accurate schemas from previous experience and knowledge, while the novices went on a case by case basis. It was also found that the experts included a possible solution in their schemas, while the novices did not.

In contrast, different results were found in a study that compared expert probation officers to a non-expert panel (Lurigio, 1983). In this study, subjects classified probation cases from a description of each case. As in Chi, Feltovich, and Glaser (1981) the schemas used by each subject were compared, but in this case, experts generated

---

[1] Sometimes pluralized as schemata

more category schemas than novices (t [38] = 2.08, $p < .05$). While this is inconsistent with the earlier studies, other aspects were consistent. Novice schemas were "simple and impoverished" (p. 139) while experts used a rich, interrelated network of traits and strategies. The experts were described as "not just stringing together facts' (Lurigio, 1983, p. 139). Like the physics experts, they seemed to use information beyond that which was given.

In a card sorting task comparing expert and novice probation officers, Lurigio's experts had a clear conceptualization before placing cards. That is, they appeared to follow a plan that included groups of cases, while the novices seemed to follow a case-by-case pattern. In fact, experts took more time to sort the cards, once again in conflict with the earlier studies (Chi, Feltovich and Glaser, 1981; Rosch et al., 1976). Lurigio interpreted this to mean that the experts were invoking complex schemas rather than making isolated decisions. These response times, it should be noted, were measured in minutes, and they involved a conscious, elaborate procedure. They are not to be confused with the high speed YES-NO RTs of Rosch and her associates.

Category Levels and Expertise

Rosch and her associates (1976) found what they termed a vertical dimension in their category structures. At the core of this structure was the basic level, for example "tree" as opposed to a specific variety or a more general taxonomic "plant." This phenomenon was more than just an average, popular name that most subjects used. It was the level of cue that led to the fastest RTs, the label that generated the most descriptive terms or features, and the level at which children and adults were best able to sort cards representing different familiar categories.

It appeared that the basic level was not just a superficial effect among familiar objects. Rosch and her associates suggested that it might reflect actual structures in the environment as well as a convenient and economic aspect of cognition.

> "Basic objects for an individual, subculture, or culture must result from an interaction between the potential structure provided by the world and the particular emphases and state of knowledge of the people who are doing the categorizing. However, the environment places constraints on categorizations. Human knowledge cannot provide correlational structure where there is none. Humans can only ignore or exaggerate correlational structures." (1976, p.430)

Tanaka and Taylor (1991) argue that structure in the environment does not necessarily determine the basic level, but that human perception, interaction, and culture have a substantial influence. They point to children of the Tzeltal Mayan tribe, who learn the equivalent of our subordinate (more specific) labels early in life. These subordinate labels help the children discriminate plants as important sources of food. Only later do they learn the so-called basic level. In effect, their basic level has been shifted downward for adaptive purposes.

Tanaka and Taylor (1991) replicated some of the methods used in Rosch et al. (1976) but with two distinct groups of subjects. One group consisted of expert dog handlers and members of a local American Kennel Club chapter. The other group consisted of experienced bird watchers. The pictures and cues used in the categorization response time task were dogs, birds, and a few other assorted filler items. Thus the dog experts served as a non-expert control group for bird cases, and the bird experts served as novices in the dog cases. Since the present study will adopt similar methodology, a close examination of Tanaka and Taylor is in order.

Their method used a computer program similar to Rosch et al. A row of plus signs appeared on screen to focus attention; after one second this was replaced by a

category label. Two and a half seconds later, a picture appeared on a projection screen. If the picture and label matched, the subjects pressed true; if not, they pressed false. Subjects used the index finger of their dominant hand and were instructed to respond as fast as possible. They were given a two-minute rest midway through the experiment.

There were a total of 128 trials: eight dogs, eight birds, each shown (on a slide projector)16 times. Each slide was shown as a TRUE case one time at each level (superordinate, subordinate, basic) and once as a FALSE case at each level; also, 32 fillers (trees & rocks) were used to prevent automatic FALSE responses for non-dog and non-bird cases. Plant was used as a superordinate level FALSE case. That is, the subjects saw a plant cue and then a dog; this match was false, but at the superordinate level (plant, animal) rather than the basic level. There were plants, so subjects had to pay attention and consider the possibility of a TRUE case. Half of the fillers were TRUE, half FALSE.

Tanaka and Taylor analyzed their results using a 3 x 2 ANOVA with category level (superordinate, basic, subordinate) and knowledge domain (expert, non-expert) as independent variables. Interaction between category level and knowledge domain on TRUE trials was significant:

$F_{(2,46)} = 12.43$, $p < .001$. For FALSE, results were also significant: $F_{(2,46)} = 14.53$; $p < .001$. As predicted in their hypothesis, experts were as fast to categorize at the subordinate level as at the basic level. In non-expert domains (that is, dog experts categorizing birds and vice versa) RTs were faster at basic level.

Experts were faster to respond FALSE at the subordinate level. For example, bird experts were faster when given a "sparrow" cue and robin picture than when given a "dog" cue and robin picture. But outside of their domains of expertise, RT at the subordinate level was slower. Novices, that is, took longer to decide on more specific labels.

Four general differences were found between expert and novice categorization performance. In the expertise domain, subordinate level categories were as differentiated as basic level categories; that is, they were as richly described and as rapidly identified. Experts used subordinate level names as frequently as basic level names to identify objects. A dog expert, for example, was as likely to call a Doberman pinscher by its specific breed as to call it a dog. Expert RTs at the subordinate level were as fast as at the basic level. And, finally, "subordinate word primes produced greater facilitation in a physical matching task than basic level

word primes" (p. iv). In other words, confirming that two pictures matched was easier when a subordinate prime (rather than a basic level prime) was given before the two pictures.

Category Boundaries and Expertise

From the very beginning, Rosch argued against discrete Platonic categories. But if there were not strict boundaries between one category and the next, then how was one category recognizable from another? This study will not answer that question in any ultimate terms, but it might shed some light on how these elusive boundaries are effected by in-depth knowledge of a specialized field.

The Tanaka and Taylor study (1991) is concerned almost exclusively with levels of categorization and expertise; boundaries are not discussed in any detail. In fact, Tanaka and Taylor specifically used exemplars that were central to the category, exemplars that the experts would recognize with no confusion. Category boundaries and atypical exemplars were avoided.

Murphy and Wright (1984), on the other hand, were quite concerned with the breadth of each group's categories. They wanted to see if the outer limits of the category were different for the more experienced clinical

psychologists when compared to the trainees. Their test for this was the extent to which subjects used the same descriptors for two or more categories. Their categories included depressed, disorganized, and aggressive children, and their descriptive labels included such traits as "throws tantrums," "feels angry," and "feels sad."

A descriptor used in two different categories implied some overlap. Novices associated descriptors with one category, and therefore, not with another. Experts, in contrast, used the same term to describe cases in several categories. With experts, children classified as depressed were described as "feeling sad" (p. 152). But the same description was used for aggressive children. A child exhibiting this particular trait, therefore, might be classified as an instance of both categories. For experts, at least, the categories might be said to overlap.

Murphy and Wright (1984) had four groups ranging from novice through experienced student to expert. The greater part of their results, however, were significant ($p < .05$) only when the expert group was compared with the others. No significant differences ($p > .10$) were found between any two of the lower level pre-expert groups. There did not seem to be evidence of a gradual accumulation of

category changes, but an all-or-none shift at the highest level of expertise.

More important to this study were Murphy and Wright's results regarding within group agreement. In listing descriptive terms, the experts agreed with each other 41% of the time, while novices agreed with each other 22% of the time. In other words, given a clinical category, experts used shared equivalent terms to describe the category more often than novices, $F_{(3, 284)} = 30.9$, $p < .001$. According to Murphy and Wright, "There is clear evidence that . . . the level of interrater agreement increases with expertise" (p. 147). This leads to the expectation that music experts might agree more often than novices when applying descriptive labels to music.

Categorization Based on Intuitive Theories

In order to make a judgment of similarity, an individual must decide which properties to ignore. This decision, according to Medin and Wattenmaker (1989), is made on the basis of the individual's ideas about the nature of the category and about the relationships between various properties. These theories can be purely intuitive, naive, heuristic, even wrong. In any case, they are based on the individual's experience with members of the category.

The implications for the present study are clear. The novices will have a variety of intuitive theories about the musical category labels they generate based on divergent musical taste and experiences. The experts, on the other hand, will have at least somewhat more consistent theories based on their advanced study. "The need to share conceptual knowledge in communication . . . places constraints on concepts. . . People attempt to reach a common ground or consensus by tacitly agreeing on certain common values or dimensions for organizing the concepts involved" (Keil, 1989, p. 193).

Current Theoretical Issues

Medin and Barsalou (1990) argue that similarity to a prototype or to a known exemplar is not a satisfactory explanation of category structure. An element of salience has to be introduced: Some features of an object are ignored while others are deemed more important. Plums and lawnmowers, they point out, have certain qualities in common (both weigh less than a ton, both were not found on earth 100,000,000 years ago), but these qualities do not seem to constitute similarity. The unanswered question is what separates the important features from the unimportant? Once again, a simplistic feature-defined model is not sufficient.

Stylistic similarities between Mozart and Haydn might be quite obvious to the musical scholar, as are differences between these 18th Century Viennese Classicists and Vivaldi, an Italian Baroque figure. But to the novice, Vivaldi, Mozart, and Haydn might fit very neatly into a category of composers who used orchestra, wore powdered wigs, and whose music is played in concert halls. Can the music expert merely dismiss these common features as unimportant?

Developmental Influences on Categorization

Keil (1989a) points out that preschoolers' notions of a clearly defined category such as "uncle" are not those found in the dictionary. His young subjects, for example, stated that an uncle was an adult friend who was handsome and brought presents. Uncles could not be seven years old and could not be ugly. Only later do children conceive of an uncle as the brother of a parent. The situation, Keil seems to argue, is not merely a matter of accurate information, but a developmental change in category structure.

Keil (1989b) describes this maturation of category formation as a characteristic-to-defining shift, where the child first identifies an item based on salient characteristics (doggie: shaggy tail, four legs, friendly)

but later learns a more formal definition for what constitutes a dog, and what differentiates it from a cat. Social function has an impact in Keil's view: As the child grows and must communicate with others regarding different concepts, there is a tendency to use a more analytic definition. The attention, however, is still on the defining features of a category; the ability to perceive, understand, and communicate these features is the developmental issue.

Neisser (1989) argues that this characteristic-to-defining shift is also found in adults learning about new categories: "something similar occurs when we move from ignorance to expertise in any new domain. Novices have no choice but to judge by appearance, sticking as close to the basic level as they can" (p. 20). This once again implies significant differences between categorical responses by experts and novices. It clearly suggests, for example, that novices will use basic level terms more often than experts.

Stability of Expert Categorization

Barsalou (1989) argues against stability in category structures. Context is so important that the same subjects will respond differently on different occasions. Readers of the word "frogs" might not associate it with

the category "things eaten by humans;" they are more likely to do so, however, if it is read in the context of "French restaurant." Context can have a dramatic affect on categorization. This suggests that neither experts nor novices would show reliable patterns in classifying musical examples under different contexts.

Brooks (1989) finds some fault with Barsalou's methodology in justifying unstable category structures. He argues that Barsalou's *ad hoc* studies (things to take out in a fire, ways to avoid the Mob) do not give the subject a chance to practice classifying the items. Subjects are presented with an unusual and somewhat artificial task, not at all like real-life categorization procedures. Given time to theorize and organize, writes Brooks, the subjects might develop some stability in their response patterns. Groups of subjects given such an opportunity might develop within-group consistency.

Brooks does not use the term expertise, but it appears that he offers it as a foil to Barsalou's changing contexts. In fact, Barsalou admits that "subjects from the same population correlate around .50 with each other. . . although agreement is generally low, it exists and must be accounted for" (p. 123). Thus some consistency is found using general knowledge, general populations, and somewhat bizarre *ad hoc* categories. Might not even more

be found in the highly-organized categories of a music history expert?

Indeed, studies using knowledge organization as a test of understanding (Diekhoff, 1983) indicated that expert instructors were more consistent and reliable in their judgments of similarity than the novice students. Diekhoff suggests using these judgments of similarity as a basis for evaluation and grading. He found a high correlation between the judgments of his top students and a panel of experts.

Murphy and Wright (1984), in comparing category description among clinical psychology professionals, found clear expert/novice differences. The experts frequently used the same or similar terms to describe the category, while novices did not. With regard to the issues taken between Barsalou and Brooks, Murphy and Wright seem to support substantial category stability among experts, but not among novices. Again, this suggests music experts will agree more often than novices.

Areas of Disagreement

Homa, Rhoads, and Chambliss (1979) found that experts had tighter category clusters than novices. That is, members of a category were rated as more similar to one another and as more different from members of other

categories. This contradicts Murphy and Wright (1984), who found broader, more inclusive, and less distinctive categories with experts. The latter authors explain that the results are reconcilable because the methodologies were different; but further investigation is needed to settle the issue.

Subjects in Homa et al. (1979) rated the similarity of objects within categories, while those in Murphy and Wright used feature lists. The earlier study involved classification skill, and the focus was on distinctiveness, while the more recent study used a more open-ended procedure. It is easier to use a descriptive label in more than one category, if one feels so inclined, than it is to classify an object in two different categories.

In the study by Lurigio (1983), expert probation officers took more time than novices to categorize criminal suspects. Initially, this seems to contradict the hypothesis of faster RTs for experts. Lurigio suggests that the time delay was caused when experts called up rich schemas from memory, whereas novices made case-by-case decisions based on relatively short feature lists. Furthermore, in this study subjects used a conscious procedure and protocols, not immediate response

patterns. The times in Lurigio were measured in minutes, not milliseconds.

Categorization by Musical Style

The first thesis in this study is that the preferred domain of response to a recorded musical example will involve style. That is to say, subjects will call the various examples names such as "Jazz", "Rock," or "Classical" more often than "vocal," "piano," or "cheerful." The pilot study gave some evidence to support this position, but most of the evidence is anecdotal and informal. The same kind of evidence supports the second thesis, that music experts will use more specific terms in their basic level response, such as "Renaissance," "Baroque," and "Romantic."

This evidence, as discussed in Chapter I, includes the following informal observations: (a) When asked what kind of music they like, people usually answer with a style term such as country, rock, or classical; (b) record stores commonly organize their shelves under similar style labels; (c) radio stations announce their specialized programming in style terms; and (d) music history courses and texts are organized around such style terminology.

In addition, support for the categorization of music by style comes from Booth and Cutietta (1991). Subjects in this study listened to a tape of 12 songs which, according to the authors, fell into four style categories: Christmas songs, rock songs, musical show tunes, and children's songs. After two subsequent hearings, 48 hours apart, subjects tried to recall the titles. Some subjects recalled titles in the approximate order they originally heard them (serial recall), others seemed to group titles by style category, and a third group seemed to use no particular strategy. Of the three groups, those using a categorical strategy had the highest recall. As with the present study, the researchers did not give the subjects any cues for the categories; categorization, they maintain, was spontaneous.

Further evidence for spontaneous categorization by style comes from "subjects who, when they could not remember a specific song, wrote responses such as 'another Christmas song' or 'another Broadway song' (p. 130). In their conclusions, Booth and Cutietta called for "more research . . . that uses free recall as a tool to explore more thoroughly the relationship between verbal labels and musical recall, as well as the cognitive processes involved with both" (p. 130).

The present study, while it uses labeling rather than recall, is an attempt to provide such research. It follows the path of categorization by the subjects (rather than the researcher), and it draws on methodology normally associated with verbal rather than musical information. It is distinctly different, however, in its focus on category levels (generic versus specific), on expert-novice differences, and on the influence of musical context.

Pilot Study

In preparation for the expert/novice study, a pilot study was conducted in the summer of 1992 at the University of Oregon. In this study, subjects were all experts--that is, they were all graduate and upper division music students. As in the larger study, these subjects listened to brief excerpts of recorded music: half all Western art music, and half mixed with Rock, Blues and Jazz. The priming question was "What style of music is this?"

In the pilot, context had a significant effect on the kinds of labels used. That is, subjects who listened to a narrow context--all "Classical"--used style terms 36.25% of the time, while those who listened to an extended context--Rock, Blues, Jazz, and Classical--used STs 63.25%

of the time. This difference proved significant (F1, 14; p = .01).

The results clearly supported context as a source of variation in the kinds of labels used by subjects in the pilot study. These results, together with the work of Barsalou (1989), and of Booth and Cutietta (1991), are the basis for the hypothesized context effects in the larger study.

Surprisingly, even though the question was "What style of music is this?" subjects answered a total of 50.25% of the time with labels that were not, by the definition adopted for this study, style terms at all. Such labels included genre (opera, lieder), performing media (orchestra, choir), and national origin (Italian, German)[2]. Since the larger study gave no such cue for style ("What kind of music is this?") any preference for STs might well be attributed to the listener rather than the study design.

Arguments Against Musical Categories

Arguments are often heard that music cannot be categorized. Grout (1973), a standard text in many

---

[2] Some subjects also tried to identify (that is, to name) the work or the composer. This option occurred so infrequently in the main study that it is included in the "other" response category.

college music history courses, states that "Historical labels are a matter of convenience, and most of them are necessarily inaccurate" (p. 159). Robinson and Winold (1976), in a text designed for choral conducting students, adopt a similar view: "We have been stressing those aspects which make one period of choral literature different from other periods. It is, however, one of the sublime glories and mysteries of music that it ultimately resists assignment to any closed chronological or geographical category" (p. 498).

If one assumes the rigid Platonic category model, as these statements seem to, then categorization is indeed difficult if not impossible. But if the more intuitive category model is used, a strong case can be made for using generally accepted style terms. If categorization is such an automated process, the listener might be labeling the music from the first few notes. That indeed is the central phenomenon of this study.

Musical Style and Cognitive Research

While there appears to be little experimental literature treating musical style in terms of cognitive categorization, there are a number of studies that do use it as a variable. Many of these studies treat style as a nominal, discrete variable, defined in terms of historical

perspective: A given selection is or is not an example of a given style. The implication is that style is a matter of certain features, and students attend to these features. Cutietta (1993) proposed a set of defining features (motion, energy, flow, fabric) and tested their reliability. Eastlund (1992) used multidimensional scaling to extract the "dominant dimensions" of musical style. These two studies have been mentioned earlier, so a careful look at others in the field is in order here.

Brittin (1991) examines the effects of overtly categorizing popular music on preference in college non-music majors. The study is concerned with correlation between preference and category labels such as Pop, Rock, and Jazz. The distinction between groups is whether they were given category labels to use in classifying examples, or whether they were allowed to make up their own.

While preference was measured on a continuous Lickert-type scale, the style categories were discrete. The process by which examples were categorized, whether by the researcher or by the subjects, was not an issue. This implies an assumption of a rigid, rule-based category structure: Any given example is simply in one category or another. The question of subjects classifying the same example differently was not raised, nor was the question

of subjects rejecting the labels assumed by the researcher.

An earlier study by Eastlund (1990) measured the influence of cognitive style on discriminating musical style. While "cognitive style" refers to a body of research not relevant here, it should be mentioned that it refers to such parameters as leveling-sharpening (the ability to discern fine size differences among geometric figures) and conceptual differentiation (the number of groups formed in a sorting task). In simple terms, it claims to measure whether a student considers things as a group or as individual cases.

Eastlund pays considerable attention to the non-discrete aspects of style categories: "Listening for style is a complex process. . . Ambiguity surrounds category boundaries. . . Members of a stylistic category appear not to be related by a single rule, but seem to be related by commonalities that are sometimes difficult to verbalize" (p. 51-52). This shows healthy respect for the non-discrete Roschian categories mentioned above, although there are no references to such cognitive literature; Eastlund does cite Gardner's study (1973) extensively.

Nevertheless, Eastlund's actual experimental methodology treated style as a discrete variable. Subjects judged two passages, some from the same

composition, some from different compositions as same or different, then rated the degree of similarity on a Lickert scale. If a subject rated two passages as same when they were from the same composition, that answer was considered correct. This approach seems reasonable enough until one considers the assumptions on which it is based.

If membership in a style category is indeed so complex and difficult to verbalize, can it be assumed that two passages from a given composition will be the same style, while a passage from some other composition will not? It seems quite possible, for instance, that a given piece might change along several dimensions such as instrumentation, tempo, or dynamics within a few measures. If students consider these features important to style, they will give a response that Eastlund might label incorrect.

Eastlund's study, it must be noted, shows greater understanding of the nature of categories than most music education literature. The assumptions made are quite reasonable within the experimental framework adopted. But it reflects the fact that music educators, when they deal with categories, deal with them in rigid terms. Eastlund writes: "Stylistic analysis concerns discovery of the attributes held in common by a group of compositions" (1990, p. 51). This implies a simplistic category

structure based on the presence or absence of specific attributes. It stands in contrast with the statements regarding the ambiguous, complex nature of listening for style. It reflects what Gardner (1985) terms the classical view, where categories are discretely defined by such attributes.

Music Education and Style Categories

The idea of using such questions and answers to measure musical knowledge is nothing new. The First National Assessment of Musical Progress (1974) and The Second Assessment of Music (1980) include a section which begins with the question "Are there any kinds of music that you like to listen to?" Students were then directed through a series of similar questions and more and more specific levels. Responses were categorized under such general headings as popular or classical, then more specific headings such as jazz, vocal and instrumental art, folk, rock, soul, and country-western.

Students were also asked to classify recorded examples as similar or different in style. Again, moving to more specific levels, they were asked to classify different sub-styles of jazz such as ragtime, boogie-woogie, and Brubeck. The results of this activity were

used to assess increases or declines in the musical knowledge among different age groups.

Barrett (1989) identifies classifying as one of the core thinking skills in music:

> Classifying involves grouping items into categories based on attributes of the items. . . As the learner becomes more sophisticated, classification systems become more subtle, requiring finer distinctions. Students might be requested to classify jazz recordings by school or traditions of playing. A crucial step in classifying is the labeling process, which aids the student to recall the item at a later time, along with items belonging to the same category. (p. 51)

Gardner (1973) found that children were able to classify different musical styles at a surprisingly early age. As part of a developmental study, Gardner had children at five age levels (between 6 and 19) listen to recordings of western symphonic music. Gardner was surprised at "the overall excellence of the Ss at the task" (p. 74) despite the grave reservations of colleagues who had listened to the recorded examples he used.

With the insight that has made him a leading voice in the application of cognitive science to education, Gardner makes a critical comment that seems overlooked in many other studies:

> What are the cues in the musical stimulus? . . . Though rhythm, melody, instrumentation, and volume are probably the most prominent cues, the

> listener may also take into account details of ornamentation or interaction among instruments or any other perceptible element. For this reason one can not state with confidence on what basis a stylistic judgment is made, nor can one insure that a certain aspect is or is not a cue. (p. 74)

One question addressed here is that of automatic or preferred ways of categorizing music. That is to say, when an individual hears music, what are the first observations and associations made, and what are the first labels consciously applied? The hypothesis states that style--Classical, Jazz, Rock, Blues--will be the most frequently used domain of labels.

There is no empirical evidence that style is the most common or preferred categorical response to music; the present study seeks such evidence. There are, however, some informal indications. Consider the commonly asked question, "What kind of music do you like?" A typical response does not refer to ensemble, nation, or mood, but to style. Radio stations announce themselves as rock, country, and classical. Record stores organize their shelves similarly. Bars and clubs tend to specialize in various rock idioms, jazz, country, or blues.

Within the realm of academia, music courses are, for the most part, organized and labeled in terms of historical style periods: Music of the Renaissance, Music

of the Baroque, and Music of the 20th Century for example. Similarly, music history texts bear such labels as Music in the Middle Ages (Reese, 1968), Music in the Renaissance (Brown, 1976), and Music in the Romantic Era (Einstein, 1947).

It should be noted that informal real-world evidence about labeling was the basis for the groundbreaking research of Rosch and her associates nearly a generation ago. Eschewing the laboratory-controlled features of randomized colors, circles and squares, Rosch used tools, animals, and furniture. These were objects with which people interact on a daily basis, objects which, when people talk to each other, they instinctively and intuitively label and categorize. Surely music might be considered in very similar terms.

Styles in Music History Literature

Palisca (1981) asks a question that is at the heart of this study: "Is it justifiable to lump together into one stylistic period and under one label such diverse modes of musical expression?" (p. 2). Cognitive science offers an answer, but not a simple one. It does indeed seem justifiable, even essential to organize information and label one collection of items differently from another. But it must be remembered that such

organizations and labels are not rigid or discrete, and that any two items so lumped might seem quite different when viewed from various perspectives and by various observers.

Seaton (1991) points out that "learning theory has clearly shown that information is only absorbed and retained when it is incorporated into some coherent pattern" (p. vi). It could well be argued that categories, style-based or otherwise, provide such coherent patterns. Surprisingly, Seaton addresses the issue of categories as more than feature-bundles without explicitly using cognitive terminology:

> "Western musical tradition is best regarded not as one of changing traits of style *per se* but as one of changing models for musical expression. These models in turn justify the articulation of music history into periods. . . They also account for historical divisions and connections at different levels." (1991, p. vi)

Seaton's "changing models" seem closely related to prototypes, with style periods being centered on specific exemplars rather than feature-bundles or rules. Far from negating the existence of style categories, these models actually "justify" them. An example should clarify this.

In musical terms, it might be argued that Mozart's Eine Kleine Nachtmusik (K. 525) is a clear (prototypical) case of 18th century classical style, while his Fantasia

in C minor (K. 475) is less typical. Nachtmusik exemplifies clear, tonal, harmonies and melodies, with classical tonic-dominant key relationships, articulated phrases and sections, and the use of sonata and rondo forms (Seaton, 1991). The Fantasia is more dramatic and chromatic, exemplifying a style more akin to 19th century Romanticism (Grout, 1973).

If one were teaching the 18th century style, Eine Kleine Nachtmusik would be the better model. It exemplifies those features and forms that characterize the First Viennese School. Keeping in mind the graded nature of the category, however, it would be misleading to ignore the Fantasia; it is an atypical member, but a member nonetheless. Presented with this atypical example, the students are likely to experience some confusion, but in the end they will have a more realistic idea of what was going on in 18th-Century Vienna.

Contributions of the Present Study

The first aspect of this study that sets it apart from existing literature is that musical style will be the domain of knowledge. No such study is available in cognitive expert-novice categorization literature. This time, a highly-structured body of information is the subject matter, not trees and birds. Similar studies have

been done with organized domains such as clinical psychiatry (Murphy & Wright, 1984) and Physics (Chi, Feltovich, & Glaser, 1981), but for the most part the research uses general knowledge domains such as dogs and birds (Tanaka & Taylor, 1989) or trees and furniture (Rosch et al. 1976). Since expertise is normally associated with highly-organized and structured domains, it seems appropriate to use music history here.

There are methodological problems raised by the use of actual recorded music. The response time device would be difficult to use because the musical example requires a number of seconds to hear and categorize. Gardner (1973), Seaton (1991) and others point out the complexity of musical style and the difficulty of verbalizing its features. But music teachers and their students must cope with this complexity on a regular basis. This study will offer some empirical information regarding the different ways students (novices) and teachers (experts) might perceive the very same recorded example.

Unique Features of the Present Study

Rosch, Mervis, Gray, and Boyes-Braem (1976) claimed to have discovered universal principles of categorization. These universals were not, they stated, the content of any category. Rather, the principles of category formation

were the universals. Thus one culture or one individual, while following the same universal principles, might come up with a different set of category members from another culture or individual.

With regard to the present study, this means that the canonical taxonomy of Baroque, Classical, and Romantic music is not a fixed list as it might appear from traditional texts. There certainly are commonalities and differences, and there certainly are categories. But the structure and contents of those categories need further investigation. Rosch's airplane mechanic, when asked to envision an airplane, did so from below, while the other subjects used a side view. Perhaps music experts, like the mechanic, hear music from a totally different angle.

Contributions of this Study to Music Education

Musical categories are obviously different in content from those of natural objects such as dogs and birds, but are they different in structure? Vertical levels and graded typicality have been demonstrated in natural object categories (Rosch et al., 1976). It would be wrong, however, to assign these characteristics to only bird, dog, and furniture categories. They have been tested and supported in such diverse fields as corrections (Lurigio, 1983), psychiatry (Murphy & Wright, 1984), and medical

diagnosis (Medin, Altom, Edelson, & Freko, 1982). Barsalou (1989) found these characteristics in what he termed goal directed ad hoc categories.

Murphy and Wright (1984) consider this issue. "Although some people do attain expertise in natural object domains usually investigated by psychologists (e.g., birds, flowers tools), many are experts in more abstract domains, like chess positions, musical styles, diseases, foreign policy, and so on. One might argue that these abstract domains are very different from object categories, but there seems to be little reason to believe that the effect of expertise should be qualitatively different in the two domains" (p.146, this writer's emphasis). In other words, the expertise effects found in object categories might well be applicable to more abstract categories such as musical style and genre.

There are indeed differences between dogs and composers, and in the way people become experts in the different fields. Much of the knowledge about famous composers has come from academic study, not direct experience. Tanaka and Taylor (1991) theorize that physical interaction gives their dog experts an advantage over bird experts, who must view from a distance. Lakoff (1987) argues that real-world physical interaction and bodily experience might be the basis for very early

categorization of objects. Could musical expertise be of a more abstract and less physical nature?

The Special Case of Music Expertise

Musicians do have an opportunity to interact with the music of famous composers, but in a very different way from experts in other fields. They play, sing, or conduct this music, all quite physical activities; they analyze it using traditional harmonic and rhythmic approaches. Experts, it seems reasonable to assume, have had more interaction with music than novices, more years to practice, play, conduct, and listen.

Summary

With context and expertise as the independent variables, this study should be of interest to music teachers at every level. Although some teachers are obviously more experienced and more knowledgeable than others, it still seems appropriate to think of them as more expert than their students. It also seems appropriate to assume that, in earning their degrees and credentials, they have been exposed to music at higher levels and in greater depth than the vast majority of their students.

As a result of these different backgrounds, the teacher hears the music in a different context from that of the student. The teacher, for example, considers the lesson plan, the important concepts to be learned, and chooses a musical selection that demonstrates those concepts. This choice is made by an informed person who has spent considerable time listening to and studying music, including the example to be played.

Students, on the other hand, come from a culture dominated by electronically reproduced commercial music. They enter a room known as the music room, where certain kinds of music are heard quite unlike that with which they are familiar. They very probably never heard the musical selection chosen by the teacher. Teacher and students listen to the same recording, but with different knowledge bases and in different contexts. There is a serious question as to whether they hear the same things.

If musical categories are not the discrete, feature-defined structures described by Plato, then an alternative model for them must be found. The organization and presentation of musical information would benefit from an empirical cognitive base. This study could serve as a beginning of an ever widening and ever deepening understanding of musical concepts, and a fresh approach to their teaching.

CHAPTER III

METHODOLOGY

The Pilot Study

In the pilot study, all subjects were music experts (N=16), and the independent variable was context. Half of the subjects listened to a narrow range of all "Classical" excerpts; this was analogous to the all-Norton Anthology tape (CTXT1) in the larger study. The other half listened to a mixture of popular styles as well as Classical, analogous to the Rock, Blues, and Jazz tape (CTXT2). Thus context was a variable, while expertise was not a factor.

Even though subjects were asked to write down the style of music they heard, their responses included genres, national origins, and other kinds of labels. The use of style terms was greater in the mixed context: $F(1,14)=8.8$, $p=0.0102$. That is, the musical context appeared to have an effect on the kind of labels people used. These results are summarized in Table 1.

TABLE 1. Analysis of Variance of Mean Use of Style Terms in Different Contexts (Pilot Study)

Source of Variation	Sum of Squares	df	Mean Square	F	p	F Critical
Between	0.2916	1	0.2916	8.812	0.010	4.6001
Within	0.4633	14	0.0331			
Total	0.7549	15				

The Larger Study: Expertise and Context

The pilot study showed some influence of context on the labels used by experts. The main premise of the larger study is that experts and novices will react differently to the change in musical context. In the larger study, novices were expected to show an even greater change in the kinds of labels they used in the two different contexts. That is, the kinds of labels used by novices should be quite different when all the examples are Classical as opposed to a mixed sampling of styles.

Subjects

All subjects were upper division or graduate students at the University of Oregon. There were 32 music majors (experts) and 32 non-majors (novices), and each subject listened to a series of 40 brief recorded examples, 12 of these being the target examples for the study. This produced a pool of 768 responses, a reasonable pool from which to test the hypotheses.

Subjects were recruited by advertisements posted in the Music building, on the Psychology "paid studies" bulletin board, at several locations around a graduate housing complex, and several other central bulletin boards. Subjects were paid $5 for participation, or given credit through the psychology department subjects program.

Upper division and graduate music students were desired because they would have taken the sequence of music history courses that would give them considerable familiarity with the Renaissance, Baroque, and Romantic styles in question. While some juniors would also have had these courses, the case for expertise could be made more securely with juniors, seniors and graduates. This also captured the experts without specifically mentioning music history or styles, a mention which might easily have influenced the kinds of responses they gave. In effect,

it avoided "tipping them off" to the researcher's desired results.

The novice sample was simply chosen to match the experts: juniors, seniors and graduates, but with no academic music background. Thus experts and novices could be assumed to be approximately equal in age, intellectual ability, and non-musical academic experience. Any differences in performance, therefore, could indeed be attributed to musical expertise.

Measures

The Instrument

As described in the preceding chapters, the categorization of music was measured by a free response form headed with the question "What kind of music is this?" While much of the categorization literature uses response times to cues, this was not practical in the case of music. Milliseconds of response time might be meaningless when the music itself takes a significant amount of time to be perceived and labeled.

Tanaka asked subjects "to say the word that names the object as quickly as possible" (1989, p.20). Upon hearing the response, Tanaka recorded whether it was a BLT or otherwise. For the musical study, an equivalent procedure

was to have the subject write his or her own response. Answering aloud might have interfered with listening to the music, would have been difficult to record when more than one subject was participating, and might have led to inaccurate judgments on the part of the researcher. Consistent with Tanaka's procedure, subjects were asked to write their first thoughts and not to change them; furthermore, only the very first word was used as data in the study. The inference was that this might best capture their first thoughts about the music. Another assumption was that "Italian Baroque" represented a different response from "Baroque--Italian. [see APPENDIX A: Response lists].

Of great importance was the exact selection of musical examples presented on the tape. The intent was to present a stable, reliable set to the subjects, but in two different contexts. While other studies (Gardner, 1973) had attempted this with different passages from the same composer or piece, this study required more precision. Also, the variable of musical context had to be manipulated with some validity: One could argue for endless factors that influence a listener's perceptions. These two problems, a reliably stable set and a valid change of context, were solved with a series of steps more stringent than those found anywhere in the literature.

Selecting and Presenting the Examples

The target examples are from the Norton Anthology of Western Music set of compact discs, and are listed in the accompanying text as Renaissance, Baroque, and Romantic examples. Classical (18th century) examples were avoided because this label might be confused with the generic basic level term. The NAWM in a CD format satisfied the conditions of (a) high quality of performance and recorded sound, (b) clearly identifiable stylistic categories, and (c) ease in capturing the exact same 20 seconds of music on several different context tapes.

Nearly any given example of music has too many attributes to list and control (Gardner, 1973). An ideal experimental situation would have been to select and present passages purely at random. Unfortunately, with 40 examples total and 12 targets, early attempts at a random list were unsatisfactory. For instance, one randomized list produced nearly all group rather than solo music, while another was nearly all vocal with very little instrumental. It was feared that these attributes might affect the labeling of the few cases that were different. That is, it would unintentionally highlight the one or two cases of solo or instrumental music; in effect, it would introduce an undesired variable or a priming effect.

To accommodate all of these concerns, a constrained randomization was used, with the commitment to a balance of vocal and instrumental and of solo and ensemble music. All NAWM Renaissance, Baroque, and Romantic selections were assigned random numbers and so ordered. From this order, examples were placed in four separate columns: vocal solo, vocal ensemble, instrumental solo and instrumental ensemble. With selections from rotating columns, the resulting target set presented a more balanced sampling of music to the subjects than that resulting from pure randomization.

One other factor played a role in the selection. Music of many styles and genres often begins with a few introductory bars. Often this passage is markedly different from the main body of the piece, as when a piano plays the opening measures of a Schubert Lied, or an *allegro* movement begins with an *adagio* prelude. The compact disk format allowed for precise timing, so that each selection was started at least 10 seconds into the piece. If this resulted in a major change (such as the entrance of a choir or soloist) very near the end of the passage, the starting point was adjusted to either include the event as a significant part of the 20 second passage or to eliminate it.

Setting the Musical Context

Once again, a myriad of features might be used to define "musical context." This study, however, focused on style as a variable, so the two musical contexts were defined on the basis of style. Context 1 was a constrained all-NAWM set; Context 2 a mixed Blues, Rock, Jazz, and NAWM set.

To ensure that some context was established by the 28 non-target examples, the targets were shifted toward the end of the sample and mixed in. Starting with the final position (number 40), the targets were spaced with one, then two, then three of the context-setters between them. This pattern was repeated until the twelfth target was placed (position number 8). Consequently, subjects heard 7 context-setting selections before hearing the first target.

Once the 12 targets were chosen, the other examples were selected in a similar constrained random order. The NAWM examples were taken from the same list as the targets, with attention to solo/ensemble and vocal instrumental features. The Blues and Jazz examples were taken from two anthologies, somewhat equivalent to the NAWM. The intent was to provide an arguably matched set

that was already labeled and categorized, creating sub-categories within the generic Jazz and Blues labels.

The Smithsonian Classic Jazz collection (1988) provides high quality CD recordings of significant Jazz works. These works are categorized by the editors under such headings as Ragtime, New Orleans, Swing, and Modern (Williams, 1987). Since these selections were context setters rather than data relevant targets, less stringent measures were used. A sampling of three tracks from each of the Early Swing and Modern categories was drawn up, then randomized for position on the tape.

The same process was used for the Blues selections, using the anthology Blues Masters (1993). These examples came from Volume 7, Blues Revival, volume 8, Delta Blues, and volume 9, Postmodern Blues. Again, subjects could respond any way they chose, but they at least had the opportunity to use subsets of the category "Blues."

For the Rock selections, no authoritative anthology was readily available. A great number of "best hits of.." collections exist, but these tend to be commercial promotions by a given company rather than carefully constructed anthologies. A collection of CDs representing at least three sub-styles was needed, so examples included early rock from the 1950s and 1960s (Elvis, Chuck Berry), British groups (Wings, Dire Straits), and Southern groups

(Lynard Skynard, Allman Brothers). As with the NAWM, Jazz, and Blues examples, a subject had the opportunity to respond with the generic "rock," a more specific label, or in any subjective manner chosen.

Finally, to avoid ordering effects, an additional tape was made in both contexts with the examples in reverse sequence. Targets still occupied the same positions (from number 8 to number 40) but in reverse order; Context setters also occupied the same positions but in reverse order. The original order was dubbed A, the reverse B; thus 4 tapes were used: 1A, 1B, 2A, and 2B. A complete listing of all titles, composers, and artists appears in Appendix B, Table 2.

The Variables

The two nominal independent variables were expertise and context. Experts were the music majors, novices the non-music majors. All subjects were upper division or graduate students, so the groups could be assumed to be matched on such factors as maturity, intelligence, and general linguistic abilities. Context 1 was the all-Classical tape, context 2 the tape using mixed styles. The "context setters" were the 28 examples which preceded and followed the 12 target examples.

As explained in Chapter 1, the dependent variable was the pool of 768 words used by subjects to describe the musical examples. This pool, however, was treated as three separate variables because the research questions examine three different portions and three different aspects of the 768 words.

The first of the three hypotheses tested was that style terms would occur more often than random. That is, words referring to a historical musical idiom period such as Classical, Renaissance, or early 16th Century would constitute a higher percentage of the pool than references to the instrument (voice, piano), the genre (opera, symphony) or identification (composer or title). Relevant to research question 1, the total number of style terms was compared to the total response set.

Basic level terms (BLTs) were by definition those words used most frequently by the novices. Studies demonstrate that experts use more specific labels than the general public. However, context might have some effect on this difference. The third hypothesis states that the use of BLTs will be different between the two contexts, and that this effect will be different for experts and novices. In other words, context and expertise will interact. This is the central issue of research question 2.

Finally, for any given target, each subject might use a unique label that no other subject used. The hypothesis, however, stated that experts would use the same term more often than novices. Within each group, every time a subject used the same term as one or more other subjects, it was counted as a SET. Research question 3 concerns the influences of expertise and context on SETs.

Validity and Reliability

BLTs are in one sense a mathematical phenomenon, in that they are the most frequently appearing words (Rosch et al., 1976). There are few threats to validity when one term appears with substantially greater frequency. SETs are more problematical and some subjective judgment is involved. Equivalent terms include words that have, for all practical purposes, the same meaning. For example, two experts using the labels "Renaissance" and "16th century" might be said to agree in their categorization of an example. "Church" and "sacred" might also be argued as equivalent terms. "Mass" and "Kyrie" would not: The latter is more specific than the former; similarly, "early Renaissance" and "Renaissance," "chamber music" and "string quartet," and "concert choir" and "a cappella chorus" would not be considered exact equivalents. A

panel of the researcher and two music history faculty members made independent judgments of equivalency.

Design and Hypotheses

This study involves six research questions addressed by three hypotheses, so it might be best to represent it in three parts. Research question 1A concerns the frequency with which style terms (such as Classical, Baroque, and Sixteenth Century) are used overall; Question 1B concerns novices using "Classical" as the BLT. Research questions 2A and 2B concern the level (generic or specific) at which subjects label examples. Research questions 3A and 3B concern how often subjects use the same labels as others in their group. Each research question, of course, has an appropriate hypothesis, but it will be clearer if each hypothesis has a subscript based on its variable. The hypothesis for style terms, that is, will be symbolized H_{ST}, while those for basic level terms and shared equivalent terms will be symbolized as, respectively, H_{BLT} and H_{SET}.

Hypothesis Number One: Style Terms (H_{ST})

If style is not a preferred basis for categorizing music (null hypothesis) then STs should occur with the same frequency as other kinds of labels. If their frequency is higher than expected, however, then the alternative hypothesis is supported. STs will be compared with genre (opera, church, symphony), media (piano, orchestra, choir), national origin (Italian, French), and other. The hypothesis states that subjects will use style terms more often than genre, media, nationality, or other terms.

$H_{ST}0$ (null): STs = Genre = Media = National = Other

$H_{ST}1$: STs > Genre ≈ Media ≈ National ≈ Other

Hypothesis Number Two: Basic Level Terms (H_{BLT})

A BLT cannot be determined until after the pool of responses is carefully examined. It is operationally defined here as the style term used most frequently by the novice population. This is consistent with studies from Rosch et al. (1972) to Tanaka and Taylor (1991).

Novices are expected to use the BLT significantly more often than experts; experts tend to use more specific

(subordinate level) terms. This is a well-documented phenomenon in categorization literature, but has not been tested in a musical setting.

The situation is more complex than that, however, when context is introduced as a variable. The prediction is that Expert-Novice differences in BLT use will be greater in one context than in the other. There is no clear basis for predicting whether the change from CTXT1 to CTXT2 will cause an increase or decrease in Novice BLTs, only that it will cause a change. The hypothesis states that Experts will use fewer BLTs than Novices, and that musical context will have an interactive effect on this between-group difference.

$H_{BLT}0$ (null):

No difference between EXP/NOV or between CTXT1/CTXT2

$H_{BLT}1$: Exp BLT < Nov BLT

$H_{BLT}2$: CTXT1 BLT ≠ CTXT2 BLT

$H_{BLT}3$: Interaction

Hypothesis Number Three:
Shared Equivalent Terms (H_{SET})

Experts should share more equivalent terms than novices. That is, they will agree more often and use the

same label (or equivalent labels) for a given target example. Murphy and Wright (1984) found this to be true with clinical psychologists, comparing experienced senior staff members with trainees. The number of SETs will be tallied for both groups and both contexts. Interaction is predicted: That is, a change of context should have a greater effect on the novices than on the experts. As with the BLT comparison, there is no clear basis for predicting whether the change from CTXT1 to CTXT2 will cause an increase or decrease in Novice SETs, only that it will cause a change. The hypothesis states that Experts will use more SETs than Novices, and that context will have an interactive effect on the between-group differences.

$H_{SET}0$ (null):

No difference between EXP/NOV or between CTXT1/CTXT2

H_{SET} 1: EXP SETs > NOV SETs

H_{SET} 2: CTXT1 SETs ≠ CTXT2 SETs

H_{SET} 3: Interaction

The study took place on the University of Oregon campus between April and July, 1993. All tapes and response sheets have been kept for possible use in further analysis. Detailed lists of all recorded examples appear

in APPENDIX B, and the subject response forms used appear in APPENDIX C.

CHAPTER IV

RESULTS AND ANALYSIS

Research Questions and Results

The study is based on three main hypotheses, each addressing two research questions. Hypothesis one is concerned with how often subjects used style terms and how often novices used a particular term, "Classical." Hypothesis two is concerned with the use of this basic level term as influenced by differences in expertise and context. Hypothesis three is concerned with how often members of a group use the same labels, and how this is influenced by differences in expertise and context. The results will be discussed in the order of the research questions as stated in Chapter 1.

Style Terms

Research Question #1A: ST Frequency

The first hypothesis stated that subjects would use style terms (STs) more often than genre, performing media, national origin, or other kinds of labels. STs consisted

of the established labels Renaissance, Baroque, Romantic, and the generic Classical. In addition, any term that referred to a historical perspective (16th century, early, modern) or a commonly used musical idiom (folk, jazz) were taken as style terms. Based on the pilot study and the actual responses gathered in this study, responses were sorted into 5 nominal categories including style terms. The other nominal categories were: genre (opera, church, symphony); media of performance (vocal, choral, piano); national or linguistic (Italian, German), and other. As a test for reliability, the researcher's classification of label types was compared with those of two music history faculty members. This panel of three agreed with 78% reliability.

This prediction was made independent of group or context, so the overall count of STs was analyzed rather than cell by cell. Subjects used STs 319 times out of a possible 768; other response types are tallied in Table 2:

TABLE 2. Summary of Label Response Types

Style	Genre	Media	Nat'l	Other
319	155	163	54	77

FIGURE 1 demonstrates even more clearly the predominance of style terms compared to those of genre, media of performance, national origin, and other kinds of labels.

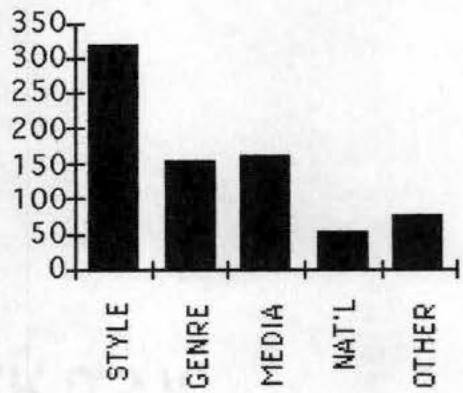

FIGURE 1. Comparison of Label Response Types

Chi Square was calculated; χ^2 = (4, N=768, p< .001). Clearly, subjects used STs more than other kinds of labels. This supports the first hypothesis.

Research Question #1B: Specifying the Basic Level Term

The BLT for style was operationally defined as the style term used most often by the Novice group (representing the general population). As predicted, "Classical" was the most frequently used label, 78 out of

a total 129 STs. The next most frequently used label was "opera," a genre rather than a style term, at 54 times. Among STs, the next most frequently used labels were "Baroque" and "Medieval" at 10 times each, not even approaching "Classical" at 78.

A chi square test proved significant, χ^2 (1, N = 129, p = .0146), clearly establishing "classical" as the most frequently used style term among novice subjects. These results are summarized in Table 3.

TABLE 3. Frequency of "Classical" as a Style Term

"Classical"	Other	Total
78	51	129

Basic Level Term Response Patterns

The Basic Level Term for Style has now been established as "Classical." The next step is to find how use of this BLT is influenced by expertise and context.

Research Question #2A and 2B: BLTs with Expertise and Context

The second hypothesis stated that novices would use BLTs more often than experts, and that this difference

would be influenced by context (i. e., interaction was predicted). A 2X2 ANOVA tested this prediction.

Main Effects

As predicted, there was a significant difference between the groups, with experts using the label "classical" 22 times and novices using it 78 times. Thus expertise was a significant source of variation. This is consistent with nearly all of the expert/novice literature, and with Tanaka and Taylor (1991) in particular.

There was, however, no significant difference due to context in either group ($p = .441$). These results are somewhat at odds with Barsalou's context dependent categorization theories (1989). That is, a change of musical context did not change the use of the generic "classical" label. The results support category stability across context. Brooks (1989) suggested stability among the more experienced group, but the stability in the Novice responses is somewhat surprising. Figure 2 shows the clear difference between groups, with no influence from context.

Comparison of BLT Use

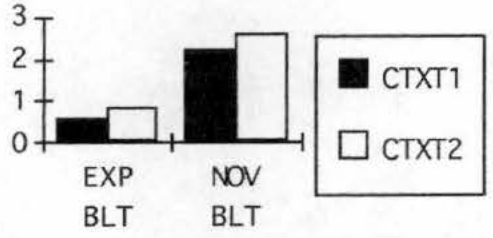

FIGURE 2. Comparison of Basic Level Term Use by Group (EXP/NOV) and Context (CTXT)

Interaction

The prediction of interaction between expertise and context was not supported by the results. Expertise alone seems to be a factor in the use of the basic level style term "classical." That is, experts persisted in using "Renaissance, Baroque, and Romantic" while the novices persisted in using "classical" whether the context was all Norton Anthology or rock, blues, and jazz.

TABLE 4. Analysis of Variance of Mean Use of Basic Level Terms

Source of Variation	Sum of Squares	df	Mean Square	F	p	F Critical
Context	1.5625	1	1.5625	0.574	0.451	4.0012
Group	49.0	1	49.0	18.02	0.00	4.0012
Interaction	0.0625	1	0.0625	0.023	0.88	4.0012
Within	163.13	60	2.7188			
Total	213.75	63				

The lack of interaction is demonstrated more clearly in FIGURE 3. The lines are nearly parallel, indicating that while the two groups differed to some degree, context had no effect within each group. Interaction would be indicated by more divergent or crossed lines.

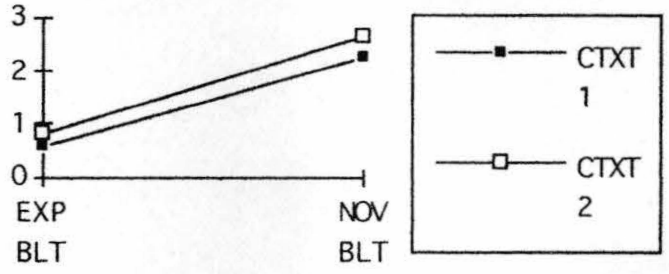

FIGURE 3. No Interaction Between Group and Context

Shared Equivalent Term Response Patterns

Research Questions #3A and 3B: SETs with Expertise and Context

SETs consisted of any term used by two or more subjects on a given example, or of any term that was an equivalent (such as "choir" and "choral"). This part of the study tested the theory that categorization by experts is more stable than categorization by novices. The hypothesis stated that a change in context will have little influence on the experts, but a significant influence on the novices (i.e., interaction between expertise and context). As with style terms, a panel of the researcher and two music history faculty independently judged for equivalency. This panel agreed with 88% reliability.

The results did not support this hypothesis: neither expertise nor context was a source of significant variation. Although experts agreed more often than novices (289 versus 254 times) the difference was not significant: $F(1, 60) = 3.12$, $p = .08$. Context had virtually no effect at all: $F(1, 60) = .06$, $p = .80$. FIGURE 4 shows a difference, though not significant, between experts and novices in their use of SETs, as well as the lack of influence of context on either group.

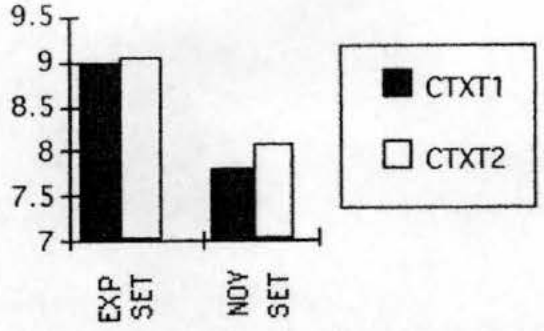

FIGURE 4. Comparison of Shared/Equivalent Term Use by Group (EXP/NOV) and Context (CTXT)

In greater detail, TABLE 5 shows the ANOVA for both groups in both contexts, and FIGURE 5 shows the lack of interaction between group and context.

TABLE 5. Analysis of Variance of Mean Use of Shared/Equivalent Terms

Source of Variation	Sum of Squares	df	Mean Square	F	p	F Critical
Context	0.3906	1	0.3906	0.064	0.802	4.0012
Group	19.141	1	19.141	3.118	0.083	4.0012
Interaction	0.1406	1	0.1406	0.023	0.880	4.0012
Within	368.31	60	6.1385			
Total	387.98	63				

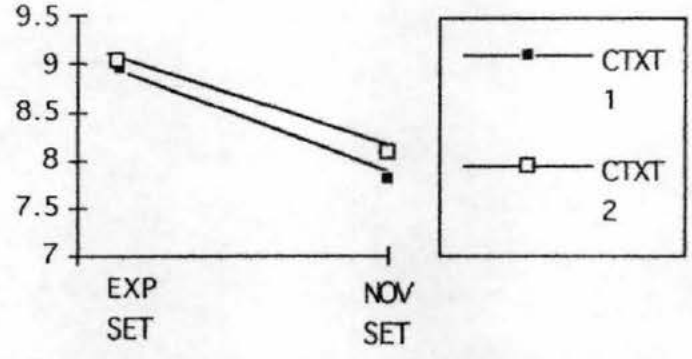

FIGURE 5. No Interaction Between Group and Context

Once again, the categorization patterns appear more stable than might be expected, even within the novice group. Subjects in both groups agreed with their peers whether the context-setting excerpts were all Renaissance, Baroque, and Romantic, or Rock, Blues, and Jazz. The expert/novice differences suggested by Murphy and Wright (1984) did not appear, nor did the context effects suggested by Barsalou (1989).

Conclusions

In some cases, the results supported the hypotheses, while in other cases they did not. Some of the results, indeed, appear rather inconsistent with previous research and with the pilot study. The following discussion will

consider the possible reasons for both the expected and unexpected results.

Style Terms and BLTs

The results support the use of style terms as the most favored manner of categorizing music, significantly more so than genre, performing media, or national origin. Among the STs, the most frequently used term for the novices was "Classical," establishing it as the basic level term for musical style.

Experts tended to use more specific level terms (Renaissance, Baroque, Medieval; Early . . ., and Late . . .), as suggested by Tanaka and Taylor (1991). The difference was significant: $F(1,60) = 16.2$, $p = .0002$. Use of BLTs, however, was not influenced by context. Even when all the examples were "Classical" (actually Renaissance, Baroque, and Romantic, CTXT1) novices used that generic term 36 times. Novices used the term only slightly more often (42 times) when the target examples were preceded by Rock, Blues, and Jazz, and the difference was not significant.

The implications are that musical experts, like those in other domains, categorize music at more specific levels. They rarely used the generic term "classical," preferring the more specific Renaissance, Baroque, or

Romantic labels. This difference persisted across the two contexts, suggesting that level of categorization is a function of expertise but not of context.

Shared Equivalent Terms

Within each of the four cells, subjects used the same or very similar labels (SETs) 543 times out of a response pool of 768. There were no differences between groups or between contexts. The number of times subjects agreed overall far exceeded that of disagreements. That is, subjects used the same label for the same example more often than they used different labels. This difference was significant when tested with a chi square: χ^2 (1, N = 768), $p < .001$. In other words, independent of group or context, all subjects agreed with their peers significantly more often than they disagreed.

The results from each part of the study clearly support category stability within a group. The only differences that appeared were expert/novice differences in the BLTs. The way individuals label or categorize a passage of music seems to be a relatively fixed response. The intensive study of music in the western art tradition leads to more specific levels of categorizing, but does not lead to more consistent agreement with one's peers.

CHAPTER V

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

Summary

This study began with a question. That question was simple and direct: "What kind of music did Mozart write?" The answer has proven considerably more complex and lengthy than one might expect. This complexity is not merely a function of Mozart's music, but of human cognition as well. To say that every person hears something different in Mozart's music is as great an oversimplification as it would be to say we all hear the same thing. The truth lies somewhere between these extremes.

There was a rich diversity in the labels used to describe the musical examples. Some subjects used highly subjective and imaginative language, including "On Golden Classical Pond" and "Medieval Monks bemoaning their celibacy." It is perhaps regrettable that such creative descriptions might be marked "wrong" on a test using teacher-generated labels. A fundamental premise of this study is that such labels are just as valid and meaningful

as any textbook terminology. The fact that they are neither style terms nor shared equivalent terms--as defined here--has no bearing on their usefulness as an indicator of the subject's knowledge and perceptions of the music.

Nevertheless, for the most part, both expert and novices used a surprisingly stable set of labels, and this stability was not significantly affected by context. Members of both groups tended to agree with their peers more often than they disagreed, and a relatively few terms (Classical, opera) dominated the pool of 768 responses.

Research Questions and Results

Research Question #1A

Will subjects use style terms (STs) more often than genre, performing media, national origin, or other kinds of labels?

The results of the study supported the hypothesis regarding use of style terms: Subjects overall used STs significantly more than any other kind of label. That is, they used labels such as Classical, Baroque, and Renaissance more often than labels such as piano, symphony, and Italian.

Research Question #1B

Will Novices (representing the general public) use the generic, basic level term "Classical" significantly more often than other style terms?

Novices did use "Classical" significantly more than any other ST, supporting the hypothesis. This establishes it as the basic level term (BLT) for style. For the purposes of this study, and in keeping with the research literature, the BLT has been operationally defined as the term used most often by the Novice group.

Research Question #2A

Will Novices use the generic, basic level term "Classical" significantly more often than Experts?

The results supported the hypothesis: Novices did indeed use the generic BLT "Classical" significantly more often than Experts. This was consistent with the findings of nearly all of the literature. It is hardly surprising that Experts would use more specialized terms in describing the musical examples, but there was some evidence that the between-group difference might be influenced by context.

Research Question #2B

Will musical context influence the use of BLTs? That is, will any difference between the two groups be greater with one tape than with the other?

Context had no significant effect on BLT use. Novices used more BLTs than Experts whether the music was all classical (CTXT1) or mixed styles (CTXT2). These results failed to support the hypothesis, and run somewhat counter to the arguments of Barsalou (1989). They also appear inconsistent with the pilot study, wherein a similar manipulation of context led to a significant change in the kinds of labels used.

Research Question #3A

Will Experts use the same label for a given example significantly more often than Novices?

The difference between the two groups was not significant, although experts agreed with each other slightly more often than novices. Again, this runs counter to some of the literature (Brooks, 1989), which had suggested that experts would be more consistent in their labeling, especially when the context was changed. It appears that the Novices used a few terms ("Classical"

and "opera" in particular) so often that it led to a high level of agreement.

This high level of agreement explains the difference between these results and those of Murphy and Wright (1984). In that study, the setting was a children's clinical psychology center. The experts, like those in the music study, were veterans who tended to agree with each other in their labeling of the cases presented. The psychology novices, however, were beginning trainees who, while not yet experienced, had some motivation for appearing astute and perceptive in their descriptions of cases that they would indeed soon be treating. The music novices, in contrast, had no motivation to use anything other than the most familiar and superficial vocabulary. Furthermore, they were almost certainly influenced by the forces mentioned early in this study: the marketing practices of record companies and radio stations and the colloquial practice of everyday informal discourse about music. Not surprisingly, the very responses used most often by the novices are those used by junior high students in Cutietta (1993): Classical and opera. This even suggests that the supposed novices <u>were indeed experts</u> in the sense that they had learned and practiced these informal categories years earlier.

It should be noted that, consistent with Tanaka and Taylor (1991), the expert SETs were at more specific levels, and carried more detailed information than the novice SETs. This was also in keeping with other expert-novice studies including Chi, Feltovich, and Glaser (1981), and Lurigio (1983). Nevertheless, novices used shared or equivalent terms just as often as the experts.

Research Question #3B

Will musical context influence the use of SETs? That is, will any difference between the two groups be greater with one tape than with the other?

There was no significant difference between the two groups or between the two contexts, and there was no interaction. Although the different groups used different terms, they tended to agree with members of their own groups. This is, perhaps, the most surprising result of the study. Experts, the literature suggested, might be fairly consistent and stable in their use of learned, specific labels. But there was much evidence that novices, with limited training and vocabulary, would use different kinds of terms in different contexts. That was not the case.

Barsalou (1989), more than any other author, argues that categorization is not a stable process. Depending on

the circumstances, he proposes, the same person might categorize the same item different ways at different times. But even with completely different "context-setting" musical excerpts (all Classical versus Rock, Blues, and Jazz) categorization patterns remained stable.

Conclusions

People--experts and novices--agreed significantly more often than they disagreed on what kind of music they were listening to. In this sense, some might argue that they "heard the same thing." But the overwhelming evidence from categorization research on which this study is based does not allow such a simplistic assumption. Individual theories about music, heuristics, experience, training, and an ever changing environment all have an impact on the choice of a label. More importantly, they have an impact on the individual's concept of what that label means. Two people using the same label, in other words, do not necessarily mean the same thing. The obvious case, returning to the original question, would be the expert and novice agreeing that Mozart wrote "Classical" music, but meaning different things.

The expert probably uses the term to designate the 18th Century Viennese Classicist School, associated with Haydn, Mozart, and early Beethoven. The novice probably

means music for an orchestra, played in concert halls by people in formal outfits. This case is so obvious, in fact, that music of the Viennese Classicists was intentionally eliminated from the Norton Anthology selections of Context 1. It is quite possible that, in the few cases where experts used the term "Classical," they were mistaking some late 17th-Century piece for an early 18th-Century piece. Such a judgment is entirely consistent with the concept of non-discrete, non-Platonic categories.

A label, then, does not carry with it the traditional Platonic discretely defined category; agreement on a label does not necessarily imply agreement on a list of defining characteristics. This is the fundamental implication found in Rosch et al. (1976), Barsalou (1989), Medin and Wattenmaker (1989), Murphy and Wright (1984), Tanaka and Taylor (1991), and nearly all of the literature on this complex subject. It is also the thought that must be held in mind while evaluating the results of this study.

Use of Style Terms

As predicted, both groups used style related terms more often than other kinds of terms. The present study makes no claims as to whether this is an effect of conditioning by radio stations, record companies, and

music courses, or an effect of some innate aspect of human cognition. Such questions might be the focus of future studies.

Cutietta (1993) assumed style to be the preferred basis for categorization and searched for defining features. He adopted his own hypothetical feature list (motion, energy, flow, fabric, and color) and asked students to rate the music on the these factors. This approach primes the subjects to attend to their own perceptions of these factors.

In contrast, the present study merely asked "What kind of music. . .;" had the question been "What style?" or "What kind of instruments?" or "When do you think this was written?" subjects would have been primed to respond in a particular manner. In effect, the wording of the question avoided the imposition of a schema and allowed each subject to construct his or her own basis for responding. Cutietta's assumptions about flow and fabric might very well be accurate, but a design free of such priming cues would lend them greater validity.

Basic Level Terms

Gardner (1973) surprised himself and his musical colleagues when he examined young children's ability to consistently relate passages from the same composer.

Musical style categories seem to develop well before any overt training or study takes place. This, in fact, might be one reason for the unexpected results of the present study. In a sense, the novices had nearly as much experience as the experts at categorizing music, but at in a less formal and rigorous atmosphere--thus the differences in level but not in consistency.

Further explanation comes from a more precise examination of the two studies most closely related to this one: Tanaka and Taylor (1991) and Murphy and Wright (1984). These two studies supplied elements of methodology and design, but obvious differences were necessitated. For example, the use of real-time musical examples made it difficult to adopt a meaningful timed response format, which was an important aspect of Tanaka and Taylor and much of the categorization literature. Subjects did, at one point, write labels under pictures, which suggested the format used here.

More important is a theoretical--or perhaps semantic--difference. Tanaka and Taylor suggested that for the experts, the more specific labels <u>were</u> the basic level terms. That is, for a bird expert, "robin," "crow," "jay," and "cardinal" are as accessible and easily identifiable terms as "bird" is for the general population. The musical analog suggests that, for music

experts, "Renaissance," "Baroque," and "Romantic" actually are the basic level terms.

It might seem merely two different ways of saying the same thing: experts automatically categorize at more specific levels, or the basic level for experts is more specific. Tanaka and Taylor clearly demonstrated differences in the level at which experts and novices make their most immediate categorical response. Once that fact has been established, the distinction has more to do with defining BLT than with further inferences about human cognition. This approach might take "Renaissance", "Baroque", and "Romantic" as a composite BLT for the experts.

Even with this alternative approach, context had little or no effect. Experts used the three STs 57 times in CTXT1, 54 times in CTXT2. Novices used their BLT ("Classical") 36 times in CTXT1, 42 times in CTXT2. A Chi Square test showed that these counts do not differ significantly: χ^2 (3, N=189), p = .102. Using Tanaka and Taylor's working definition of BLT, then, it appears that there is little difference even between groups. Once again, category stability overcomes individual differences.

Shared Equivalent Terms

Neither expertise nor context had any significant effect on SETs. On the surface, it appears that experts and novices agreed with members of their own group at about the same rate. Any explanation of these unexpected results is, at this point, conjecture. Such conjecture might very well be the basis for further study, as new hypotheses arise to be tested.

The central question is still one of category stability versus changes due to expertise and context. Much of the literature as well as the pilot study suggested that response patterns--SETs in particular--would be influenced by these two variables. Since the results did not follow the predictions, a more careful scrutiny of the literature might yield an explanation.

As discussed above, the novices might have become quasi-experts in that they had learned and practiced describing music in basic level terms early in their development. The very terms they used most often were no different from those used by Cutietta's junior high students: Classical and opera (1993). This logic is somewhat analogous to Tanaka and Taylor's view of the basic level itself changing with expertise.

Tanaka and Taylor's (1991) view assumes that both experts and novices were categorizing at their own basic level: Basic level, they argued, is not inherent in the material, but in the eyes of the beholder. There was no fundamental difference in the way the two groups categorized, only in the specificity of the labels. It can be argued similarly that the music novices were "experts," but at a basic level of terminology. That is, they were just as experienced at classifying music as "Classical" and "opera" as the experts were at classifying "Renaissance" and "Baroque."

Another possibility is that the two groups agreed for different reasons, reasons that seem consistent with the relevant literature. Experts might agree because they rely on a similar base of training and study, using discrimination skills and specific labels beyond the capabilities of the novice group. Novices, on the other hand, quite possibly agree because they rely on stock, stereotypical labels that they have relied on for years.

Much of the research suggests that experts form rich, elaborate schema for a category, while novices rely on specific items or rules (Lurigio, 1983; Chi, Feltovich, & Glover, 1981; Murphy & Wright, 1984). It is possible that, following some rule for classifying music, the novices used similar terms to label the examples. This,

in fact, runs counter to the argument that novice categories will be less consistent. Perhaps the tendency to use more rigid rule-based schema overcame any influence of context.

Such rigid categorization, however, belies nearly all of the literature since Rosch. It seems to take us back to the Platonic, all-or-none kinds of categories that have been so discredited by empirical study. This apparent conundrum is the very problem that makes the teaching of musical style (and other abstract concepts) so difficult. And it is here that this study will make its most important contribution to those who would teach such concepts.

Music educators must take responsibility not only for the kinds of verbal labels their students associate with music, but also for the meaning of the labels and the processes they entail. Although context did not prove a significant factor in this study, it might be an important factor in understanding the real world situation. Among themselves, music experts readily recognize the broad, rich, and diverse nature of a category such as "Renaissance music." They know that it includes the esoteric contrapuntal devices of Okeghem, the intensely emotional and chromatic madrigals of Gesualdo, the light-hearted dance music of the English Virginal School.

When these same experts turn to their students, however, they sometime adopt a more rigid Platonic approach. In order to simplify things for the less experienced young listeners, they use definitions based on short, easily memorized feature lists. They point out salient features in a few examples, and they proclaim "That is Renaissance music." Attentive students can recite the feature list definition, and with very little practice, correctly label recorded examples.

Cognitive category literature and at least some of the results of this study point out the errors inherent in such a simplistic approach. Listening for a specific set of defining features is not necessarily listening to the music. Reciting the definition or using the approved label does not necessarily indicate in-depth musical knowledge.

Recommendations

Recommendations for Further Study

This study took theoretical and methodological approaches from categorization research and applied them--for the first time--to musical styles. In this sense, it was a first step, and, as with any new study, leaves considerable work to be done by further research. There

are many possible directions this research might follow, but it is appropriate to suggest a few here.

The range of target examples was quite limited. The idea was to present a reasonable number of context-setting excerpts, plus the twelve targets, in a reasonable amount of time. The hypotheses being tested were focused on the twelve targets because these were the only examples heard by every subject. A follow-up study might examine the entire response set, a much larger pool of labels. That is, one might examine every word written by every subject for every example. This might lend some insights into the results.

Another direction for future research might be subtle changes in the methodology. The data here consisted of the very first word written by each subject for the twelve targets. It is possible, with more elaborate equipment, to record the spoken comments of subjects, as Tanaka and Taylor (1991) did. Perhaps the very act of writing led to a less spontaneous, more considered response pool. All the subjects were college students, and all were accustomed to taking written tests. They may have experienced some interference from a "testing" mind set, even though they were assured that there was no "right or wrong" answer, and that all results were anonymous.

A third direction is perhaps the most important and the most obvious. If the results are ultimately to be considered useful to classroom music teachers, the study should be replicated with real teachers and real classroom students (i. e., children). Serafine (1991) found some important developmental trends in the perception of brief musical patterns, and these might very well have their counterpart in perceptions of style in recordings of actual music. Cutietta (1993) states that junior high students categorize Western art music as "classical," "opera," and "church" music. These category schemas seem very close to those of the University of Oregon Novice group, but no extensive generalizations should be made without more empirical support.

Recommendations for Application

What kind of music did Mozart write?

If students and teachers continue to think of categories as rigid rule-defined concepts, then they will have to sort out the contradictions when they encounter music that does not follow these rules. If "Classical" must mean music composed by the 18th Century Viennese School, then every student who uses the word in a generic sense--as well as record stores and radio stations--is wrong. They have failed the music test.

The teacher's task, under the Platonic view, is to erase the student's defining rules and replace them with those of the authorities. The student is clearly wrong, and the teacher is clearly right: "nor can any other opposite . . . simultaneously become or be its own opposite" (Plato, 1991, p.188). This either/or situation is not a function of musical style, student acumen, or teacher skill. It is a function of an entrenched view of category structure.

If the strong case for individualized, theory-driven categories is applied to the same situation, we have not only a more realistic view but also a chance for better communication. Both student and teacher categorization systems can coexist; students can indeed learn much about music without necessarily adopting a textbook list of defining features. The goal should not be to call Mozart's music "Classical" and Bach's music "Baroque." It should be to understand and appreciate fundamental differences and fundamental similarities between these two great composers. The mere use of a category label does not constitute such an understanding.

Style terms as category labels are unavoidable. It is possible to have fuzzy, heuristically defined concepts, but it is impossible to have no concepts at all. Whether concepts are "stored in our memory as prototypes"

(Jonassen, 1988) or whether individual memory traces form a concept in recall (Hintzman, 1986), the concepts are there. We need words to communicate these concepts. Knowing that the same word means different things to different people should, at least, lead to less misunderstanding. At best, it might lead to more effective communication and a better sense of mutual respect among teachers and students.

Final Philosophical Considerations

The student who complains that "all classical music sounds alike" and the teacher who believes in a rigid definition are, in a sense, victims of the same misconception. Both have concepts that are more concerned with rigid similarities than with rich diversity. Both are concerned with (arguably) superficial features rather than musical substance.

Category labels, or words for concepts, serve a vital function in our thought and in our communication. But they must be used with the knowledge that their meaning is flexible and subjective. Rosch, Barsalou, Medin, and others have struggled with the psychological and linguistic aspects of this issue, and have been duly cited throughout this study. But in the end, the final question is philosophical. Plato (1991), Aristotle, (1991), Kant

(1965), and Wittgenstein (1958) struggled with the meaning of categories and names. In examining this idea and those of other profound thinkers, one figure seemed to most clearly articulate the position and advice that emerged here.

> The facts of the world in their sensible diversity are always before us, but our theoretic need is that they should be conceived in a way that reduces their manifoldness to simplicity. Our pleasure at finding that a chaos of facts is the expression of a single underlying fact is like the relief of the musician at resolving a confused mass of sound into melodic or harmonic order. . . But alongside this passion for simplification there exists a sister passion, which in some minds--though they perhaps form the minority--is its rival. This is the passion for distinguishing; it is the impulse to be acquainted with the parts rather than to comprehend the whole. . . It prefers any amount of incoherence, abruptness, and fragmentariness (so long as the literal details of the separate facts are saved) to an abstract way of conceiving things that, while it simplifies them, dissolves away at the same time their concrete fullness. Clearness and simplicity thus set up rival claims, and make a real dilemma for the thinker. (William James, 1952, p. 4-5)

James' "chaos of facts" is too often associated with the student, while the "passion for distinguishing" is too often associated with the teacher. On the contrary, James holds that we must all face the chaos, and we must all satisfy our passion for distinguishing. The rival claims

of clearness and simplicity set up a dilemma for thinking students and thinking teachers. We must face the dilemma together.

APPENDIX A

RESPONSE LISTS

(Exact replication of Subjects' written responses with first word, fragments of following words, crossed-out words, and occasional misspelled words)

SUBJ	EXAMPLE	LABEL	SUBJ	EXAMPLE	LABEL
101	1	CHORAL MUSIC	110	1	CHORAL
101	2	MID 1900-OPE	110	2	BAROQUE ARIA
101	3	RENAISSANCE	110	3	RENAISSANCE D
101	4	CONTEMPORAR	110	4	ROMANTIC LIE
101	5	SYMPHONIC M	110	5	BAROQUE DOUB
101	6	CLASSICAL MA	110	6	SOLO GUITAR
101	7	PATRIOTIC	110	7	ROMANTIC
101	8	ORGAN MUSIC	110	8	BAROQUE ORGA
101	9	EARLY MOTET	110	9	EARLY COUNTE
101	10	PIANO SONATA	110	10	ROMANTIC SOL
101	11	MOTET-ISH	110	11	CHORAL - COUN
101	12	REN. FRENCH TU	110	12	RENAISSANCE S
108	1	CLASSICAL - CH	111	1	BAROQUE
108	2	BAROQUE OPER	111	2	EARLY CLASSIC
108	3	INSTRUMENTA	111	3	LATE MEDIEVAL
108	4	GERMAN ART SO	111	4	ROMANTIC LIE
108	5	SYMPHONY OR	111	5	BAROQUE CONC
108	6	GUITAR - SPAN	111	6	BAROQUE
108	7	NATIONALISTI	111	7	MID/LATE ROM
108	8	BAROQUE ORGA	111	8	LATE BAROQUE
108	9	MELISMATIC/	111	9	MEDIEVAL
108	10	ROMANTIC PIA	111	10	MID ROMANTI
108	11	RENAISSANCE -	111	11	BAROQUE
108	12	LATE MIDDLE A	111	12	LATE MEDIEVAL
109	1	CHORAL	117	1	CLASSICAL CHO
109	2	ITALIAN OPERA	117	2	BAROQUE VOCA
109	3	ENGLISH DANC	117	3	RENAISSANCE/
109	4	GERMAN LIEDE	117	4	ROMANTIC W/
109	5	SYMPHONY - V	117	5	CLASSICAL ORC
109	6	XX	117	6	CLASSICAL GUI
109	7	PIANO & STRIN	117	7	PIANO CONCER
109	8	ORGAN FANTAS	117	8	BAROQUE ORGA
109	9	EARLY LATIN M	117	9	MELISMATIC O
109	10	PIANO SONATA	117	10	ROMANTIC PIA
109	11	CHORAL FUGUE	117	11	CHOIR
109	12	15TH CENT CHA	117	12	MEDIEVAL MOT

SUBJ	EXAMPLE	LABEL	SUBJ	EXAMPLE	LABEL
118	1	A CAPELLA CHO	124	1	BRUCKNER VIR
118	2	SOPRANO W/IN	124	2	ITALIAN ORATO
118	3	RENAISSANCE	124	3	RENAISSANCE C
118	4	GERMAN ROMA	124	4	MARIAN ANDER
118	5	ORCH, CLASSIC	124	5	VIVALDII OR C
118	6	SOLO GUITAR FA	124	6	EARLY SOLO PI
118	7	PNO TRIO, ROM	124	7	STRING QUARTE
118	8	ORGAN MUSIC.	124	8	ORGAN/BAROQ
118	9	"KYRIE ELEISO	124	9	CATHOLIC MAS
118	10	PIANO, ROMAN	124	10	ROMANTIC SOL
118	11	CHORUS IMITA	124	11	EARLY CHORAL
118	12	FR TROUVERE SO	124	12	FRENCH - EARL
119	1	CHORAL	125	1	XX
119	2	OPERA ARIA - E	125	2	XX
119	3	EARLY MEDIEV	125	3	CONSORT /REN
119	4	ART SONG	125	4	XX
119	5	STRING QUARET	125	5	BAROQUE/ORCH
119	6	EARLY MUSIC O	125	6	PIANO/QUINTE
119	7	STRING/PIANO	125	7	ORGAN/BAROQ
119	8	ORGAN - CHURC	125	8	XX
119	9	CHANT STYLE C	125	9	PIANO/ROMAN
119	10	ROMANTIC PER	125	10	XX
119	11	CHORAL, CANO	125	11	XX
119	12	A TROUBADOUR	125	12	XX
120	1	CHORAL, SACRE	126	1	CHORAL
120	2	L'ORFEO, GLUCK	126	2	BAROQUE
120	3	CHAMBER MUS	126	3	BAROQUE
120	4	~~GERMAN~~ FRENC	126	4	VOCAL SOLO
120	5	SYMPHONIC, C	126	5	CONCERTO GRO
120	6	GUITAR OR LUT	126	6	XX
120	7	SYMPHONIC, P	126	7	QUNTET
120	8	BUXTEHUDE, OR	126	8	BAROQUE
120	9	MALE CHORUS,	126	9	CHORAL
120	10	CHOPIN, RUBAT	126	10	PIANO SONATA
120	11	CHORAL, SACRE	126	11	BAROQUE VOCA
120	12	CHANSON, REN	126	12	FRENCH IMPRO

SUBJ	EXAMPLE	LABEL
128	1	FULL CHOIR, BA
128	2	BAROQUE (?) AR
128	3	REN. DANCE MU
128	4	(ALTO) SONATA
128	5	BAROQUE CONC
128	6	GUITAR SONATA
128	7	STRING QUARTE
128	8	ORGAN SONATA
128	9	MEN'S CHOIR -
128	10	PIANO SONATA
128	11	BAROQUE CANT
128	12	RENAISSANCE D
129	1	CHORAL
129	2	BAROQUE ARIA
129	3	RENAISSANCE -
129	4	LIEDER
129	5	BAROQUE CONC
129	6	CLASSICAL GUI
129	7	ROMANTIC - PI
129	8	BAROQUE - ORG
129	9	GREGORIAN CH
129	10	CLASSICAL - PI
129	11	XX
129	12	XX
130	1	RENAISSANCE
130	2	BAROQUE OPER
130	3	RENAISSANCE,
130	4	ROMANTIC ~~LIE~~
130	5	BAROQUE, CONC
130	6	GUITAR/LUTE
130	7	ROMANTIC PIA
130	8	BAROQUE , ORG
130	9	MEDIEVAL MOT
130	10	PIANO, ROMAN
130	11	MOTET
130	12	MEDIEVAL CHA

SUBJ	EXAMPLE	LABEL
131	1	CLASSICAL - CH
131	2	BAROQUE ARIA
131	3	EARLY BAROQU
131	4	ROMANTIC LIE
131	5	BAROQUE CONC
131	6	RENAISSANCE C
131	7	ROMANTIC TRI
131	8	BAROQUE ORGA
131	9	EARLY RENAISS
131	10	CLASSICAL PIA
131	11	LATE RENAISSA
131	12	RENAISSANCE F

SUBJ	EXAMPLE	LABEL	SUBJ	EXAMPLE	LABEL
102	1	MONOPHONIC	105	1	CHORAL OCTAV
102	2	SOP. ARIA WIT	105	2	BAROQUE (ITAL
102	3	X	105	3	RENAISSANCE D
102	4	FRENCH ART SO	105	4	GERMAN ART SO
102	5	ORCH - STRING	105	5	BAROQUE CHAM
102	6	BAROQUE SUITE	105	6	CLASSICAL GUI
102	7	MARCH - STRIN	105	7	TRASHY CLASSI
102	8	ORGAN - POLYP	105	8	ORGAN PRELUD
102	9	CHORAL POLYP	105	9	MEDIEVAL/EAR
102	10	SOLO PIANO ET	105	10	CHOPIN ~~PRELU~~
102	11	CHORAL MONOP	105	11	CHORAL - RENA
102	12	CELTIC SONG W	105	12	RENAISSANCE B
103	1	A CAPELLA CHO	106	1	LATE BAROQUE
103	2	ITALIAN - BAR	106	2	EARLY BAROQU
103	3	EARLY INSTRU	106	3	RENAISSANCE E
103	4	SOP. SOLO GERM	106	4	LATE ROMANTI
103	5	BAROQUE CONC	106	5	CLASSICAL (OR
103	6	GUITAR ARRAN	106	6	BAROQUE GUITA
103	7	XX	106	7	CLASSICAL QUI
103	8	ORGAN SOLO CO	106	8	LATE BAROQUE
103	9	ORGANUM	106	9	EARLY BAROQU
103	10	SOLO PIANO - M	106	10	ROMANTIC PIA
103	11	A CAPELLA - MA	106	11	LATE BAROQUE
103	12	ITALIAN	106	12	RENAISSANCE C
104	1	SACRED MOTET	107	1	SACRED CHORA
104	2	EARLY BAROQU	107	2	EARLY OPERA M
104	3	RENAISSANCE	107	3	RENAISSANCE I
104	4	LIEDER	107	4	ROMANTIC VOC
104	5	VIVALDI CONC	107	5	BAROQUE ORCH
104	6	LUTE - RENAISS	107	6	RENAISSANCE G
104	7	BRAHMS QUINT	107	7	ROMANTIC ERA
104	8	BACH ORGAN FU	107	8	BAROQUE ORGA
104	9	MEDIEVAL	107	9	MEDIEVAL CHU
104	10	CHOPIN NOCTU	107	10	ROMANTIC ERA
104	11	BAROQUE CHOR	107	11	RENAISSANCE S
104	12	RENAISSANCE	107	12	EARLY RENAISS

SUBJ	EXAMPLE	LABEL
112	1	CONCERT CHOR
112	2	BAROQUE ITAL.
112	3	REN/BAR LUTE
112	4	ART SONG - FRE
112	5	SYMPHONIC -
112	6	BAR/REN. LUTE
112	7	STRING QT & PN
112	8	BAROQUE ORGA
112	9	EARLY REN CHO
112	10	PIANO SOLO - L
112	11	CHORAL - GENE
112	12	14TH C. - ARS
113	1	BAROQUE CHOR
113	2	BAROQUE ARIA
113	3	RENAISSANCE C
113	4	GERMAN LIED (
113	5	BAROQUE CONC
113	6	CLASSICAL GUI
113	7	CLASSICAL PNO
113	8	BAROQUE ORGA
113	9	RENAISSANCE C
113	10	CHOPIN PIANO
113	11	BAROQUE CHOR
113	12	FRENCH RENAI
114	1	HOMOPHONY
114	2	ARIA (CLASSIC
114	3	IMITATIVE POL
114	4	ARIA-LIKE (W/
114	5	CLASSICAL (FRO
114	6	SPANISH GUITA
114	7	FRENCH JAZZ
114	8	EARLY ORGAN
114	9	GREGORIAN CH
114	10	ROMANTIC (CH
114	11	POLYPHONY (I
114	12	MELISMATIC P

SUBJ	EXAMPLE	LABEL
115	1	BAROQUE
115	2	CLASSICAL VOC
115	3	RENAISSANCE
115	4	ROMANTIC VOC
115	5	CLASSICAL
115	6	CLASSICAL GUI
115	7	CLASSICAL
115	8	BAROQUE
115	9	RENAISSANCE
115	10	ROMANTIC
115	11	RENAISSANCE
115	12	RENAISSANCE
116	1	CHORAL, A MAS
116	2	ARIA FROM BAR
116	3	BAROQUE
116	4	LIEDER (A GOOD
116	5	BAROQUE
116	6	GUITAR INSTRU
116	7	STRING QUARTE
116	8	ORGAN - BACH?
116	9	CHORAL, MEN'S
116	10	AH! CHOPIN, P
116	11	CHORAL, A CAP
116	12	BAROQUE? VOC
121	1	CHORUS MASS -
121	2	VOCAL - OPERA
121	3	BAROQUE
121	4	VOCAL SOLO W/
121	5	CLASSICAL SYM
121	6	CLASICAL SOLO
121	7	STRING QUARTE
121	8	BAROQUE FUGU
121	9	MASS/RENAISS
121	10	OMANTIC PIAN
121	11	CHORUS/CLASS
121	12	VOCAL/OPERA

SUBJ	EXAMPLE	LABEL
122	1	BAROQUE CHOR
122	2	OPERA
122	3	MEDIEVAL (LAT
122	4	FRENCH CABAR
122	5	BAROQUE ORCH
122	6	RENAISSANCE L
122	7	ORCHESTRAL
122	8	VERY EARLY OR
122	9	GREGORIAN CH
122	10	ROMANTIC
122	11	CHORAL
122	12	MEDIEVAL
123	1	ORATORIO/ CHU
123	2	BAROQUE ITALI
123	3	RENAISSANCE I
123	4	FRENCH SONG
123	5	CLASSICAL SYM
123	6	SPANISH GUITA
123	7	PIANO QUINTE
123	8	BAROQUE ORGA
123	9	CHANT
123	10	PIANO IMPRO
123	11	ORATORIO/ANT
123	12	FRENCH MEDIE
127	1	MEDIEVAL CHA
127	2	EARLY BAROQU
127	3	MEDIEVAL CON
127	4	~~GERMAN~~ FRENC
127	5	CONCERTO GRO
127	6	MEDIEVAL LUTE
127	7	PIANO QUINTE
127	8	MEDIEVAL ORG
127	9	MEDIEVAL CHA
127	10	CHOPIN
127	11	MEDIEVAL CHO
127	12	MEDIEVAL SON

SUBJ	EXAMPLE	LABEL
132	1	CHORAL
132	2	ITALIAN ARIA-
132	3	RENAISSANCE-
132	4	GERMAN LIEDE
132	5	BAROQUE CONC
132	6	RENAISSANCE-
132	7	LATE CLASSICA
132	8	BAROQUE-ORGA
132	9	MASS
132	10	CHOPIN PIANO
132	11	MADRIGAL
132	12	TROUVERE SONG

SUBJ	EXAMPLE	LABEL	SUBJ	EXAMPLE	LABEL
201	1	OPERA CHORUS	206	1	CLASSICAL - W
201	2	DIGNIFIED, CH	206	2	OPERA
201	3	BAROQUE INSTR	206	3	THE KIND WE W
201	4	VOCAL, ALTO, G	206	4	GERMAN OPERA
201	5	CLASSICAL, VIO	206	5	18TH C. BAVAR
201	6	FLAMENCO PIE	206	6	SPANISH GUITA
201	7	DRAMATIC PAR	206	7	PIANO W/STRI
201	8	CHURCH "FILLE	206	8	CLASSICAL - BA
201	9	XX	206	9	MUSIC FROM A
201	10	WALTZ	206	10	SOFT, PIANO
201	11	CHRISTIAN CHO	206	11	CHOIR, CLASSI
201	12	ORIENTAL	206	12	BALLAD, IRISH
202	1	XX	207	1	CHOIR/CHORAL
202	2	OPERA	207	2	SOLO PIECE/CH
202	3	RENAISSANCE	207	3	SYMPHONY/FL
202	4	ACCOMPANIM	207	4	SOLO OPERA PI
202	5	CLASSICAL	207	5	SYMPHONY/BE
202	6	SPANISH GUITA	207	6	FOREIGN/SPAN
202	7	DRAMATIC	207	7	SYMPHONY
202	8	ORGAN	207	8	CHURCH/ORGA
202	9	MONASTIC CHO	207	9	MEN'S CHORAL
202	10	PIANO	207	10	SOLO PANO
202	11	CHOIR	207	11	CHORAL
202	12	MINSTREL	207	12	FRENCH SOLO
205	1	CHORAL - CLAS	213	1	CLASSICAL
205	2	ITALIAN CLASS	213	2	OPERA
205	3	MEDIEVAL INS	213	3	CHAMBER
205	4	GERMAN CLASS	213	4	OPERA
205	5	CLASSICAL ORC	213	5	CLASSICAL ORC
205	6	CLASSICAL GUI	213	6	CLASSICAL
205	7	CLASSICAL INS	213	7	~~CLASSICAL~~ - CH
205	8	CLASSICAL ORG	213	8	CHAMBER
205	9	MEDIEVAL VOC	213	9	OPERA
205	10	CLASSICAL PIA	213	0	CLASSICAL (PIA
205	11	MEDIEVAL CHO	213	11	CHIOR
205	12	FRENCH VOCAL	213	12	CHAMBER MUS

SUBJ	EXAMPLE	LABEL
216	1	CHORAL
216	2	OPERA
216	3	XX
216	4	OPERA
216	5	CLASSICAL
216	6	CLASSICAL GUI
216	7	CLASSICAL
216	8	CHAMBER MUS
216	9	CHANT
216	10	CONCERTO
216	11	OPERA
216	12	XX
218	1	MODERN CLASS
218	2	OPERA
218	3	BAROQUE - REN
218	4	MODERN
218	5	ROMANTIC CLA
218	6	MEDIEVAL
218	7	MODERN CLASS
218	8	BAROQUE
218	9	GREGORIAN
218	10	PIANO ROMAN
218	11	CLASSICAL
218	12	BAROQUE
222	1	CHURCH CHANT
222	2	MIDDLE AGES C
222	3	RENAISSANCE F
222	4	CLASSICAL OPE
222	5	CLASSICAL EUR
222	6	MIDDLE AGES L
222	7	CLASSICAL EUR
222	8	MIDDLE AGES C
222	9	GREGORIAN CH
222	10	CLASSICAL PIA
222	11	CHURCH CHANT
222	12	FOLK - MIDDLE

SUBJ	EXAMPLE	LABEL
223	1	CHOIR
223	2	OPERA
223	3	XX
223	4	GARBAGE
223	5	SYMPHONY
223	6	GUITAR SOLO
223	7	FUNERAL PROC
223	8	HAUNTED HOUS
223	9	CHURCH CHOIR
223	10	LOUNGE MUSIC
223	11	CHOIR
223	12	FOREIGN
224	1	CHOIR
224	2	OPERA
224	3	MIDDLE AGES F
224	4	OPERA
224	5	CLASSICAL
224	6	GUITAR FOLK
224	7	CLASSICAL
224	8	CHURCH ORGAN
224	9	CHURCH
224	10	CLASSICAL
224	11	CHOIR
224	12	XX
225	1	BALLET MUSIC
225	2	OPERA
225	3	ORCHESTRA OLD
225	4	FOREIGN
225	5	CLASSICAL
225	6	FOLKLORE
225	7	CLASSICAL
225	8	BIBLE
225	9	MONK CHOIR
225	10	SAD
225	11	CHOIR W/ ORCH
225	12	OPERA OR PLAY

SUBJ	EXAMPLE	LABEL
226	1	CHOIR
226	2	OPERA/BAROQU
226	3	BAROQUE
226	4	OPERA
226	5	CLASSICAL
226	6	CLASSICAL GUI
226	7	CHAMBER
226	8	CLASSICAL MED
226	9	MONK CHANTI
226	10	LOUNGE PIANO
226	11	CHOIR
226	12	ROYAL CLASSIC
227	1	~~SYMPHONIC~~ CH
227	2	OPERA
227	3	QUINTET
227	4	SOLO CONCERT
227	5	SYMPHONY OR
227	6	~~DUET~~ SOLO
227	7	ORCHESTRA
227	8	PARLOUR MUSI
227	9	OPERA
227	10	SOLO CONCERT
227	11	SYMPHONY OR
227	12	OPERA
228	1	CHOIR
228	2	OPERA
228	3	FOLK MUSIC
228	4	OPERA/ CHEESY
228	5	CLASSICAL / SY
228	6	CLASSICAL MU
228	7	XX
228	8	CARNIVALESQU
228	9	CHANTING
228	10	CLASSICAL PIA
228	11	CHOIR
228	12	OPERA/FOLK

SUBJ	EXAMPLE	LABEL
230	1	CHOIR
230	2	OPERA FEMA
230	3	CHAMBER M
230	4	FEMALE GER
230	5	CLASSICAL S
230	6	CLASSICAL G
230	7	TRIO CLASSI
230	8	BAROQUE OR
230	9	GREGORIAN
230	10	CLASSICAL P
230	11	CHOIR
230	12	TRIO CLASSIC

SUBJ	EXAMPLE	LABEL
203	1	QUIRE [ESL] MU
203	2	OPERA
203	3	(CLASSICAL MU
203	4	OPERA
203	5	CLASSICAL MU
203	6	CLASSIC GUITA
203	7	CLASSICAL MU
203	8	ORGAN IN A CH
203	9	QUIRE [ESL] IN
203	10	LIGHT MUSIC (P
203	11	QUIRE [ESL]
203	12	OPERA (FLUTE)
204	1	CLASSICAL
204	2	OPERA
204	3	MEDIEVAL CHA
204	4	CLASSICAL
204	5	BAROQUE CLASS
204	6	CLASSICAL GUI
204	7	BAROQUE CLASS
204	8	FUGUE
204	9	GREGORIAN CH
204	10	CLASSICAL
204	11	CLASSICAL
204	12	MEDIEVAL CHA
208	1	CHORAL - I ALW
208	2	ITALIAN OPERA
208	3	CHAMBER QUAR
208	4	GERMAN W/TH
208	5	17TH CENT/CLA
208	6	GUITAR, SATUR
208	7	MODERN - RETR
208	8	CHUCK EVANS A
208	9	MONKS BEMOA
208	10	NOCTURNE - ES
208	11	VISIONS OF PA
208	12	ITALIAN BEAUT
209	1	CHOIR
209	2	OPERA
209	3	CLASSICAL
209	4	OPERA
209	5	CLASSICAL
209	6	CITAR MUSIC
209	7	CLASSICAL
209	8	ORGAN MUSIC
209	9	CHOIR
209	10	CLASICAL PIAN
209	11	CHOIR MUSIC
209	12	OPERA
210	1	CLASSICAL CHO
210	2	OPERA
210	3	EARLY CLASSIC
210	4	OPERA
210	5	CLASSICAL ORC
210	6	CLASSICAL GUI
210	7	CLASSICAL
210	8	CLASSICAL/HA
210	9	CLASSICAL-BA
210	10	CLASSICAL
210	11	CLASSICAL/CH
210	12	OPERA
211	1	CHURCH
211	2	OPERA
211	3	CHAMBER
211	4	OPERA
211	5	CLASSICAL
211	6	FLAMENCO
211	7	CHAMBER
211	8	CHAMBER
211	9	GREGORIAN
211	10	CLASSICAL
211	11	CLASSICAL
211	12	CHAMBER

SUBJ	EXAMPLE	LABEL
212	1	CHOIR
212	2	CHOIR
212	3	FOLK
212	4	FOLK
212	5	JAZZ
212	6	FOLK
212	7	JAZZ
212	8	CHOIR
212	9	CHOIR
212	10	XX
212	11	CHOIR
212	12	CHOIR
214	1	CHORAL
214	2	OPERA
214	3	CLASSICAL
214	4	OPERA
214	5	SYMPHONY CL
214	6	CLASSICAL GUI
214	7	CLASSICAL
214	8	ORGAN HYMNA
214	9	CHOIR BOYS - G
214	0	PIANO CONCER
214	11	OPERA
214	12	CLASSICAL
215	1	CATHOLIC CHUR
215	2	OPRA (MAYBE I
215	3	EUROPEAN REN
215	4	GERMAN
215	5	CLASSICAL
215	6	SPANISH GUITA
215	7	VICTORIAN CLA
215	8	DARK AGES CAT
215	9	MONKS IN MOU
215	0	CLASSIC PIANO
215	11	OPRA - CLASSIC
215	12	EUROPEAN REN

SUBJ	EXAMPLE	LABEL
217	1	SOARING - MOS
217	2	OPERATIC, BUS
217	3	CLASSICAL, NI
217	4	NOT SOMETHIN
217	5	SYMPHONIC -
217	6	SPANISH - CLA
217	7	ROYAL - POMP
217	8	TINNY MUSIC
217	9	GREGORIAN CH
217	10	PIANO
217	11	UPLIFTING
217	12	A PERFORMANC
219	1	BIG CHOIR CLA
219	2	19TH CENTURY
219	3	COURT MUSIC
219	4	DEPRESSING W
219	5	CLASSICAL ORC
219	6	SEGOVIAN SPA
219	7	CLASSICALY CO
219	8	HAPPY PIPE OR
219	9	MONESTARY
219	10	ON GOLDEN CLA
219	11	CHEERY OPERA
219	12	ELECTRIC "SOU
220	1	GOSPEL/CHURC
220	2	OPRA
220	3	ORCHESTRA
220	4	MUSICAL PLAY
220	5	CLASSICAL
220	6	ITALIAN MUSI
220	7	SILENT MOVIE
220	8	CHURCH ORGAN
220	9	OLD RELIGIOUS
220	10	WALTZ TYPE M
220	11	OLD ENGLISH M
220	12	OPRA

SUBJ	EXAMPLE	LABEL
221	1	CHRISTMAS MU
221	2	OPERA
221	3	OLD MID-AGE T
221	4	OPERA
221	5	CLASSICAL
221	6	CLASSICAL
221	7	ORCHESTRA - C
221	8	MID. CENTURY
221	9	CHURCH MUSIC
221	10	CLASSICAL
221	11	FROM PLAYS?
221	12	OPERA
229	1	CHIOR
229	2	OPERA
229	3	ORCRASTRA
229	4	OPERA
229	5	SYMPHONY OR
229	6	SPANISH GUITA
229	7	ORCRASTRA (SP
229	8	ORGAN
229	9	CHOIR
229	10	PIANO
229	11	OPERA
229	12	OPERA
231	1	1850'S ORCHES
231	2	GERMAN OPRA
231	3	BAROQUE 1650
231	4	GERMAN CHURC
231	5	1850'S ORCHES
231	6	SPANISH GUITA
231	7	FULL ORCHESTR
231	8	PIPE, GERMAN
231	9	PASTORAL, 150
231	10	PIANO, 1700'S
231	11	CHOIR, GERMA
231	12	BAROQUE, ITAL

SUBJ	EXAMPLE	LABEL
232	1	RELIGIOUS
232	2	OPERA
232	3	CLASSICAL
232	4	OPERA
232	5	CLASSICAL (LOV
232	6	CLASSICAL GUI
232	7	CLASSICAL
232	8	RELIGIOUS
232	9	RELIGIOUS
232	10	CLASSICAL
232	11	RELIGIOUS
232	12	OPERA

APPENDIX B

RECORDED EXAMPLES

(Target examples indicated
with an asterisk*)

CTXT1A

Exmp #	NAWM #	Composer	*Title*	Vol/CD/Track	Timing
1	96	Bach	*Wenn wir*	I 7 19	0:10
2	69	Monteverdi	*Poppea*	I 5 8	0:10
3	140	Weber	*die Frieschutz*	II 3 38	0:40
4	42	Dufay	*Adieux*	I 3 44	0:10
5	88	Bach	*Nun Komm*	I 6 13	0:10
6	127	Liszt	*Mazeppa*	II 2 14	0:00
7	56	Weelkes	*O Care*	I 4 40	0:10
8*	143	Bruckner	*Virga Jesse*	II 4 45	0:10
9	100	Frescobaldi	*Toccata*	I 7 34	0:10
10	130	Mendelssohn	*Midsummer's ND*	II 2 26	0:10
11*	76	Scarlatti	*Griselda*	I 5 27	0:20
12	45b	Narvaez	*Mille Regretz*	I 3 48	0:10
13*	58	Attaignant	*Danseries*	I 4 44	0:10
14	91	Corelli	*Trio Sonata*	I 6 32	0:10
15	136	Mahler	*Kindertotenlied*	II 3 17	0:20
16	59	Morton	*L'homme arme*	I 4 46	0:10
17*	135	Wolf	*Kennst du*	II 3 9	0:20
18	74	Purcell	*Fairy Queen*	I 5 22	0:10
19	139	Meyerbeer	*Les Hugenots*	II 3 29	0:00
20*	92	Vivaldi	*Concerto Grosso*	I 7 2	0:10
21	50	Cara	*Io non compro*	I 4 12	0:10
22*	60	Milan	*Fantasia XI*	I 4 48	0:00
23	83	Carissimi	*Jephthe*	I 6 2	0:10
24	128	Liszt	*Nuages gris*	II 2 19	0:10
25	40	Palestrina	*Missa Papa Marcelli*	I 3 35	0:10
26*	132	Brahms	*Piano Quintet*	II 2 34	0:00
27	101	Gautier	*Gigue*	I 7 35	0:10
28	129	Berlioz	*Symph fantas*	II 2 20	0:20
29*	94	Buxtehude	*Danket*	I 7 15	0:10
30	98	Dowland	*Lachrymae Pavane*	I 7 24	0:10
31*	38	Taverner	*Missa Gloria*	I 3 29	0:10
32	73	Lully	*Armide*	I 5 15	0:10
33	133	Schubert	*Kennst du*	II 3 1	0:10
34	61	Tye	*In nomine*	I 4 53	0:10
35*	126	Chopin	*Nocturne*	II 2 12	0:10
36	78	Gay	*Beggar's Opera*	I 5 36	0:10
37	125	Field	*Nocturne*	II 2 11	0:10
38*	80	Gabrieli	*Hodie*	I 5 40	0:10
39	39	Arcadelt	*Missa Noe*	I 3 41	0:10
40*	41	Dufay	*Resveillies vous*	I 3 41	0:10

CTXT1B

Ex #	NAWM #	Composer	*Title*	CD/Track	Timing
1	39	Arcadelt	*Missa Noe*	I 3 41	0:10
2	125	Field	*Nocturne*	II 2 11	0:10
3	78	Gay	*Beggar's Opera*	I 5 36	0:10
4	61	Tye	*In nomine*	I 4 53	0:10
5	133	Schubert	*Kennst du*	II 3 1	0:10
6	73	Lully	*Armide*	I 5 15	0:10
7	98	Dowland	*Lachrymae Pavane*	I 7 24	0:10
8*	41	Dufay	*Resveillies vous*	I 3 41	0:10
9	129	Berlioz	*Symph fantas*	II 2 20	0:20
10	101	Gautier	*Gigue*	I 7 35	0:10
11*	80	Gabrieli	*Hodie*	I 5 40	0:10
12	40	Palestrina	*Missa Papa Marcelli*	I 3 35	0:10
13*	126	Chopin	*Nocturne*	II 2 12	0:10
14	128	Liszt	*Nuages gris*	II 2 19	0:10
15	83	Carissimi	*Jephthe*	I 6 2	0:10
16	50	Cara	*Io non compro*	I 4 12	0:10
17*	38	Taverner	*Missa Gloria*	I 3 29	0:10
18	139	Meyerbeer	*Les Hugenots*	II 3 29	0:00
19	74	Purcell	*Fairy Queen*	I 5 22	0:10
20*	94	Buxtehude	*Danket*	I 7 15	0:10
21	59	Morton	*L'homme arme*	I 4 46	0:10
22*	132	Brahms	*Piano Quintet*	II 2 34	0:00
23	136	Mahler	*Kindertotenlied*	II 3 17	0:20
24	91	Corelli	*Trio Sonata*	I 6 32	0:10
25	45b	Narvaez	*Mille Regretz*	I 3 48	0:10
26*	60	Milan	*Fantasia XI*	I 4 48	0:00
27	130	Mendelssohn	*Midsummer's ND*	II 2 26	0:10
28	100	Frescobaldi	*Toccata*	I 7 34	0:10
29*	92	Vivaldi	*Concerto Grosso*	I 7 2	0:10
30	56	Weelkes	*O Care*	I 4 40	0:10
31*	135	Wolf	*Kennst du*	II 3 9	0:20
32	127	Liszt	*Mazeppa*	II 2 14	0:00
33	88	Bach	*Nun komm*	I 6 13	0:10
34	42	Dufay	*Adieux*	I 3 44	0:10
35*	58	Attaignant	*Danseries*	I 4 44	0:10
36	140	Weber	*die Freischutz*	II 3 38	0:40
37	69	Monteverdi	*Poppea*	I 5 8	0:10
38*	76	Scarlatti	*Griselda*	I 5 27	0:20
39	96	Bach	*Wenn wir*	I 7 19	0:10
40*	143	Bruckner	*Virga Jesse*	II 4 45	0:10

CTXT2A

Ex #	NAWM #	Artist/ Composer	*Album / Title*	Vol/Track	Time
1		Wings	*Wings Greatest/ Junior's Farm*	4	:30
2		L. Johnson	*Blues Masters/ On the Wall*	VIII 5	:30
3		Dire Straits	*Money for Nothing/ Tunnel of Love*	5	:40
4		Can Heat	*Blues Masters/ On the Road Again*	VII 15	:40
5		L Skynard	*Gold & Platinum/ You Got That Right*	5	:20
6		JR Morton	*Classic Jazz/ King Porter Stomp*	I 9	:10
7		R Johnson	*Blues Masters/ Crossroad Blues*	VIII 8	:10
8*	143	Bruckner	*Virga Jesse*	II 4 45	0:10
9		Sara Vaughn	*Classic Jazz/ All Alone*	IV 4	:20
10		Band	*Best of/ Life Is a Carnival*	5	:30
11*	76	Scarlatti	*Griselda*	I 5 27	0:20
12		Howl Wolf	*Blues Masters/ Sittin on Top of the World*	VIII 13	:30
13*	58	Attaignant	*Danseries*	I 4 44	0:10
14		Paker	*Classic Jazz/ Embraceable You*	III 14	:30
15		Fitzgerald	*Classic Jazz/ You'd Be So Nice*	II 12	:10
16		Armstrong	*Classic Jazz/ Big Butter & Egg Man*	I 13	:10
17*	135	Wolf	*kennst du*	II 3 9	0:20
18		Taylor	*Blues Masters/ I'm a Woman*	IX 8	:30
19		Oliver	*Classic Jazz/ Dippermouth Blues*	I 5	:10
20*	92	Vivaldi	*Concerto Grosso*	I 7 2	0:10
21		Coltrane	*Classic Jazz/ Alabama*	V 6	:10
22*	60	Milan	*Fantasia XI*	I 4 48	0:00
23		Cray	*Blues Masters/ Phone Booth*	IX 10	:10

24		Elvis	*Top 18 Hits/ Ring Around Your Neck*	13	:10
25		Waters	*Blues Masters/ Got my mojo*	VII 8	:20
26*	132	Brahms	*Piano Quintet*	II 2 34	0:00
27		Mayall	*Blues Masters/ Death of JB Lenoir*	VII 14	:10
28		Henderson	*Classic Jazz/ Stampede*	II 1	:20
29*	94	Buxtehude	*Danket*	I 7 15	0:10
30		Domino	*Best of Fats/ Be my guest*	8	:20
31*	38	Taverner	*Missa Gloria*	I 3 29	0:10
32		L Richard	*18 Greatest Hits/ Heebie-Jeebies*	6	:10
33		Basie	*Classic Jazz/ Doggin Around*	II 20	:10
34		Allman	*Filmore East/ Hot'Lanta*	II 1	:30
35*	126	Chopin	*Nocturne*	II 2 12	0:10
36		Berry	*R&R Rarities/ Come on*	5	:10
37		Beatles	*Rubber Soul/ The Word*	6	:10
38*	80	Gabrieli	*Hodie*	I 5 40	0:10
39		St R Vaughn	*Blues Masters/ Pride & Joy*	IX 13	:30
40*	41	Dufay	*Resveillies vous*	I 3 41	0:10

CTXT2B

Ex #	NAWM #	Artist/ Composer	*Album / Title*	Vol/Track	Time
1		St R Vaughn	*Blues Masters/ Pride & Joy*	IX 13	:30
2		Beatles	*Rubber Soul/ The Word*	6	:10
3		Berry	*R&R Rarities/ Come on*	5	:10
4		Allman	*Filmore East/ Hot'Lanta*	II 1	:30
5		Basie	*Classic Jazz/ Doggin Around*	II 20	:10
6		L Richard	*18 Greatest Hits/ Heebie-Jeebies*	6	:10
7		Domino	*Best of Fats/ Be my guest*	8	:20
8*	41	Dufay	*Resveillies vous*	I 3 41	0:10
9		Henderson	*Classic Jazz/ Stampede*	II 1	:20
10		Mayall	*Blues Masters/ Death of JB Lenoir*	VII 14	:10
11*	80	Gabrielli	*Hodie*	I 5 40	0:10
12		Waters	*Blues Masters/ Got my mojo*	VII 8	:20
13*	126	Chopin	*Nocturne*	II 2 12	0:10
14		Elvis	*Top 18 Hits/ Ring around your neck*	13	:10
15		Cray	*Blues Masters/ Phone Booth*	IX 10	:10
16		Coltrane	*Classic Jazz/ Alabama*	V 6	:10
17*	38	Taverner	*Missa Gloria*	I 3 29	0:10
18		Oliver	*Classic Jazz/ Dippermouth Blues*	I 5	:10
19		Taylor	*Blues Masters/ I'm a Woman*	IX 8	:30
20*	94	Buxtehude	*Danket*	I 7 15	0:10
21		Armstrong	*Classic Jazz/ Big Butter & Egg Man*	I 13	:10
22*	132	Brahms	*Piano Quintet*	II 2 34	0:00
23		Fitzgerald	*Classic Jazz/ You'd be so nice*	II 12	:10
24		Parker	*Classic Jazz/ Embracable you*	III 14	:30

25		Howl Wolf	*Blues Masters/ Sittin on top of the world*	VIII 13	:30
26*	60	Milan	*Fantasia XI*	I 4 48	0:00
27		Band	*Best of/ Life is a carnival*	5	:30
28		Sara Vaughn	*Classic Jazz/ All Alone*	IV 4	:20
29*					
30		R Johnson	*Blues Masters/ Crossroad Blues*	VIII 8	:10
31*	135	Wolf	*kennst du*	II 3 9	0:20
32		JR Morton	*Classic Jazz/ King Porter Stomp*	I 9	:10
33		L Skynard	*Gold & Platinum/ You Got That Right*	5	:20
34		Can Heat	*Blues Masters/ On the road again*	VII 15	:40
35*	58	Attaignant	*Danseries*	I 4 44	0:10
36		Dire Straits	*Money for Nothing/ Tunnel of Love*	5	:40
37		L. Johnson	*Blues Masters/ On the Wall*	VIII 5	:30
38*	76	Scarlatti	*Griselda*	I 5 27	0:20
39		Wings	*Wings Greatest/ Junior's Farm*	4	:30
40*	143	Bruckner	*Virga Jesse*	II 4 45	0:10

APPENDIX C

SUBJECT RESPONSE FORMS

Subject Profile

Age _____ Gender (circle) M F

Major or Department: _____

Please circle: undergraduate graduate GTF faculty

Have you taken any music history courses in college? YES NO

Have you had any musical training while in college?

No Yes, college courses Yes, private lessons

If **yes**, were these courses:

Music history _____ Voice or instrument lessons _____

Choir, band, or orchestra _____

If you are a faculty member, how many years have you been teaching in your subject area? 1-3 more than 3

Is there any other music related experience or training that you might have had? (a brief answer is fine).

_____

_____

Instructions:

You will hear a series of musical examples lasting about 20 seconds each. Listen for the first few seconds, and decide what kind of music you think this selection represents. You decide what this question means and how you will answer it. You may change your mind about how to answer after you have heard a few examples. That is perfectly alright, but you must not go back and change earlier answers.

Write down your answer quickly, since the next example will start a few seconds after the end of the preceding one. You will hear the number each time so you will not get lost. If you cannot decide on an answer, mark an X and go on. Once you have begun, we cannot stop the tape. The whole procedure should take about 15-20 minutes.

This is not a "name that tune" test. You might recognize some of the music, but the question is: what kind of music is this.

Please do not ask any questions aloud if you are taking the study with other people, since your questions might influence their thinking. If something is unclear, raise a hand and point to the area that is confusing.

YOU decide what is meant by "kind of music," so do not ask for examples or explanations.

A reminder:

I am interested in your first decisions, so do not go back and change an answer. It is quite possible that you will change your mind as you hear more examples, but please do not go back.

WAIT TO TURN THE PAGE

Tape: 1A 1B 2A 2B

What kind of music is this?

1. _____

2. _____

3. _____

4. _____

5. _____

6. _____

7. _____

8. _____

9. _____

10. _____

11. _____

12. _____

13. _____

14. _____

15. _____

What kind of music is this?

16. _____

17. _____

18. _____

19. _____

20. _____

21. _____

22. _____

23. _____

24. _____

25. _____

26. _____

27. _____

28. _____

29. _____

30. _____

What kind of music is this?

31. _____

32. _____

33. _____

34. _____

35. _____

36. _____

37. _____

38. _____

39. _____

40. _____

Thank you for participating in this study. I would appreciate your not discussing it with anyone who might be a future subject, since your answers might influence theirs. If you have questions about the nature of the study or the musical selections used, I will be happy to provide you with the information. Thank you once again.

BIBLIOGRAPHY

American Psychological Association. (1989). Publication Manual (3rd ed.). Washington, D.C.: American Psychological Association.

Anderson, J. R. (1990). Cognitive psychology and its implications (3rd ed.). New York: W.H. Freeman.

Aristotle (1991). Categories. In R. E. Allen (Ed. and Trans.) Greek philosophy: Thales to Aristotle (3rd ed., pp. 285-291)

Armstrong, S. L., Gleitman, L., & Gleitman, H. (1983). What some concepts might not be. Cognition, 13, 263-308.

Ausubel, D. (1968). Educational psychology: A Cognitive view. New York: Holt, Reinhart & Winston.

Barrett, J. R. (1989). Core thinking skills in music. In E. Boardman, Ed. Dimensions of musical thinking. (pp. 45-56). Reston, VA MENC.

Barsalou, L. (1983). Ad hoc categories. Memory and Cognition, 11, 211-227.

Barsalou, L. (1989). The instability of graded structure: Implications for the nature of concepts. In U. Neisser (Ed.), Concepts and conceptual development: Ecological and intellectual factors in categorization (pp. 101-141). Cambridge, UK: Cambridge University Press.

Bartlett, F. C. (1932). Remembering: A Study in experimental and social psychology. London: Cambridge University Press.

Blues Masters. (1993). Rhino Records R2 71130.

Booth,G. D., & Cutietta, R. (1991). The applicability of verbal processing strategies to recall of familiar songs. Journal of Research in Music Education, 39, 121-131.

Borg, W., & Gall, M. (1989). Educational Research (3rd ed.). Longman, New York.

Bradley, J. (1990). The role of cognitive structure in expertise. Unpublished doctoral dissertation, University of Texas, Arlington.

Brittin, R. V. (1991). The effect of overtly categorizing music on preference for popular music styles. Journal of Research in Music Education 39, 143-151.

Brooks, L. R. (1989) Decentralized control of categorization: The role of prior processing episodes. In U. Neisser, (Ed.), Concepts and conceptual development: Ecological and intellectual factors in categorization (pp. 141-174). Cambridge, UK: Cambridge University Press.

Brown, H. M. (1976). Music in the Renaissance. Englewood Cliffs, NJ: Prentice Hall.

Bruner, J. (1973). Beyond the information given: Studies in the psychology of knowing. New York: W. W. Norton.

Bruner, J. S., Goodnow, J. J., & Austin, G. A. (1957). A study of thinking. New York: Wiley & Sons.

Chi, M., Feltovich, P., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. Cognitive Science, 5, 121-152.

Cole, M., & Means, B. (1981). Comparative studies of how people think. Cambridge, MA: Harvard University Press.

Conrad, C. E. (1971). An analysis of a hierarchical model of semantic memory organization. Unpublished masters thesis, University of Oregon, Eugene.

Corbett, A. T. (1975). Artificial scattering and semantic distance in a categorization task. Unpublished masters thesis, University of Oregon, Eugene.

Corrigan, J. (1990). Effects of expertise on category knowledge: changes in basic level. Unpublished doctoral dissertation, University of Delaware, Newark.

Cutietta, R. (1993). Musical elements: Who said they're right? Music Educators Jounal, 79, 48-53.

Deutsch, D. (Ed.). (1982). The psychology of music. New York: Academic Press.

Diekhoff, G. M. (1983). Testing through relationship judgments. Journal of Educational Psychology, 75, 227-233.

Donington, R. (1973). A performer's guide to baroque music. London: Novello.

Dowling, W. J., & Harwood, D. L. (1986). Music cognition. San Diego, CA: Academic Press.

Eastlund, J. (1990). Musical style discrimination as a function of selected cognitive style variables. Contributions to Music Education 17, 50-66.

Eastlund, J. (1992). A Multidimensional scaling analysis of musical style. Journal of Research in Music Education, 40, 204-215.

Edwards, R. (1988). Transfer and performance instruction. In C. Fowler, (Ed.). The Crane Symposium: Toward an understanding of the teaching and learning of music performance. (pp. 119-142). Potsdam, NY: Crane School of Music

Einstein, A. (1947). Music in the romantic era. New York: Norton.

First national assessment of musical progress. (1974). Report No. 03-MU-40. Denver, CO: National Assessment of Educational Progress, Education Commission of the United States.

Fowler, C. (Ed.). (1988). The Crane Symposium: Toward an understanding of the teaching and learning of music performance. Potsdam, NY: Crane School of Music

Frederick, W. C., & Klausmeier, H. J. (1965). The effects of instructions, concept complexity, method of presentation, and order of concepts upon a concept attainment task. (Report No. BR-5-0216-TR-3). Madison, WI: Wisconsin University Research and Development Center for Learning and Re-Education. (ERIC Document Reproduction Service No. ED 010-203)

Gagne, E. (1985). Cognitive psychology of school learning. Boston: Little, Brown, & Co.

Gahart, M. (1987). Toward an understanding of social schemata: expertise, level of categorization, and schema use. Unpublished doctoral dissertation, University of California, Los Angeles.

Gardner, H. (1973). Children's sensitivity to musical styles. Merrill-Palmer Quarterly 19, 67-77.

Gardner, H. (1985). The mind's new science: A History of the cognitive revolution. New York: Basic Books.

Garnham, A. (1985). Psycholinguistics: Central topics. New York: Methuen.

Givon, T. (1989). Mind, code, and context. Hillsdale, NJ: Lawrence Earlbaum Associates.

Gregory, R. (Ed.). (1989). Oxford companion to the mind. Oxford, UK: Oxford University Press.

Grout, D. (1973). A History of western music. New York: Norton.

Haack, P. (1982). A study of high school music participants' stylistic preferences and identification abilities in music and the visual arts. Journal of Research in Music Education, 30, 213-220.

Harnad, S. (1990). Category induction and representation. In S. Harnad (Ed.), Categorical perception: The Groundwork of cognition (pp. 535-555). Cambridge, UK: Cambridge University Press.

Harnad, S. (Ed.). (1990). Categorical perception: The groundwork of cognition. Cambridge, UK: Cambridge university press.

Hintzman, D. (1986). Schema abstraction in a multiple-trace memory model. Psychological Review 93, 411-428.

Homa, D., Rhoads, D., & Chambliss, D. (1979). The evolution of conceptual structure. Journal of Experimental Psychology: Human Learning and Memory, 5, 11-23.

Houston, W. B. (Ed.). (1990). Handbook of research on teacher education. New York: MacMillan.

Howell, P., Cross, I., & West, R. (1985). Musical structure and cognition. London: Academic Press.

Iuele, P. (1991). Representation and feedback in the formation of physical science concepts. Unpublished doctoral dissertation, Rutgers University, New Brunswick, NJ.

Jonassen, D. H. (1988, January). Prototype development and acquisition in an interactive video concept lesson. Paper presented at the annual meeting of the Association for Educational Communications and Technology, New Orleans, LA. (ERIC Document Reproduction Service No. ED 295-650)

Kant, I. (1929). Critique of pure reason (N. K. Smith, trans.). New York: St. Martin's Press. (Original work published 1781).

Katsuyama, R. M., & Hoffarth, G.D. (1978). Salience and developmental effects upon the learning of rules. Biennial Southeastern Conference on Human Development, Atlanta, GA. (ERIC Document Reproduction Service No. ED 165-889)

Keil, F. (1989) Conceptual development and category structure. In U. Neisser (Ed.), Concepts and conceptual development: Ecological and intellectual factors in categorization (pp.175-200). Cambridge, UK: Cambridge University Press.

Keil, F., & Kelly, M. (1990). Developmental changes in category structure. In S. Harnad (Ed.), Categorical perception: The groundwork of cognition (pp. 491-510). Cambridge, UK: Cambridge University Press.

Kellogg, W. A. (1985). Knowledge and instantiation of meaning in reading: On-line studies of expertise. Unpublished doctoral dissertation, University of Oregon, Eugene.

Kennedy, M. (1987). Inexact sciences: Professional education and the development of expertise. In E. Rothkopf, (Ed.), Review of research in education (Vol. 14). Washington, D.C.: American Educational Research Association.

Lakoff, G. (1987). Women, fire, and dangerous things: What categories reveal about the mind. Chicago: University of Chicago Press.

Lakoff, G. (1989) Cognitve models and prototype theory. In U. Neisser (Ed.), Concepts and conceptual development: Ecological and intellectual factors in categorization (pp. 63-100). Cambridge, UK: Cambridge University Press.

Langer, S. K. (1975). Philosophy in a new key: A study in the symbolism of reason, rite, and art. Cambridge, MA: Harvard University Press.

Lao Tzu (1973). Tao Te Ching (Ch'u Ta-Kao, trans.). New York: Samuel Weiser.

Laughlin, P. R. (1973). Focusing strategy in concept attainment as a function of instructions and task complexity. Journal of Experimental Psychology, 98, 320-327.

Lurigio, A. J. (1983). Schematic approach to understanding probation officers' judgments of criminals. Unpublished doctoral dissertation, Loyola University, Chicago.

McCauley, R. (1989). The role of theories in a theory of concepts. In U. Neisser (Ed.), Concepts and conceptual development: Ecological and intellectual factors in categorization (pp. 288-310). Cambridge, UK: Cambridge University Press.

McKinney, C. W. (1987). Effects of best examples, critical attributes, definitions, and practice on concept acquisition and prototype formation. Journal of Social Studies research, 11, 1-14.

Medin, D., Altom, M., Edelson, S., & Freko, D. (1982). Correlated symptoms and simulated medical classification. Journal of Experimental Psychology: Learning, Memory, and Cognition, 8, 37-50.

Medin, D., & Barsalou, L. (1990). Categorization process and categorical perception. In S. Harnad (Ed.), Categorical perception: The groundwork of cognition (pp. 455-490). Cambridge, UK: Cambridge University Press.

Medin, D., & Wattenmaker, W. (1989). Category cohesiveness, theories, and cognitive archeology. In U. Neisser (Ed.), Concepts and conceptual development: Ecological and intellectual factors in categorization (pp. 25-62). Cambridge, UK: Cambridge University Press.

Merrill, M. D., & Tennyson, R. D. (1977). Teaching concepts: An instructional design guide. Englewood Cliffs, NJ: Educational Technology Publications.

Merrill, M. D., & Tennyson, R. D. (1978). Concept classification and classification errors as a function of relationships between examples and nonexamples. Improving Human Performance Quarterly, 7, 351-364.

Murphy, G., & Wright, J. (1984). Changes in conceptual structure with expertise: differences in real-world experts and novices. Journal of Experimental Psychology: Learning, Memory and Cognition, 10, 144-155.

Neisser, U. (1989). From direct perception to conceptual structure. In U. Neisser (Ed.), Concepts and conceptual development: Ecological and intellectual factors in categorization (pp. 1-10). Cambridge, UK: Cambridge University Press.

Neisser, U. (Ed.). (1989). Concepts and conceptual development: Ecological and intellectual factors in categorization. Cambridge, UK: Cambridge University Press.

Newell, A. (1973). Production systems: Models of control structures. New York: Academic press.

Norton Anthology of Western Music. (1988). (Compact Disk No. A5 20523). New York: Norton.

Palisca, C. (1981). Baroque music. Englewood Cliffs, NJ: Prentice-Hall.

Parker, I. V. (1990). The reliability, sensitivity, and criterion-related validity of concept comparisons and concept mapping for assessing reading comprehension. Unpublished doctoral dissertation, University of Oregon, Eugene.

Plato (1991). Parmenides. In R. E. Allen (Ed. and Trans.) Greek philosophy: Thales to Aristotle (3rd ed. p. 257-269). New York: Macmillan.

Reese, G. (1940). Music in the middle ages. NY: Norton

Reimer, B. (1994, July). A profile of the 1992 NSSE Yearbook: The Arts, education, and aesthetic knowing. Paper presented at the meeting of the International Society of Music Educators, Tampa, FL.

Rosch, E. (1973). Natural categories. Cognitive Psychology, 4, 328-350.

Rosch, E. H., & Lloyd, B. (1978). Cognition and categorization. Hillsdale, NJ: Earlbaum & Assoc.

Rosch, E. H., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. Cognitive psychology, 8, 382-439.

Rosch, E. H., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. Cognitive psychology, 8, 382-439.

Rosch, E. H., Mervis, C. B., Gray, W., Johnson, D., & Boyes-Braem, P. (1975). Basic objects in natural categories. (Working paper no. 43). Berkeley, CA: California University, Language Behavior Research Laboratory. (ERIC Document Reproduction Service No. ED 142-051)

Rothkopf, E. (1987). Review of research in education (Vols. 14-15). Washington, D.C.: American Educational Research Association.

Russell, B. (1972). A history of western philosophy. New York: Simon & Schuster.

Schmidt, H., & Boshuizen, H. (1990, April). Effects of activation of prior knowledge on the recall of a clinical case. Paper presented at the annual conference of the sociation, Boston, MA.

Seaton, D. (1991). Ideas and styles in western musical tradition. Mt. View, CA: Mayfield.

Silverman, D. (1990). The effect of category structure on judgments of similarity. Unpublished doctoral dissertation, University of New Mexico, Santa Fe.

Skinner, B. F. (1957). Verbal behavior. New York: Appleton-Century-Crofts.

Stepich, D. (1991). From novice to expert: Implications for instructional design. Performance and Instruction, 30, 13-17.

Steve, M. H., & Tennyson, R. D. (1974, April). Prior memorization of definitions, examples and nonexamples with conjunctive and disjunctive concept learning tasks. Paper presented at the annual meeting of the American Educational Research Association, Chicago. (ERIC Document Reproduction Service No. ED 091-430)

Taebel, D. K. (1971). The effect of various instructional modes on children's performance of music concept tasks. (From Dissertation Abstracts International, 33, 6957A)

Tait, M. and Haack, P. (1984). Principles and processes of music education: New perspectives. New York: Teachers College Press.

Tanaka, J. (1989). Object categorization and expertise. Unpublished doctoral dissertation, University of Oregon, Eugene.

Tanaka, J. & Taylor, M. (1991). Object categories and expertise: Is the basic level in the eye of the beholder? Cognitive Psychology 23, 457-482.

Taylor, J. (1989). Linguistic categorization: Prototypes in linguistic theory. Oxford UK: Clarendon.

Tennyson, R. D. (1973). The effect of negative instances in concept acquisition using a verbal learning task. Journal of Educational Psychology, 64, 247-260.

Trabasso, T. R. (1963). Stimulus emphasis and all-or-none learning in concept identification. Journal of Experimental Psychology, 65, 398-406.

Tversky, A., & Gati, I. (1978). Studies of similarity. In E. H. Rosch, & B. Lloyd, Cognition and categorization. Hillsdale, NJ: Earlbaum.

Tversky, A., & Kahneman, D. (1973). Availability: A Heuristic for judging frequency and probability. Cognitive Psychology, 5, 207-232.

Vessey, I. (1988). Expert-novice knowledge organization: An empirical investigation using computer program recall. Behavior and Information Technology, 7, 153-171.

Vickers, J. (1986). Expert-novice differences in knowledge structures of action. Ottowa: National Library of Canada.

Whorf, B. (1956). Language, Thought, and Reality: Selected writings of Benjamin Lee Whorf, ed. J. B. Carroll. Cambridge, MA: MIT Press.

Wittgenstein, L. (1958). Philosophical investigations. New York: Macmillan.