

THE EFFECT OF INDIVIDUAL VERSUS COLLECTIVE CREATIVE PROBLEM
SOLVING EXPERIENCES ON FOURTH- AND FIFTH-GRADE STUDENTS'

COMPOSITIONAL PRODUCTS

Beatriz E. Aguilar, B.M., M.M.E.

Dissertation Prepared for the Degree of
DOCTOR OF PHILOSOPHY

UNIVERSITY OF NORTH TEXAS

December 2004

APPROVED:

Debbie Rohwer, Major Professor and
Coordinator of the Ph.D. Program in Music
Education

Dahryl Ramsey, Minor Professor

Robin Henson, Committee Member

Warren Henry, Committee Member and

Chair of the Music Education Department

James C. Scott, Dean of the College of Music

Sandra L. Terrell, Dean of the Robert

Toulouse School of Graduate Studies

Aguilar, Beatriz E., The effect of individual versus collective creative problem solving experiences on fourth- and fifth-grade students' compositional products. Doctor of Philosophy (Music Education), December 2004, 208 pp., 27 tables, 7 figures, references, 76 titles.

The purpose of the study was to explore the effects that individual vs. collective structured creative musical problem solving tasks had on students' compositional products. Subjects in a convenience sample of 32 fourth-graders and 32 fifth-graders were randomly assigned to either the individual or collective condition. The 3 treatment sessions were characterized by an open-ended creative problem solving task, which included questions intended to guide subjects through 3 stages of the creative problem solving process: *Understanding the Problem*, *Generating Ideas*, and *Planning for Action*. Subjects participated in the pre- and posttest individually. Three experienced music educators assessed the compositional products in terms of pattern use, cohesiveness, and creativity.

The originally intended MANCOVAs could not be carried out because the data did not meet the necessary assumptions. Pretest and posttest scores were explored with individual ANOVAs. The Bonferroni technique was used to adjust the alpha level. The statistical analyses showed that subjects exposed to the individual condition obtained higher scores than subjects exposed to the collective condition on six of the eight explored subtests, but these differences were not significant. The level of interjudge reliability decreased at each of the three measurements of the study: pilot test, pretest, and posttest.

The study's results suggest that music educators interested in observing specific characteristics of individual students' compositional products, such as the levels of cohesiveness, creativity, and pattern use, could do so regardless of the condition under which students were exposed to compositional tasks, either individually or collectively. Recommendations for future research include the use of a measurement instrument specifically designed for open-ended tasks, and the exploration of the current study's measurement instrument with closed-ended tasks. The study highlights the need for appropriate measurement instruments designed for the compositional tasks at hand, and the need for research results reported clearly, so that more advancement of this field is possible.

Copyright 2004

by

Beatriz E.Aguilar

ACKNOWLEDGMENTS

My most sincere thanks go to the following people, without whom this research project would not have been possible: the National Autonomous University of Mexico (Universidad Nacional Autónoma de México) for funding my graduate studies; the students who participated in the project, for their willingness to share their music with me; the teachers Jennifer Florez, Laura Kinnaird, Cecile Johnson, Julie Scott, Melissa Hartshorn, and the Principal Teresa Andress, for offering their time, effort, and access to their students and facilities; to Dr. Dahryl Ramsey and Dr. Henry for sharing their insights and serving as members of my committee, and for allowing me to adopt them as members of my family while far away from home; to Dr. Robin Henson for his help with the statistical analyses, and for serving as member of my committee; to Dr. Hildegard Froehlich for sharing her wisdom with me, and providing the original seed for this project; and special thanks to Dr. Debbie Rohwer, for all her support and wisdom throughout my graduate studies, and for serving as my major advisor, giving me her guidance and excellent feedback throughout all of the stages of this project.

I want to thank my parents, Miguel and Lucy, for setting an example of dedication and discipline, and their endless support as I worked on this project; to my siblings Miguel, Alejandro, and Lucy for their support and love; to Miss Hudnall and Paula Rives for all their help and friendship; to my friends at North Texas, specially Nora Ceja, the Sequeras, the Kinnairds, the Renguilfos, the Morales, Balam Vazquez, Cliff, and Dr. Tyler-Wood for their endless help; to my husband Rodrigo and my son Rodriguito, whose constant love, patience, and support gave me the courage to take up and to finish a project of this magnitude. I could not have done it without all of you. Thank you!

TABLE OF CONTENTS

	Page
ACKNOWLEDGMENTS.....	iii
LIST OF TABLES.....	vi
LIST OF FIGURES.....	viii
 Chapters	
I. INTRODUCTION	1
Background.....	1
Purpose of the Study.....	10
Limitations of the Study	11
Assumptions	12
Definition of Terms	12
II. REVIEW OF LITERATURE	15
Research on Collective versus Individual Creative Problem Solving	15
Research on Collective versus Individual Creative Problem Solving in Music.....	20
Research on Creative Problem Solving Strategies	28
The Exploration and Assessment of Musical Compositions.....	31
III. METHODOLOGY	51
Design.....	51
Sample Size	53
Host School.....	54
The Task.....	55
Independent Variable	56
Sampling Procedures.....	64
Pretest Procedures	66
Treatment Sessions 1, 2, and 3 Procedures	68
Posttest Procedures.....	69
Dependent Variable	70

	Statistical Procedures	74
	Validity	77
	Pilot Tests Procedures	78
	The Training of the Judges	82
	Reliability.....	86
IV.	RESULTS	88
	Interjudge Reliability.....	88
	Homogeneity of Variance.....	91
	Subtests	92
V.	SUMMARY, DISCUSSION, AND RECOMMENDATIONS.....	137
	Summary.....	137
	Discussion.....	146
	Recommendations for Music Educators.....	157
	Recommendation for Future Research	159
Appendices		
A.	CONSENT FOR PROJECT PARTICIPATION FORMS.....	163
B.	SAMPLING AND TREATMENT ASSIGNMENT	168
C.	TREATMENT DESIGN	172
D.	JUDGES' SCORING FORMS FOR MAIN STUDY	174
E.	CONSENT FOR PROJECT PARTICIPATION FORMS (PILOT TEST).....	179
F.	LETTERS OF APPROVAL	184
G.	JUDGES' TRAINING MATERIAL AND SCORING FORMS FOR PILOT TEST.....	187
H.	INTERJUDGE RELIABILITY.....	193
	REFERENCES.....	200

LIST OF TABLES

Table 1. Overall Interjudge Reliability	90
Table 2. Descriptive Statistics for the Use of Repeated Melodic Pattern Log	
Transformed Posttest Data	98
Table 3. Analysis of Variance for the Use of Repeated Melodic Pattern Log Transformed	
Posttest Data	99
Table 4. Descriptive Statistics for the Use of Developed Melodic Pattern Log	
Transformed Posttest Data	104
Table 5. Analysis of Variance for the Use of Developed Melodic Pattern Log	
Transformed Posttest Data	105
Table 6. Descriptive Statistics for the Use of Repeated Rhythmic Pattern Inverse	
Transformed Posttest Data	110
Table 7. Analysis of Variance for the Use of Repeated Rhythmic Pattern Inverse	
Transformed Posttest Data	111
Table 8. Descriptive Statistics for the Use of Developed Rhythmic Pattern Square Root	
Transformed Posttest Data	116
Table 9. Analysis of Variance for the Use of Developed Rhythmic Pattern Square Root	
Transformed Posttest Data	117
Table 10. Descriptive Statistics for the Tonal Cohesiveness Posttest Data	121
Table 11. Analysis of Variance for the Tonal Cohesiveness Posttest Data	122
Table 13. Analysis of Variance for the Metric Cohesiveness Posttest Data	136
Table 14. Descriptive Statistics for the Fluency Posttest Data	131
Table 15. Analysis of Variance for the Fluency Posttest Data	132

Table 16. Descriptive Statistics for the Originality Posttest Data	135
Table 17. Analysis of Variance for the Originality Posttest Data	136
Table 18. Overall Interjudge Reliability, Pilot Test.....	194
Table 19. Interjudge Reliability for Pattern Use, Pilot Test	194
Table 20. Interjudge Reliability for Cohesiveness, Pilot Test.....	195
Table 21. Interjudge Reliability for Creativity, Pilot Test	195
Table 22. Interjudge Reliability for Pattern Use	196
Table 23. Interjudge Reliability for Pattern Use by Each Pair of Judges	197
Table 24. Interjudge Reliability for Cohesiveness	198
Table 25. Interjudge Reliability for Cohesiveness by Each Pair of Judges.....	198
Table 26. Interjudge Reliability for Creativity	199
Table 27. Interjudge Reliability for Creativity by Each Pair of Judges	199

LIST OF FIGURES

Figure 1. Example of a melodic sequence based on the visual context..	156
Figure 2. Sample size and treatment assignment	169
Figure 3. Example of treatment and testing sessions assignments.....	170
Figure 4. Examples of Comment Sheets.....	171
Figure 5. Treatment Design.....	173
Figure 6. Transcription of Composition Example A for Judges Training Session	191
Figure 7. Transcription of Composition Example B for Judges Training Session	192

CHAPTER I

INTRODUCTION

Background

Encouraging the development of children's problem solving skills has been an important consideration for North Americans. In 1994, the U.S. Congress created and passed the Goals 2000: Educate America Act. In this act, the Congress declared that one of the National Education Goals was, among others, that "...every school in America will ensure that all students learn to use their minds well, so they may be prepared for responsible citizenship, further learning, and productive employment in our Nation's modern economy" (U.S. Congress, 1993, Third clause, p. 5). The second objective of this goal was that "the percentage of all students who demonstrate the ability to reason, solve problems, apply knowledge, and write and communicate effectively will increase substantially" (U.S. Congress, 1993, Third clause, p. 5).

Throughout history, there have been several efforts to understand and define intellectual skills such as the ability to reason, solve problems, apply knowledge, and write and communicate effectively. As a post-Sputnik reaction during the 1950s, educators became more interested in processes of problem solving. The writings of Jerome Bruner (1960) and Benjamin Bloom (1956) were examples of this interest in problem solving processes. Bloom (1956) stated that even more important than the teaching of information, was the need for "some evidence that students can do something with their knowledge, that is, that they can apply the information to new situations and problems" (p. 38). According to Bloom,

This has been labeled as "critical thinking" by some, "reflective thinking" by Dewey and others, and "problem solving" by still others. In the taxonomy we

have used the term “intellectual abilities and skills.” The most general operational definition of these abilities and skills is that the individual can find appropriate information and techniques in his previous experience to bring to bear on new problems and situations. This requires some analysis or understanding of the new situation; it requires a background of knowledge or methods that can be readily utilized; and it also requires some facility in discerning the appropriate relations between previous experience and the new situation. (p. 38)

According to Robertson (2001), a problem solving situation arises when we do not know exactly what steps to take to get to the goal and we have to take some “mediating action” (p. 4). If the student knows the needed course of action, that means the goal can be achieved by making use of some specific previous experience the student had with an equal problem solving situation. In other words, once we know how to solve a problem, it is not a problem anymore. Treffinger, Isaksen, and Dorval (1994) stressed the idea that creative problem solving “might be termed a problem if, and only if, the person does not already know [the] solution” (p. 226).

Treffinger, Isaksen, and Dorval (1994) are not alone when it comes to include the words “creativity” and “problem solving” in the same sentence. Several authors within the field of educational research have seen problem solving as closely connected to creativity. Mumford and Connelly (1994) recognized that the combination and reorganization of existing knowledge structures must be used to generate “new ideas or novel problem solutions...creativity achievement is reflected in the production of useful, new ideas or products” (p. 2). In fact, Kay (1994) defined creativity as “a process whereby the individual finds, defines, or discovers an idea or problem not predetermined by the situation or task” (p. 117).

The creative problem solving process has also been explored in the area of music education. According to DeLorenzo, “creative thinking, linked with the cultivation

of discovery and invention, seems an essential and necessary component in the educational process, since it is through the generation of new questions and ideas that a discipline perpetuates itself” (p. 4).

Further, Elliott (1995) said that music making and musicianship should involve not merely the completion of tasks. The process of solving problems, or tasks, should be done with enlightenment, with a critical consciousness. When using the word *praxis* to describe the nature of music making and musicianship, Elliott said that,

Praxis means action committed to achieving goals (*telos*) in relation to standards, traditions, images, and purposes (*eidos*) viewed as Ideas that are themselves open to renewal, reformulation, and improvement...Put in another way, to act artistically as a music maker is to engage in music making and music listening...as *praxis*. (p. 69)

The Need for Research on Composition as a Creative Musical Problem Solving Task

Many music educators have recognized the importance of involving students in activities that encourage the development of creative problem solving skills. In 1989, the Music Educators National Conference published *Dimensions of Musical Thinking* (Boardman, 1989), where problem solving was identified as a “prerequisite for human survival” (p. 45). Throughout the history of music education in the United States, there have been many efforts to encourage the development of creative problem solving through composition in the music classroom. The importance of composition was stressed by the creators of different music education curricula, such as the Manhattanville Music Curriculum Program (MMCP) (U.S. Congress, 1989). Ronald Thomas, director of the MMPC, stated as one of the basic premises of the MMPC that “If [music] is a creative art, learning means creating” (U.S. Congress, 1989, p. 4). These efforts to encourage the development of creative problem solving through composition

in the music classroom are also reflected in the National Standards for Arts Education (Music Educators National Conference, 1994). The fourth of these standards stresses the importance of “composing and arranging music within specified guidelines” (p. 27).

Despite the importance that professional institutions and individual music educators have given to it, according to Sloboda (1985), “composition is the least studied and least well understood of all musical processes” (p. 103). The lack of understanding that music educators have about children’s compositions might be one of the reasons for which composition is a rather atypical activity in the music classroom. Another possible reason for the lack of compositional activities in music classrooms may be that music educators do not feel comfortable themselves as composers. They may hold the idea of a composer as an idealized individual with supernatural powers touched by the muses. This idea affects both the way they perceive themselves as composers, and the way in which they think about children as potential composers. If educators do not perceive themselves as having these supernatural powers of composition, they may not feel comfortable composing or providing composition experiences for their students. Educators may also think that only a musically gifted child needs early exposure to composing experiences. Campbell (1998) believed that because of this idealized concept about musicians, children who were not perceived as talented were not being stimulated enough in the schools. They were mostly being “exposed,” but not stimulated to their full potential. Campbell (1998) said that,

This Eurocentric concept of musical talent evokes thoughts of Mozart, the wunderkind, composing at five and performing in the general European courts as a schoolchild. It creates images of musical participation for the very few: richly endowed geniuses who manifest enormous musical memories while still in the nursery...composers who think profound and “universal” musical thoughts and structures to commit notes on paper...these prodigies may be seen as so

distinguished as t to possess superhuman abilities that separate them from the “norm,” the mainstream. (p. 169)

Even when music educators have added composition to their curricular activities, assessing the compositional products has been yet another challenge. According to Perkins (1984), there has been a propensity among schools to rely on activities that require simple cognition and where learning outcomes can be clearly defined and measured with standardized instruments. As a result, there is a tendency to perceive music composition as depending on the ability to read and notate music. Further, DeLorenzo (1987) pointed out that, while music curricula often reflect an emphasis on the acquisition of music reading skills, fingering patterns, and factual information, it was possible that students’ musical thinking could be at a more advanced level than their music reading and writing skills. Alas, DeLorenzo (1987) observed during her study, music classroom activities “did not appear to enhance the student’s quality of thinking or attitude toward artistic process, goals that teachers deemed important in the instructional program” (p. 11).

Clearly, music educators do not have enough information about the ways in which children compose. Research exploring the creative musical problem solving skills involved in the compositional process, and the compositional products that result from these process, is relatively new in music education (Richardson & Whitaker, 1992). Research focusing on the ways in which children compose in various settings and with different tasks structures, and on the music compositions children create, may provide music educators with insight as to how to structure creative musical problem solving experiences in the classroom. It could also give them ideas for the assessment of children’s compositions in a practical manner.

Individual versus Collective Musical Problem Solving

When exploring children's compositional process, we need to understand music's social nature. Early in childhood, children are exposed to musical laws specific to the culture within which they are born. The musical environment to which children are exposed is often reflected in their compositions. According to Elliott (1995), when individuals compose, they are not acting alone. Their culture situates the composition in three ways:

First, the musicianship required to compose particular kinds of music develops in relation to the thinking of other composers and performers, past and present, who have immersed themselves in the achievements and authority (or de standards and traditions) of particular compositional practices...Second, composing is highly contextual in that composers do not generate and select musical ideas in abstraction...They compose particular forms of music...Third, among the most important guidelines that composers use to channel and evaluate their composing are performance practices. (p. 162)

Another social aspect of music is reflected in the ways in which children use music. As Campbell (1998) said, children socialize through music. They get together and sing, play, dance, and even exercise to music. Immigrant children can learn basic rules of interaction and language of their new social group through songs and musical games. Children can make new friends through their participation in music. Campbell added that "Music appears to be everywhere in the lives of children" (p. 168).

However, there are times in which children do not share music with anyone else. They "take music into themselves at their most private of times.... They interpret it for its messages to them and absorb and rework it in new configurations as their very own music" (Campbell, 1998, p. 168). As a result of these moments, children's compositions speak of their own individual musical preferences, their own individual musical values.

Because the process of composition involves these two aspects of music, the collective and the individual, it is not easy to determine if children who are exposed to composition activities would benefit more from a collective experience or an individual one. It might be that students absorb musical concepts presented by their peers and use them to expand their own musical vocabulary. But it might also be that children have their musical vocabulary shaped early in life, and interaction with their peers would not necessarily enrich their musical creations.

In her book, Wiggins (1990) suggested the use of both collective and individual composition activities. According to Wiggins, “Group composition...reinforces social skills such as the building of interpersonal relationships and skills in group discussion, organization, listening, and valuing the ideas of others” (p. 2). Individual composing activities, Wiggins suggested, provide the teacher with the opportunity to assess individual students’ musical knowledge and ability. But we do not have enough information about the comparison of individual and collective creative musical problem solving experiences in relation to children’s compositional products. Clearly, and as it was suggested by Kaschub (1999), “music educators need information regarding the efficacy of task structure and the learning potential inherent in individual and collaborative creative engagements as both conditions are possible within the classroom and rehearsal hall” (p. 7).

Historically, there has been a controversy within the area of educational psychology over whether collaborative or individual problem solving is more conducive to learning. Marzano, Brandt, et al. (1988) and Marzano, Pickering, et al. (1997) perceived complex thinking as a social process and insisted on the importance of

children's interactions with their teachers and peers. However, in a contrasting position, Dewey (1910) perceived the processes that take place during a reflective thinking experience as highly individualistic. Different people may spend more or less time at each phase, or even proceed through the phases in a different order. This may depend on the way in which each person perceives the problem, as well as on the degree to which the problem is perceived to be personally relevant (Dewey, 1910). Another similar contrast can be found between Piaget (1967), who saw children's developmental process going from individualistic to social, and both Vygotsky (1978) and Mead (1970), who believed that individuals develop from social to more individualistic.

According to Piaget (1967), it is not until children are seven years old that they can have a real interchange of ideas. Before then, children "speak as if they were talking to themselves" (p. 20). Based on Piaget's ideas, collaborative problem solving cannot take place in reality during children's early stages of development. If children were asked to solve a problem together, there would be egocentric monologues, but there would not be a real interchange of ideas, or a joint effort to solve the problem at hand. At this age, children's social behavior remains "midway along the road between egocentrism and true socialization" (Piaget, 1967, p. 21). At about age seven, children become able to concentrate when working individually, and they also become able to collaborate effectively in a group. Piaget (1967) said that at this stage,

...the child becomes capable of cooperation because he no longer confuses his point of view with that of others. He is able both to dissociate his point of view with that of others and to coordinate these different points of view... True discussion are now possible in that the children show comprehension with respect to other's point of view and search for justification of proof with respect to their own statements. (p. 39)

Piaget believed that children's development moves from highly individualistic, or egocentric, towards social cooperation. Collaborative problem solving could take place and would be more effective as the older children reach the stage of formal operations, in which they would not merely be imitating adults or more experienced peers, but they would be identifying, comprehending, and respecting the viewpoints of others. Because of this recognition of others' points of view from their own, older children or young adults could take the input provided by others during different problem solving processes, assimilate it to their own, and collaborate in making a final decision.

In contrast to Piaget's developmental process going from individualistic to social, Mead (1970) and Vygotsky (1978) believed that individuals develop from social to more individualistic. In fact, Mead (1970) said that,

The Self...is essentially a social structure, and it arises in social experience. After a self has arisen, it in a certain sense provides for itself its social experiences, and so we can conceive of an absolutely solitary self. But it is impossible to conceive of a self arising outside of social experience. When it has arisen we can think of a person in solitary confinement for the rest of his life, but who still has himself as a companion, and is able to think and converse with himself as he had communicated with others. (p. 140)

According to Mead, the very process of thinking is "simply an inner conversation that goes on" (p. 141). Once the self has been established and the individual is able to undergo inner conversations, there is a need for an audience that listens to one's ideas. The individual internally identifies the importance of what s/he is going to express, and prepares the discourse before actually expressing it. The individual would then think out his idea, expressing it perhaps in the form of a book or in the form of a musical discourse, but this expression would still be a part of a social intercourse "in which one is addressing other persons and at the same time addressing one's self" (p.142).

Thinking, Mead said, “is a conversation of gestures which in its completion implies the expression of that which one thinks to an audience” (p. 141).

Vygotsky (1978) agreed with both Piaget and Mead in that speech was of central importance for complex thinking processes such as problem solving. In relation to problem solving, Vygotsky said that a child who is able to use speech divides a problem solving activity into two consecutive parts. First, the child plans how to solve the problem through speech and then carries out the prepared solution through overt activity. Vygotsky thought that egocentric speech, as described by Piaget, increased as the difficulty of the problem to be solved by the child increased. Vygotsky (1978) hypothesized that “children’s egocentric speech should be regarded as the transitional form between external and internal speech” (p. 27). When children are not able to solve a problem by themselves, they turn to an adult, verbally describe the part of the problem solving process that they had not been able to carry out by themselves, and ask for help. According to Vygotsky,

The greatest change in children’s capacity to use language as a problem-solving tool takes place somewhat later in their development, when socialized speech (which has previously been used to address an adult) *is turned inward*. Instead of appealing to the adult, children appeal to themselves; language thus takes on an *intrapersonal function* in addition to its interpersonal use. (p. 27)

Purpose of the Study

The purpose of the study was to explore the effects that individual versus collective structured creative musical problem solving tasks had on students’ compositional products. The null hypothesis stated that there would be no difference on scores obtained from students trained under collective condition with those obtained by students trained under individual condition. The specific research questions are:

1. Is there a significant difference between the level of tonal and metric cohesiveness identified on the compositional products of fourth- and fifth-grade students who participated in individual creative problem solving activities, and the level of tonal and metric cohesiveness identified in the compositional products of students who participated in collective creative problem solving activities?
2. Is there a significant difference between the melodic and rhythmic pattern-use identified on the compositional products of fourth- and fifth- grade students who participated in individual creative problem solving activities, and the melodic and rhythmic pattern-use identified on the compositional products of fourth- and fifth- grade students who participated in collective creative problem solving activities?
3. Is there a significant difference between the level of fluency and of originality found in the compositions of fourth- and fifth- grade students who participated in individual creative problem solving activities from that found on the compositions of fourth- and fifth- grade students who participated in collective creative problem solving activities?

Limitations of the Study

Because there were time and space limitations at the host school, it was not possible to select students through true randomization. Therefore, the findings of the study can not be generalized to every fourth- and fifth-grade student. Also because of these time and space limitations, the study took place during a five-week period, which included one week for the pretest, three weeks for the treatment sessions, and one

week for the posttest. It was not possible to have a longer experimental condition, which might have produced different results.

Assumptions

The following assumptions were made in this study:

1. It is possible to familiarize subjects with the different stages required by the creative problem solving process through the posing of questions related to this process, and without an open discussion between the subjects and the person in charge of the treatment sessions.
2. All subjects possessed normal physical and mental abilities.
3. All subjects possessed basic musical knowledge that enabled them to work on the task.
4. All subjects had been acculturated within the traditional western music style. Therefore, it was possible to explore their compositions in terms of their tonal and metric cohesiveness, and the use of melodic and rhythmic patterns as basic structural components.

Definition of Terms

For the purpose of this study, *creative musical problem solving* was operationally defined as the process through which children (a) became aware of the compositional challenge, analyzed the situation, considered musical options, technical concerns, and any other circumstances involved in the composition of a final product; (b) gathered their musical knowledge and previous experiences in order to test different combinations of notes and silences, and identify specific musical ideas; (c) examined

promising sounds and silence combinations to determine the steps needed to achieve the ultimate goal of solving the problem; (d) presented a final compositional product.

Musical composition was operationally defined as the combination of sounds and silences that subjects declared to be the result of their creative problem solving process, throughout which they had the time to revise their work. Subjects were neither requested nor forbidden to write down their compositions.

For the purpose of this study, the terms included in the measurement instrument were defined as follows:

Tonal cohesiveness was defined as the degree to which the pitches in a composition are constructed around a tonal center or tonal centers (Kratus, 1994, p. 121).

Metric cohesiveness was defined as the degree to which the durations in a composition are constructed of regularly occurring accented and unaccented beats (Kratus, 1994, p. 121).

Melodic pattern was defined as two to seven pitches that form a distinct pitch pattern that is perceived as a unified whole (Kratus, 1994, p. 121).

Repeated melodic pattern was defined as a melodic pattern that is identical to a previously occurring melodic pattern (Kratus, 1994, p. 121).

Developed melodic pattern was defined as a melodic pattern that is similar to, yet different from, a previously occurring melodic pattern. In most cases, the rhythm is held the same, the melody changes, and the melodic contour is similar (Kratus, 1994, p. 121).

Rhythmic pattern was defined as two to seven durations that form a distinct duration pattern that is perceived as a unified whole (Kratus, 1994, p. 121).

Repeated rhythmic pattern was defined as a rhythmic pattern that is identical to a previously occurring rhythmic pattern (Kratus, 1994, p. 121).

Developed rhythmic pattern was defined as a rhythmic pattern that is similar to, yet different from, a previously occurring rhythmic pattern. In most cases, the melody is the same or similar and the rhythm changes (Kratus, 1994, p. 121).

Musical fluency was defined as the ability to generate a number of clearly defined musical ideas (Webster, 1977, p. 137).

Musical originality was defined as the ability to create a composition that contains elements of unique expression (Webster, 1977, p. 137).

CHAPTER II

REVIEW OF LITERATURE

The theoretical foundations for this study are found in different areas, including individual and collective problem solving, creative problem finding and solving, and children's musical composition processes and products. The following review of the literature covers these related theories and research, which have served as the theoretical framework for the present study.

Research on Collective versus Individual Creative Problem Solving

Most researchers within and outside the realm of music education have explored creative problem solving processes and products in collective settings. The comparison between collective and individual creative problem solving experiences has seldom been investigated. Most studies comparing the effect that collective versus individual problem solving experiences have on children's problem solving products, have established that the collective experience benefited low-ability subjects, but it did not have a significant effect on high-ability subjects (Azmitia, 1988; Duran & Gauvain, 1993; Garton & Pratt, 2001). Average-ability subjects were not included in these analyzes. When comparing students without grouping them according to their ability, Chiara, Schuster, Bell, and Wolery (1995) found that individual performance in creative problem solving tasks improved when subjects were trained individually. Larey and Paulus (1999) found that individuals preferred to work in groups, but that subjects working individually outperformed subjects working collectively. From these studies, then, it might be inferred that individual students may not benefit from the collective creative

problem solving experience as much as they do from the individual creative problem solving experience.

Azmitia's study (1988) investigated children's solitary and collaborative performance on a construction task that involved reproducing a Lego model. If it was found that subjects benefited from interaction, the author wanted to find out if this also benefited subjects' subsequent individual performance. Also, if subjects benefited from interaction, the author was interested in identifying the features of interaction that promoted learning.

As a result of the pretest, where subjects were asked to copy a Lego model, Azmitia (1988) classified subjects as either novices or experts. Then, the author assigned the subjects to one of three experimental conditions: (a) individual, (b) same ability dyads, and (c) expert-novice dyads. Dyads were formed pairing same-sex children. During the second and third sessions, subjects in the three experimental conditions were asked to build a new house copying a Lego model. As a posttest, subjects were tested with the same task, but all of them worked individually.

Data analysis revealed that experts' performance did not change over sessions. Novices paired with novices, and those who worked individually, did not show a significant improvement. In contrast, novices who worked with experts showed a significant improvement in each of the treatment sessions. The significant difference between novices who were exposed to the individual condition and novices who were exposed to the expert-novice condition was maintained in session 4, when all novices were working by themselves.

Azmitia (1988) also found that high-ability subjects spent more time on the task than low-ability subjects. Experts and novices working with experts tended to look more frequently at the model than novices working alone or with other novices. Same-ability dyads engaged in more task-related conversation than mixed-ability dyads. Experts gave more correct explanations and demonstrations than novices, and more explanations and demonstrations were given during the first than during the second interactive session. The author concluded that during the preschool years, collaboration seemed to be a more effective learning condition, but only for novice subjects. Expert subjects did not seem to benefit from the collective experience. However, in the case of expert subjects, the lack of improvement could not be interpreted easily because it was possible that, prior to the study, they had already reached their highest building skill level. The author stressed that only novices who worked with an expert showed a significant improvement during the individual posttest. Experts seemed to contribute to novice's learning.

In a study exploring the role of age and ability in collective problem solving, Duran and Gauvain (1993) compared 7-year-old high-ability subjects paired with 5-year-old novice-ability subjects, to 5-year-old experts paired with same-age novices. The author found that the posttest performance of novice subjects working with same-age experts was significantly better than the performance of subjects paired with 7-year-old high-ability subjects. Duran and Gauvain (1993) also found that novices working with same-age experts had more involvement in the task than novices in different-age dyads. The performance of 7-year-old high-ability subjects, and that of the 5-year-old expert subjects, did not improve significantly after the treatment. The author concluded

that the collective experience benefited novice subjects when they were paired with high-ability subjects in the same age group, but did not have a significant effect on the performance of high ability subjects.

Garton and Pratt (2001) explored the extent to which the creation of a context which facilitates conversation and communication would enhance the cognitive capacities of 4- and 7-year old children working on a problem solving task involving sorting skills. The authors assigned subjects to one of three experimental conditions: (a) individual, (b) same ability dyads, and (c) expert-novice dyads. Similarly to Azmitia (1988) and Duran and Gauvain (1993), Garton and Pratt (2001) found that problem solving facilitated novice subjects' subsequent individual problem solving when they were paired with higher-ability children.

Chiara, Schuster, Bell, and Wolery (1995) compared the effect that training in small groups versus individual settings had on preschool children's capacity to solve a problem. For the small group condition, subjects were not interacting or solving the problem together. All subjects participated in both conditions and were present during both conditions. For both conditions, subjects were taught the name of one pair of pictures. Another pair of pictures was not taught "and served as a control for history and maturation effects" (p. 206). For the small group condition, the teacher asked each subject to touch each picture and to identify what it was. No prompts were used to have subjects attend to each others' trials. For the individually administered trial, the teacher approached each individual subject and asked for the identification of the pictures. The researchers found that the individualized approach resulted in "at least minimally more efficient learning...for most children" (p. 214).

According to Peel (1998), “problem-based learning in groups has been widely used in medical education” (p. 35). Peel investigated a problem-solving strategy known as Think Tanks. The Think Tanks in this study involved subjects working together in small groups for about seven weeks. Each group was assigned to solve a problem and was to present the results as a poster display. By exploring subjects’ evaluations of the project, the author found that most of the subjects were very positive about the opportunity to work with peers in the project. However, the effectiveness of collective work was not assessed. The authors reported only subjects’ preference for working in groups.

Larey and Paulus (1999) explored the impact that preference for working in groups had on subjects’ interactions during a brainstorming session. These authors developed a multiple-item scale to measure general preference for working in groups. Subjects were placed in either high-preference interactive, high-preference nominal, low-preference interactive, or low-preference nominal groups. The interactive condition involved face-to-face brainstorming. The nominal group condition involved the exchange of written ideas. All subjects were asked to solve a problem. At the end of the activity, subjects were given a questionnaire that dealt with perception of performance on the brainstorming task. Larey and Paulus (1999) found that nominal groups were significantly more active than interactive groups. Interactive high-preference groups were more active than interactive low-preference groups, and nominal low-preference participants were more active than nominal high-preference participants. Nominal groups generated more ideas than interactive groups. Nominal low-preference groups were more productive than nominal high-preference groups. Nominal groups explored

significantly more options than interactive groups, but they indicated that they would have generated more ideas in a group setting.

Research on Collective versus Individual Creative Problem Solving in Music

The effects that either individual or collective creative problem solving experiences could have on students' compositions have seldom been researched in the area of music education. Most of the investigations looking at social factors affecting students' problem solving skills have focused on the effects of individual versus collective activities on students' pitch perception or production, or some other music variable other than creative problem solving skills. For example, McCarthy (1979) investigated the effect that individual and collective instruction and students' demographic characteristics had on two measures of music reading and dropout of fifth- and sixth-grade beginning instrumental music students. McCarthy concluded that "individualized instruction by itself was not a significant positive or negative influence on either of the music achievement measures, nor did it interact with any of the remaining variables used in the study" (McCarthy, 1979, p. 68).

Researchers in the area of music education have tended to study either individual or collective creative processes and products. There are few examples of studies that compared the individual with the collective creative problem solving experience. When analyzing the research studies that have taken place in the area of music education, it is interesting to notice that the findings are in contrast to those obtained in other educational areas. As was seen earlier in this chapter, studies exploring the effect that individual versus collective experiences have on subjects' creative solving products have tended to favor individual experiences for students in

general, or have tended to favor the collective experience only for the low-ability subjects. In contrast, research studies in the music education area have tended to favor collective experiences, or collective and individual experiences equally.

Kaschub (1999) explored sixth-grade students' descriptions of their individual and collective music composition processes and products. Subjects in this study were asked to complete one prompted and one unprompted task. For the prompted task, subjects were given a poem and were asked to set it to a piece of music. The unprompted task did not include the text. Subjects in both conditions completed the prompted task before the unprompted one. At the end of the compositional process, subjects described and assessed their compositional processes and products.

According to Kaschub (1999), subjects working individually reported difficulty with generating ideas, and often used the first idea they came up with as the germinal idea for their final composition. Subjects working in the collective condition seemed to have generated more ideas than subjects working in the individual condition, and often used ideas that "emerged during latter portions of compositional work time" (p. 239). According to the author, individual composers reported taking time to plan their composition, which Kaschub (1999) qualified as "hesitancy and uncertainty" (p. 239). In contrast, collaborators initiated their compositions immediately, by choosing their individual musical instruments. Finally, individuals reported "moderate levels of interest while collaborators reported a slightly higher level of interest overall" (p. 240).

In relation to the evaluations of their final products, individual composers tended to describe individual musical elements, such as melody or style, whereas collaborators expressed more of their feelings about the overall quality of their compositions. In

relation to the task, Kaschub (1999) concluded that setting the text to music was difficult for the subjects, whose evaluation of the final composition was affected by how much they liked or disliked the poem. The author also found that, “although not significantly different by proportion, the ratings given to musical ideas and performances associated with the products generated from the unprompted task were slightly higher overall than the ratings given to compositions generated from the prompted task” (p. 242). Kaschub (1999) suggested that, “students should be exposed to both collaborative and individual compositional experiences so that they may enjoy the support of their peers and also experience the freedoms afforded by individual work” (p. 250).

Among the few studies that have explored some aspect of students’ participation in collective musical problem solving experiences, MacDonald and Miell’s study (2000) examined the impact of social factors on children’s interaction while engaged in music making or listening. The authors assumed, based on Vygotsky’s ideas, that children co-construct knowledge “through joint action and discourse, mediated through the use of cultural tools and artifacts” (p. 59). Further, the authors stated that “by actively engaging with each other’s ideas, modifying and extending them, the children develop their understanding and their critical thinking skills” (p. 59).

In their study, MacDonald and Miell (2000) asked 10- and 11-year-old children to compose a piece of music in any style they chose which reflected the idea of a rain forest. Half of the children were asked to work with their best friend, and the other half worked with a child from a different class who was not their friend. Since the authors were interested not only in the compositional product, but also in the process, the sessions were videotaped and recorded on audiocassette. Subjects’ talk was coded as

un-transactive or transactive, with transactive communication being defined as communication that builds upon and extends ideas that have already been voiced. A teacher from another school who was unaware of the hypothesis being tested and of the experimental conditions rated the final compositions for quality.

Based on their results, the authors concluded that the compositions produced by friends were ranked significantly higher than those produced by non-friendship pairs. Also, musical and verbal communication were considered qualitatively different between the two groups, with the friend groups speaking and producing more music than the non-friend groups. The style of interaction presented by friend-pairs was found to be significantly related to the teacher's higher score for these pairs, leading the authors to conclude that the presence of more transactive communication was what led to the higher quality compositions from the pairs of friends.

From a qualitative perspective, Claire (1993) examined the nature and importance of peer social interactions, "as influenced by the structure of classroom work processes, for fifth-graders' academic and artistic creative activity" (p. 21). The author observed 3 fifth-grade classes in three different schools with teachers committed to fostering children's creative abilities, and made comparisons both within and across classrooms. Claire found the terms "hierarchical" and "mutual" as the ones that most accurately and fully described the nature of differing work processes. According to the author, hierarchical work processes have a linear delineation of power and control over decision-making. Under this circumstance, one child or the teacher accepts, claims, and/or has responsibility for all decisions and the others must comply. In contrast,

mutual work processes have a collaborative structure for sharing decision-making among teachers and students.

Claire (1993) found that, during mutual work processes, subjects focused on their work. There were few and brief interruptions of concentration. In mutually oriented work processes, subjects generally appeared neither to copy nor to compete with each other. The author said that “instead of concentrating themselves with obtaining ‘right’ answers or products that were ‘better’ than their peers, subjects usually worked with their own ideas until they achieved satisfaction. Subjects also took more frequent risks to explore their own ideas” (p. 24).

In contrast, in hierarchical work processes, the subjects’ focus of attention shifted from deep involvement with work to issues of safety, preoccupation to achieve the right answer and acquire status, control, fun, and sociable exchange. Subjects’ interactions were less frequent, and there were evidences of power struggles. Subjects “rarely considered risking exploration of their own ideas and even less frequently pursued such risks” (p. 25). The author concluded that the structure of the work process shaped peer social processes which influenced subjects’ creative activity. The author suggested that “music teachers have the responsibility to design creative musical activities which are structurally congruent with the creative process, and to establish mutually oriented classroom work processes which facilitate student creative activity” (p. 27).

In a descriptive field study, Christensen (1992) developed a protocol for assessing children’s musical understanding. Subjects for this study were fourth-grade students, grouped in three teams of 4 subjects each. One of the teams comprised girls only. The other two teams comprised boys only. Assignments to the teams were done

through self-selection. Christensen (1992) asked subjects to collectively create and perform instrumental compositions, to invent visual representations for notating the compositions, and to reflect in writing and orally about themselves as musicians and about their compositional processes and products.

From subjects' written reflections, tape recorded oral interviews, and individual sketches and group representations of the compositions, the researcher drew conclusions about subjects' understanding of music, subjects' responses to collaboration, and subjects' metacognition about the compositional process. Christensen (1992) concluded that "music learning is enhanced when based on an instructional model that includes perception, production, and reflection" (p. 172). In relation to the collaborative work structure, Christensen (1992) concluded that it proved to be a positive component of the compositional task. The author reported, however, that subjects were hesitant to participate in collective activities, and became at ease once they were informed that the members of the teams would be self-selected. "The researcher interpreted this as an indication that there may have been some aspects of working in the highly organized cooperative styles that were not always satisfying to students" (p. 207).

Also from a qualitative perspective, Wiggins (2000) explored the ways in which shared musical understanding could influence students' problem solving processes. The author selected examples from 600 audiotapes and videotapes that the author had collected through previous studies of children's creative processes in classroom settings. With these data, Wiggins explored the nature of shared understanding about

music as it was present in six representative situations in which subjects were engaged in either composing or improvising with peers and/or with their teacher.

For the first situation, data included videotapes of two target students as they participated in a general music class, and a journal of the teacher's plans, reflections, and field notes. In the analysis, the author found that the subjects were engaged in a series of unrelated verbal conversations as they were composing a piece for soprano recorder. Wiggins concluded that "the students' shared understanding of both their process and product was so complete that the composing did not seem to occupy their full attention" (p. 72).

For the second situation, Wiggins (2000) transcribed a music lesson in which a third-grade general music class worked together with the teacher to write a song. In this instance, the author observed that motivic material was generally conceived by individuals within the group and then suggested to the group. Wiggins also observed that there was a common understanding of how the final product should sound.

In the third situation, Wiggins (2000) explored the data obtained through interviews, audiotapes, videotapes, and observations as 5 fifth-grade students were working together to compose an original song. In this instance, Wiggins found that most of the conversations between the group members took place in the form of melodic phrases or motives. The verbal interactions were mostly evaluative comments. According to the author, the shared understanding of what the subjects thought the song ought to sound like guided the subjects' choices. Wiggins concluded that "the group's broader vision of what would constitute an appropriate song seemed larger than each individual idea, providing a basis for evaluation of individual ideas as each was

suggested to the group” (p. 79). However, Wiggins also reported that there were some disagreements between the group members, which generated the need on the part of individuals to justify, defend, and reshape their ideas. According to Wiggins, “the need for constant justification also promoted a more solid formulation of the thematic material” (p. 79).

For her fourth, fifth, and sixth situations, Wiggins selected three audiotapes from the original 600 recordings of students’ compositions and improvisations. The author considered the selected three recordings as particularly successful products that resulted from shared understanding among children. In the first tape, the teacher encouraged a group of first-graders to create a musical conversation. One student volunteered to start the conversation by asking a musical question on the xylophone. The teacher answered, and the conversation continued under these lines. According to Wiggins, it was evident that the student understood the role of each individual in the conversation, and that this understanding grew throughout the experience. Wiggins concluded that this was “a clear example of how one student’s participation in a collaborative creative experience empowered his individual ability to generate original musical ideas” (p. 81).

For the fifth and sixth situations, Wiggins analyzed the improvisation of 2 fifth-grade students, and the composition of 4 fifth-grade students respectively. For the improvisation task, the students played chord tones on resonator bells. Through an analysis of the melodic product, Wiggins concluded that the students showed a high level of shared musical understanding. The author arrived at the same conclusion when she analyzed the sixth situation, the composition of the 4 fifth-graders. In this case, the

author pointed out that a work with the level of complexity found in the composition, “could only have been developed by four individuals who possessed an extremely high level of shared understanding of both their intentions and what it took to carry those intentions out” (p. 83). In conclusion to the analysis of the six musical situations, Wiggins said that “when individuals need to make their ideas public, explain, justify, defend, and alter them to accommodate someone else’s viewpoint, the process strengthens the individual’s ability to initiate ideas” (2000, p.87).

Research on Creative Problem Solving Strategies

Some educational psychologists have theorized that students’ creative problem solving skills within a specific content area would improve through the learning of general problem solving strategies. There have been some general education and music education researchers who have explored the effect that teaching students strategies required by the creative problem solving process has had on various educational goals. The findings of most of these research efforts seem to indicate that when students are provided with specific problem solving strategies, they can solve creative problems more efficiently.

Within the music education area, Eisman (1975) devised and tested the relative effectiveness of two methods of teaching listening skills. Eisman compared the lecture-demonstration approach with a problem-solving approach. For the lecture-demonstration approach, the researcher established the class objectives. For the problem-solving approach, the class had to discover the lessons’ objectives. Toward this end, the researcher posed questions to encourage subjects to offer solutions to the problem. Finally, subjects’ questions were rephrased, and presented as problems to be

solved. As a pretest and a posttest, Eisman used Colwell's (1969) Music Achievement Test (MAT) to measure listening skills. The researcher found that there were no significant differences between the test scores of the groups. It is important to point out that the researcher found the MAT to be inappropriate for this investigation, given the fact that it measures components such as pitch and rhythm perception, while the study's subjects were to perceive less specific music qualities, such as melodic and rhythmic repetition or variation and tone color.

Outside of music education, Cope and Murphy (1981) investigated if the learning of specific problem solving strategies was a necessary step for successful problem solving. A known strategy, Cope and Murphy sustained, could be used as a guide to organize the solutions that have been used to solve problems in the past, into a form that could achieve the final solution of a problem to be solved in the present. Both experimental and control groups were given two problems. Subjects in the experimental group received an extra document in which the elements of a successful problem solving strategy were explained. The authors found that 78 percent of the experimental group managed to solve at least one of the problems when they used the suggested strategy, whereas only 27 percent of the control group had similar success. The authors concluded that, even when subjects had no previous experience in higher mathematics, they had a better chance at solving the posed problem when they used a problem solving strategy. The authors said that "the possession of a strategy is a necessary condition for problem solution" (p. 15).

Tomic (1995) investigated the effects that a specific training program designed to foster the development of inductive reasoning and problem solving skills had on

children's performance on inductive reasoning tasks. First, Tomic wanted to find if there was any effect. Second, the author was interested in observing if the effect had any durability. Third, Tomic wanted to determine the range or types of transfer induced by the training program. Analysis of the results showed that the experimental group scored significantly higher than the control group. When the subjects were again measured four months later, the results showed that the experimental group continued to score significantly higher than the control group. The author concluded that training in inductive reasoning was successful in teaching third-graders to transfer inductive reasoning procedures and strategies and that the effects of this training had a long-term effect on students' inductive reasoning and problem solving skills.

Mumford, Feldman, Hein and Nagao (2001) explored variables that could influence the performance of groups and individuals on creative problem solving tasks. Interested in comparing individual and collective performance, these authors explored the effect that providing subjects with a wide range of alternative solutions for use in problem solving would have for both, groups and individuals. A group of 432 undergraduate students attending management courses watched a video intended to provide a common mental model that might be used in working on creative problem solving tasks. After the presentation of the video and the surveys, subjects were assigned to either individual or collective condition and were asked to work on one of two creative problem solving tasks. The first task was cognitive in nature and the second task was social in nature. The subjects had to provide alternative solutions to the problems. Subjects working individually were asked to write a two or three paragraph description of the proposed solution. Subjects working in groups were asked

to discuss alternatives, reach a consensus decision, and then provide a two or three paragraph description of the proposed solution.

There were two dependent variables: (a) the number of alternative solutions generated by each individual or group, and (b) the level of creativity of the solutions proposed by each individual or group. Mumford, Feldman, Hein and Nagao (2001) found that the presentation of the video that provided subjects with a wider range of alternative solutions for use in solving the problem had a positive effect on subjects' performance in the creative problem solving task. The authors also found that having more ideas available for the solution of a creative problem led to better individual performance, but not for a better collective performance. The authors believed that "the availability of a larger number of alternative solutions was, in turn, related to the tendency of individuals to produce more creative solutions to the cognitive and social problems" (p. 16). In a collective setting, however, it appeared as if the availability of a large number of diverse alternative solutions tended to disrupt group performance.

Mumford, Feldman, Hein and Nagao (2001) concluded that,

...groups...tended to produce more workable, but not necessarily more original, solutions to novel problems. Thus, when highly original solutions are needed, and a wide range of alternatives must be considered, individuals, relatively unconstrained by preconceptions and external structure, may show better performance. (p. 19)

The Exploration and Assessment of Musical Compositions

Compositional processes and products have been assessed from both qualitative and quantitative perspectives, depending on the theoretical framework that supported each research study. Some researchers have seen verbal protocols, or what the composers say is going through their minds during their compositional processes, as

the data to be analyzed in order to better understand such processes. Other researchers have explored the sounds and silences that take place during the compositional processes and products as the data to be analyzed. Either qualitatively or quantitatively, researchers have looked at adult music education experts, theoreticians, or composers, as well as at children themselves, as the assessors of children's compositional products.

DeTurk (1989) stressed the difficulty of evaluating thinking processes in general, but especially in the area of music. DeTurk suggested that essays produced by students reflected their musical thinking. A content analysis of essays could, therefore, be used as an assessment instrument. Some authors have used content analysis of texts to describe musical thinking processes. An example of this type of research is that of Whitaker (1989). Whitaker's study (1989) investigated the decision making of professional musicians while addressing real life problems. The content analysis was done using Dewey's definition of critical thinking as a framework.

Whitaker (1989) used two sources of data for her study. First, the author selected writings published in the form of books and articles written by pianists, conductors, and arrangers. Second, the researcher devised several tasks to generate verbal protocols of musical thinking. The tasks consisted of: (a) examining the score of a piece of music from the literature of their area of expertise, (b) studying a score, (c) practicing a piece of music, (d) reflecting on the practice session, (d) performing at least part of the piece, and (e) reflecting on the performance after listening to the taped performance.

The subjects were asked to think aloud and record their talking while working on the prescribed tasks. Based on her analysis, Whitaker concluded that individuals used

reflective thinking according to the problem and the context in which they were trying to solve the problem. The use of reflective thinking depended also on the subjects' past experiences. The author explained this by saying that,

One individual, encountering a situation may, in Dewey's terms, read the situation at sight and find no element of doubt: he has encountered the problem at some point previously and has resolved it. The solution is in his funded experience, and he resolves the problem using non-reflective thinking. When the second individual encounters a similar problem he experiences a doubt or perplexity that causes him to experience some form of reflective thinking. (Whitaker, 1996, p. 10)

The author concluded that it was important to acknowledge the individual's past academic knowledge, feelings, imagination and experiences in the music classroom in order to foster reflective thinking. Whitaker also suggested that it was important to provide students with a "variety of experiences involving the direct experience of music as a performer, listener, and creator: the 'objective' conditions" (Whitaker, 1996, p. 12).

Instead of using verbal materials for her analysis, DeLorenzo (1987, 1989) assumed that the exploration and evaluation of musical material, manifested through production, selection, and organization of sounds, reflected the students' inner musical thought processes. From her observations of sixth-grade students involved in musical problem solving tasks, DeLorenzo identified four characteristics of musical problem solving processes and used them as a framework for interpreting, describing and comparing students' thinking processes. These four characteristics were the degree of openness with which the students perceived the task, the degree with which the musical events determined the compositional form, the students' capacity to sense musical options, and the students' degree of personal investment. DeLorenzo concluded that, although the problems were constructed by the teacher and this might have influenced

musical decision making to some degree, “it seemed that the student’s perceptions of the problem structure – how the student cognitively represented the problem structure to himself or herself – had a far greater effect on the student’s perception of choice in the problem situation than the teacher-constructed choices” (DeLorenzo, 1989, p. 195). The students organized their musical compositions along the following lines: “(a) The given problem structure served as a matrix for inserting sound material, (b) a nonmusical plan such as a story or chain of events became the organizing structure, or (c) the musical character of sound events directed the shape of the resulting piece” (DeLorenzo, 1989, p. 195).

In relation to the degree of personal involvement, the author concluded that the more fully a student participated in the musical decision making process, “the greater investment he or she demonstrates in the emerging creative product” (p. 196). This, according to the author, may have had to do with the subjects’ perception of possibilities to solve the problem, and their capacity to shape musical events. In addition, “the perceived relevance of the task, as it relates to the student’s personal definition of music, may also have influenced individual involvement” (DeLorenzo, 1989, p. 197). Finally, DeLorenzo observed that, while the teachers participating in the project encouraged divergent thinking, they placed little emphasis on students evaluating the expressive import of sound material or organizing sound events into musical ideas.

Besides the need to qualitatively describe the processes and products of musical problem solving, there have been some attempts to quantitatively measure creative problem solving in music. In a quantitative approach to the exploration of musical thinking processes, Vold (1986) developed a Measure of Musical Problem Solving

(MMPS) to assess kindergarten children's fluency and flexibility, relating sound to feel-states, and divergent thinking with musical ideas. The study explored the relationship between the MMPS scores and scores on the test Thinking Creatively with Action and Movement (Torrance, 1981). The author expected that there would be no correlation between scores on the MMPS and teacher ratings of general classroom activities, and teacher ratings of general classroom creativity.

In order to measure divergent thinking, subjects in Vold's (1986) study were asked to produce different sounds using percussion instruments. To measure flexibility, the author observed the subjects' ability to produce sounds with different duration, intensity, and timbre. The author found that the more flexible the musician was in approaching the musical problem from different angles, the more musical ideas he was able to generate (divergent thinking). The author concluded that it was possible to construct a measure of musical problem solving ability in young children.

In a study described earlier in this chapter, Kaschub (1999) used sixth-grade children's descriptions of their individual and collective music composition processes and products as the focal data. The author believed that "the voices of the students must be heard in order for practitioners to structure effective and meaningful musical experiences for their students" (p. 245). Kaschub (1999) explored both oral interviews and written questionnaires answered by the subjects participating in the compositional projects.

Also interested in students' perspectives, McCoy (1999) explored the effect that open-ended versus well-defined problem solving tasks and guided self-reflection, had on students' compositional quality, self-assessments, and attitudes towards

composition. For the well-defined tasks, the participants were engaged in activities that focused on specific skills and concepts, and that progressed from easy to more difficult. McCoy labeled this type of task as problem solving. The final product of this process was an original composition. The open-ended tasks emerged from a philosophy of free play and exploration. In this case, participants started where the participants assigned to the problem solving condition ended. They were asked to compose an original piece by selecting a poem and developing it into a song. Subjects were expected to discover the need for the skills and concepts that had been compulsory for the participants assigned to the problem solving condition, and to take the necessary steps to attain these skills and concepts. McCoy (1999) labeled this type of task as problem finding. For the second independent variable of his study, self-reflection, McCoy (1999) designed a questionnaire intended to encourage participants to engage in an ongoing monitoring of their compositional processes, expressing and assessing their intentions, feelings, motivations, and efforts during and about the compositional process.

Participants were distributed into three different conditions: problem solving, problem solving with guided self-reflection, and problem finding with guided self-reflection. There was no control group. For the data collection, McCoy (1999) designed three measurement instruments. The first one was based on Amabile's (1982) Consensual Assessment Technique, and asked experienced music educators to indicate with a point on a line, how musical each composition was. Following Amabile's (1982) suggestions, judges were asked to use their own definitions of musicality. The second measurement instrument employed Likert-type, dichotomous, and open-ended items designed to identify participants' understandings of the compositional process and

participants' attitudes about their compositional experiences. For the third measurement instrument, participants were asked to rate their compositions using Likert-type and dichotomous scale items.

McCoy (1999) found that the subjects assigned to the problem solving with guided self-reflection condition obtained lower scores for their compositions' musicality than subjects assigned to the other two conditions. Although the differences were not significant, the author reported that the effect size was considerably large, which could mean that the results could have been significant if the sample size had been larger. In relation to the self-assessment variable, McCoy (1999) found that participants assigned to the problem finding task structure were more critical of specific aspects of their compositions, such as rhythm and melody, than participants assigned to the problem solving task structure. McCoy also found that participants assigned to the problem finding condition were more likely to indicate that they looked forward to composing music in the future than participants assigned to the problem solving with guided self reflection condition.

Hickey (2001) tested the reliability of Amabile's (1982, 1983) Consensual Assessment Technique on fourth and fifth-grade children's musical compositions. The author also compared the different reliability coefficients provided by music teachers, composers, theorists, seventh-grade children, and second-grade children. The five groups of judges were divided into (a) 10 instrumental music teachers, (b) 4 mixed-experience teachers, (c) 3 general/choral music teachers, (d) 3 professional composers, (e) 4 theorists, (f) 14 seventh-grade children, and (g) 24 second-grade children.

For the collection of data, Hickey (2001) used original compositions generated by fourth- and fifth-grade subjects. The subjects recorded their compositions in a MIDI file format. There were no compositional parameters, and the subjects were encouraged to record their compositions as many times as they wanted. The composers and the theory teachers independently listened to and assessed the compositions. The music teachers and the groups of children listened to the compositions together and then rated them independently. Hickey (2001) indicated that,

Following Amabile's suggestion for proper consensual assessment technique procedures (1996), the judges were instructed to rate products relative to one another rather than against some "absolute" standard in the domain of music.... Some craftsmanship and aesthetic-quality items were also included on the music composition rating forms, as recommended by Amabile (1983), in order to assure discriminant validity between these areas and creativity.

The rating form contained 18 items that fell under the dimensions of creativity, craftsmanship, and aesthetic appeal. All of the items consisted of 7-point Likert scales with anchors marked low, medium, and high. As recommended by Amabile (1983), judges were asked to rate the compositions' degree of creativity by using their own subjective definition of creativity.

Hickey's (2001) findings showed interjudge reliability coefficients ranging from .04 for the music composers, to .81 for the general/choral music teachers. The author reported that not only the level of reliability among the composers and between the composers and the other groups was very low, but also there were "instances of negative correlations between the composers with themselves and others" (p. 240). The author concluded that "the Consensual Assessment Technique for rating the 'creativity' of musical compositions proves to be reliable when used by the most knowledgeable (in that context) group of judges" (p. 242).

In a study exploring the relationship between children's compositional strategies and products, Daignault (1996) used Amabile's Consensual Assessment Technique (1982) to assess the creative and craftsmanship qualities of the compositional products. Daignault also explored the compositional processes using a combination of quantitative and qualitative methods. For her study, Daignault asked 10 to 11-year old students to improvise three to eight musical ideas using an electronic keyboard connected to a computer. Each of these improvisations was recorded into a sequencer computer program. Then, participants were asked to listen to all recorded improvisations and select one as the germinal idea for a composition. The final compositions were assessed by five judges according to the categories of craftsmanship and creativity (Amabile, 1982).

Following Amabile's suggestions, the judges were instructed to rate children's compositions relative to their own understanding of the abilities of children aged 10 to 11. Daignault (1996) concluded that Amabile's (1982) Consensual Assessment Technique "can be applied to children's musical compositions with an adequate level of reliability" (Daignault, 1996, p. 218). The author reported, however, that the categories of craftsmanship and creativity were highly correlated, and that the judges did not have adequate levels of agreement on the middle cluster of scores. The author concluded that there was a need to use research design that "systematically excludes the middle one-third groups of participants of the comparative analysis, and concentrates exclusively on the highest and lowest rated products" (p. 218).

In another study using Amabile's (1983, 1996) Consensual Assessment Technique, Byrne, Macdonald, and Carlton (2003) examined the link between the

concept of flow (Csikszentmihalyi, 1992) and the creative output of college students' compositions. The authors were interested in finding "whether or not optimal experience would occur if three of the conditions of flow [clear goals, immediate feedback, and balance between challenge and skill] were present during completion of the composing task" (p. 282). The authors used 16 compositions produced by 45 college students composing in groups of no more than 3 subjects. The compositions were rated by 31 of the subjects participating in the study, by 9 experienced music teachers at the college level, and by 15 postgraduate students in education, using Amabile's (1983, 1996) Consensual Assessment Technique. The experienced music college teachers also assessed the compositions on a standard set of criteria.

Byrne, Macdonald, and Carlton (2003) found a significant correlation between the standard criteria scores and the creativity scores. The authors also reported a significant correlation between the levels of creativity and the levels of optimal experience, as well as significant differences between the creativity scores provided by each of the three groups of judges. The Consensual Assessment Technique was found to be reliable at the .89 level. No individual reliability coefficients were provided for each of the three groups of judges. The authors concluded that the Consensual Assessment Technique "has a potential to be used as an alternative to more problematic criteria-based approaches" (p. 286). The authors suggested that it was possible that experienced judges may take into account the factors included in a criteria-based assessment method "while subjectively rating musical compositions on the single dimension of creativity. However, further research is required in order to ascertain fully the psychometric validity and reliability of this approach" (p. 286).

Webster and Hickey (1995) discussed the idea that decisions on what approaches are used to assess students' compositions should be based "in large part on whether the focus of the work is on the objective analysis of content in the products of composition or if the goal is to make some kind of overall quality judgment about the product" (p. 28). In their study, Webster and Hickey (1995) wanted to provide information about interjudge reliability and concurrent validity when using a set of scales that were based on two styles (explicit and implicit), and two types of musical content (specific and global).

Webster and Hickey (1995) asked 10 fifth and sixth-graders to compose a piece of music. Four music educators were asked to rate the compositions using two different rating forms. Webster and Hickey (1995) found that "the rating scale which contained items that were very open-ended and implicit in nature was in fact extremely reliable, and in the case of global music content items, significantly more reliable than subscores from the form which was explicit" (p. 36). The authors concluded that the use of rating forms that were very explicit in the definitions of music qualities did not contribute to a better understanding of the constructs of originality, creativity, and aesthetic value. "The subscores of global ratings from the implicit form prove to be the most predictive for the constructs of the originality/creativity qualities and aesthetic value of children's compositions" (p. 36).

In an earlier study, Webster (1977) intended to establish a set of valid and reliable criteria measures for the assessment of potential creative thinking ability during composition, improvisation, and analysis experiences. He also investigated the relationship between these criteria and selected measures of musical ability, general

creative ability, and general intelligence. Webster looked at the relationships between these criteria measures and the variables of grade level, age, sex, performance medium, and years of piano study.

For the purposes of his study, Webster (1977) assumed that there were three separate kinds of musical creative products: (1) compositions, (2) improvisation, and (3) analysis. Webster designed three separate measures that reflected each mode independently. Each set of measures assessed fluency, flexibility, elaboration, and originality for each kind of musical creative product.

Webster defined musical fluency in composition as the ability to generate a number of clearly defined musical ideas through a notational system. Musical flexibility in composition was defined as the ability to generate a number of different ideas through a notational system. Musical elaboration in composition was defined as the ability to extend, reshape, and refine musical ideas through a notational system. Musical originality in composition was defined as the ability to create a musical composition containing elements of unique expression.

For the mode of improvisation, Webster defined musical fluency as the ability to respond freely to musical ideas as a performer; musical flexibility as the ability to generate a number of different ideas through performance; musical elaboration as the ability to extend, reshape, and refine musical ideas as a performer; and musical originality as the ability to create an improvisation containing elements of unique expression.

For the mode of analysis, Webster defined musical fluency as the ability to generate a number of clearly expressed ideas about the structure of music; musical

flexibility as the ability to generate a number of different ideas about the structure of music; musical elaboration as the ability to add to or extend the original ideas; and musical originality as the ability to seek out and find relationships or observations that represent unique thinking.

In his analysis of a pretest study using a test-retest approach, Webster found levels of interjudge reliability for the three measures of musical creativity ranging from .75 to .85. A panel of composers, performers with improvisation skills, and theorists, established the content validity of the test. Face validity was established during the pretest phase of the study. When comparing the three musical creativity measures with other variables, Webster found that music achievement strongly correlated with creative potential. Figural creativity was significantly related to both improvisation and analysis. I.Q. was significantly related only to improvisation. Age and grade level had no significant relationship to any of the three musical creative modes.

Using Dewey's (1910) definition of reflective thinking as synonymous to creativity, Laycock (1992) investigated whether relationships existed between the musical characteristics of original compositions produced by high-school students, aged 15 to 18 years old, and their musical experience, musical aptitude, self-concept as musicians, age, and academic achievement as represented by their GPA. The researcher viewed composition as a problem-solving task. In the study, subjects had at least 10 minutes and no more than 20 minutes to compose and rehearse an original melody on an acoustic piano. The subjects were asked to start on middle C, and to use only the white keys on the keyboard in order to avoid harmonic conflicts.

Once the subjects finished their composition, they were asked to perform their melody twice while the researcher recorded them. These two performances were analyzed and graded by the researcher and two other teachers. The grading was based on the following concepts: tonal strength as demonstrated by key centrality; metrical coherence; rhythmic and melodic motive variation and coherence; phrasing; originality, as demonstrated by melodic and rhythmic content “entirely new to the norm for the population of the experiment” (p. 70); general cohesiveness; ability to reproduce the composed melody; and complexity. Laycock found that the image that participants had about themselves as musicians was “the variable most frequently correlated with the musical characteristics of the original composition” (p. 117). From 14 variables, 12 were correlated with self-concept, 10 with in- and out-of-school musical experience, 8 with aptitude, 5 with age, and 2 with GPA.

In another quantitative effort to explore creative musical problem solving, Kratus (1989) examined the amount of time that children of different ages, genders, and proficiency levels spent on various compositional processes while creating a melody. Based on Webster’s ideas (Webster, 1987), Kratus believed that the creative process alternated between convergent and divergent thinking. Kratus theorized that the transition between convergent and divergent thinking took place through four stages: preparation, incubation, illumination, and verification.

Kratus (1989) believed that the phase of preparation was equivalent to the exploration that takes place during the compositional processes of children and adults. In this phase, composers improvise music patterns subconsciously. In a conscious phase, equivalent to the idea of incubation, the composer elaborates and develops

musical ideas through the application of musical principles such as association, contrast, and balance. The phase of illumination takes place as the relationship between the parts and the whole become clear for the composer. Finally, the composer verifies the work through the repetition of selected rhythmic and melodic patterns.

In order to explore the ways in which these four phases take place during children's compositional processes, Kratus (1989) randomly selected 60 children, ages 7, 9, and 11, from an elementary school. The participants were asked to compose a song using an electronic keyboard. The restrictions were that subjects could only use white keys, begin on middle C, and have 10 minutes to compose their songs. After 10 minutes, participants were asked to play their compositions twice. For the analysis of the data, the 10-minute sessions were divided into 120 intervals of five seconds each. Each interval was categorized as being exploration, if the music sounded unlike anything played earlier; development, if the music sounded similar yet different from anything played earlier; repetition, if the music sounded the same as music played earlier; or silence, if no music was heard because the participant was silent or because the participants or the researcher were talking. Two independent judges were in charge of evaluating the degree to which each participant's song and its replication sounded alike.

Kratus (1989) found that as children become older, repetition is more prevalent during their compositional processes. According to the author, "the creative act of composition for the 7-year-olds was very similar to the act of improvisation, and they used compositional time to explore new ideas rather than modify ideas" (p. 17). Kratus found that 9- and 11-year-old participants spent more time developing their ideas and

significantly less time on exploration than 7-year-olds did. Comparing his results with those of other authors, Kratus concluded that 9 and 11-year-olds use compositional processes “that are similar to those used by adult composers” (p. 17). Kratus also concluded that his results support the idea that participants’ emphasis on exploration at the beginning of the composition period “is indicative of *preparation*; the emphasis on development is indicative of *incubation*; and the emphasis on repetition toward the end of the period is indicative of *verification*” (p. 17). Kratus did not, however, find these stages as discrete and clearly defined. They seemed to be intermingled processes, with one process predominating at various times.

In reference to the replication of their compositions, Kratus (1989) found that 7-year-old participants, who had spent little time developing their ideas, were significantly less able to replicate their compositions than 9- and 11-year-olds. Kratus (1989) concluded that participants who understood that in order to solve the problem they needed to repeat the melody they were composing were product oriented. “On the other hand, the subjects who were unable to replicate their songs...were process oriented” (p. 18). Finally, Kratus concluded that 7-year-olds can engage in creative musical improvisation, and that 9- and 11-year-olds can meaningfully compose and shape their musical ideas.

In a later study, Kratus (1994) examined the relationship among music audiation, the process of composition, and the musical characteristics of songs composed by 9-year-olds. Kratus selected 40 third-graders who were 9 years old at the time of the testing, and who had not taken piano lessons for longer than a year. To measure their audiation skills, Kratus administered the Intermediate Measures of Music Audiation

(Gordon, 1982). For this study, Kratus used a compositional task similar to that used in the 1989 study, in which the author gave the subjects 10 minutes to compose a song. The only difference was that, for this study, Kratus connected the electronic keyboard to a computer running Performer 2.31, which provided a transcription of the composed piece. Similar to the 1989 study, two judges determined when the subjects were exploring, developing, repeating, or silent during the compositional processes. Two other judges rated the compositions' tonal and melodic cohesiveness, and repetition or development of melodic and rhythmic patterns.

Based on the obtained results, Kratus concluded that audiation was related to the subjects' sense of tonality and meter. Subjects with higher levels of audiation tended to spend less time in exploration. According to Kratus (1994), "Too much exploration does not, in itself, produce a musical result.... Tests of musical creativity measure primarily divergent behaviors, like exploration, but composition apparently requires other behaviors as well" (p. 127). In contrast, the positive correlation between the process of development and audiation led Kratus to think that audiation allowed subjects to think about a musical idea before developing it. There was also a positive correlation between audiation and the use of silence. Kratus noticed that, during silence intervals, children tended to move their fingers over the keyboard without pressing the keys, as if reviewing their songs silently. In regard to the final products, Kratus found that subjects who spent more time on exploration created songs that were less structured and more extensive than did subjects who spent more time on development, repetition, and silence.

In another study, Kratus (2001) investigated the effect that different melodic configurations of an Orff xylophone had on fourth-graders' compositional processes and products. The underlying principle in this study was that subjects needed some restrictions in order to be able to create, because the lack of restrictions could make subjects' creative choices more difficult. Since Kratus had earlier found (Kratus, 1989, 1994) that children younger than 9-years-old had more difficulty composing songs that they could replicate, he selected 48 fourth-graders, ages 9 to 10, for this study. Subjects were randomly assigned to one of four groups of equal size. The first group composed on a xylophone that had a pentatonic scale with 5 bars (C-D-E-G-A). The second group's xylophone had a pentatonic scale with 10 bars (C-D-E-G-A-C'-D'-E'-G'-A'). The third group's xylophone had a harmonic minor scale with 5 bars (C#-D-E-F-G). The fourth group composed on a xylophone that had a harmonic minor scale with 10 bars (C#-D-E-F-G-C#'-D'-E'-F'-G').

As in previous studies (Kratus 1989, 1994), Kratus tested subjects individually, giving them 10 minutes to compose their songs. He found that no subject required the full 10 minutes. Following the procedures that he used in previous studies, Kratus divided the time spent on the compositional processes into five-second intervals in order to determine how much time was spent exploring, developing ideas, repeating ideas, and silent. The characteristics of the composed products included in the analysis were: length of the song in seconds, final pitch of the song, tonal cohesiveness, metric cohesiveness, the use of melodic patterns, the use of rhythmic patterns, and the subjects' ability to replicate the song. Except for the length of the song, which was

determined by the researcher, two independent judges evaluated all the variables using a 7-point rating scale.

Kratus (2001) found that a more limited number of pitch options resulted in less use of exploration, shorter songs, and songs that were more readily replicated. The available tonalities and pitch options had no effect on the compositional processes of development, repetition, or silence, or on the cohesiveness or use of patterns in the final compositions. Kratus also found that subjects composing on diatonic xylophones were three times as likely to end on the starting tonic pitch (D in this case) than were subjects working on pentatonic configurations. Kratus suggested that “music educators may wish to include the leading tone in composing projects as a means to encourage students to develop a stronger sense of tonality” (p. 304).

Using Kratus’ measurement instruments of compositional processes and products, Henry (1995) investigated if the compositional processes and products of fourth-grade students were affected by subjects’ music aptitude and different instructional methods. Music aptitude was measured using the Intermediate Measure of Music Aptitude (Gordon, 1982). For the treatment, one class received repeated composing opportunities on a keyboard and pattern instruction; another received repeated composing opportunities on a keyboard, but no pattern instruction; the third class received only pattern instruction; and the fourth class served as the control group. The treatment took place during a 12-week period. After the treatment sessions, subjects met with the researcher, and were asked to compose an original song. The subjects had 10 minutes to compose their songs. After the 10 minute session, subjects were asked to play their compositions twice.

For the analysis of the data, the 10-minute sessions were divided into 120 intervals of 5 seconds each. Two independent judges categorized each interval as exploration, development, repetition, or silence. Each category was totaled and averaged to represent the amount of time each subject spent in each activity. Using a seven-point Likert scale, two independent judges analyzed the compositional products, and rated them for melodic cohesiveness, metric cohesiveness, and replication. The two judges also evaluated the use of melodic and rhythmic patterns in the compositions. For the data analysis, the researcher used an ANOVA for each independent variable.

Based on his results, Henry (1995) concluded that repeated keyboard compositional activities combined with pattern instruction seemed to be the most effective instructional method. Pattern instruction alone seemed to have a positive effect on the compositional process, but the effect was not significant when combined with repeated composing opportunities.

CHAPTER III

METHODOLOGY

Several considerations were important in determining the present study's design, number of subjects, experimental procedures, and data collection. First, the sample size had to be sufficient to approximate normality, and at the same time it had to respond to the schedule, time and location limitations of the host school. Second, in response to the request made by the host school's principal, the study needed to take place only during the 50 minute general music classes. Third, there were only four locations at the host school where the experimental sessions and the pre- and posttests could take place: the second grade, third grade, fifth grade, and special education work rooms.

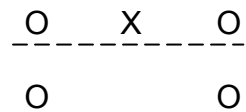
As it will be described later in this chapter, the pre- and posttests were designed to last a maximum of 25 minutes per subject. With these limitations on time and space, only 4 subjects could record their pre or posttest at the same time, which left space and time for the recording of eight individual tests during each music lesson period. Fourth, in order to avoid a threat to the internal validity of the study's design, the posttest ideally needed to take place the week after the subjects participated in the treatment. Therefore, it was decided that only 8 subjects would be randomly chosen from each class to participate in the study. The study's design, sample, statistical procedures, and independent and dependent variables, will be described in the following sections of this chapter.

Design

The selected design for this study was the nonequivalent control group design (Campbell & Stanley, 1963). This quasi-experimental design is used in instances in

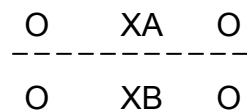
which there is a pretest and a posttest, but where there is not a possibility for true randomization, when there are not formal means “of certifying that the groups would have been equivalent had it not been for the X” (Campbell & Stanley, 1963, p. 12).

Campbell and Stanley (1963) diagramed this design as follows:



Campbell and Stanley (1963) also specified that the second group in this design is not necessarily a control group in which there is not a treatment. It can be a group in which another form of treatment will take place. That is the case for the present study.

In this case, the diagram of the study would be better represented as follows:



Even though subjects participating in the study were pulled out from intact classrooms, the selection for participation in the study and the assignment of either one of the two treatments were random and under the experimenter’s control, as indicated by Campbell and Stanley (1963).

A consent letter was sent to every parent. In the letter, the parents were asked to give their permission to have their child participate in a study exploring creative problem solving through musical composition. The letter stressed the importance for students to participate in every stage of the study, as well as their freedom to drop from the study at any moment they considered there was a need. The letter also explained that the activities that the students would be involved in for the study were part of the music curriculum of their grade level, and were specified in the National Standards for Arts

Education (MENC, 1994). Therefore, the students would not be losing valuable time from their music learning, but continuing with their musical growth while participating in the study. Finally, the letter explained that, because of time and space restrictions, subjects participating in the study would be randomly chosen from each class, and that students who did not participate in the study would participate in similar activities with their music teacher while the randomly chosen students participated in the study's activities. The consent letter is displayed in Appendix A.

The pretest took place during the first week of the study. The subjects were measured on an individual basis. There were 3 weeks of treatment, during which 4 subjects from each class participated in treatment A and 4 in treatment B. The posttest took place during the 5th week of the study. The activities that took place during the treatment sessions will be described in the section in this chapter devoted to the independent variable. The pretest and posttest will be described in the section in this chapter devoted to the dependent variable.

Sample Size

Subjects participating in this study were 32 fourth graders and 32 fifth graders attending general music classes at one elementary school, making a total of 64 subjects. The subjects were randomly chosen from eight different classes receiving general music instruction from one music educator. There were 8 subjects randomly chosen from each of the eight classes. Four of these classes were fourth grade, and four were fifth grade. The selection procedures took place as follows: the subjects whose parents returned the consent letter were listed in alphabetical order. Using the computer program SPSS® 10.0 for Windows® 1998, 8 subjects were randomly chosen

from each of the eight classes. Hence, there were four groups of 8 fourth graders and four groups of 8 fifth graders participating in the study, making a total of 64 subjects. A graphic description of the sample is displayed in figure 2. The sampling procedures will be described in more detail later in this chapter.

Host School

Three public elementary schools with established general music programs were identified as possible sites for the study. Each of the three schools had only one general music teacher. The three music teachers were interviewed to find out which site could provide the space and time requirements for the study. All music teachers received a description of the procedures, and of the space and time requirements of the study. Two of the music teachers could not find the time to schedule the study. However, both of these teachers agreed to participate in the pilot study. The school chosen for the study was located in a suburban area in northern Texas. The principal of the school agreed to participate in the study, and offered the use of five working rooms. The researcher was asked to limit the students' participation in the study to the hours of their music classes.

Subjects

The population attending the host school was predominantly middle class. Approximately 74% of the population were white Americans, 12% African-Americans, 7% Hispanics, and the other 7% were Asians and Middle Easterners. There was only one general music educator in the school who taught all of the kindergarten to fifth grade classes. The groups participating in the present study were receiving a 50 minute music lesson once a week. Fifth graders met from 8:00 a.m. to 8:50 a.m., and fourth graders met from 9:00 a.m. to 9:50 a.m. Fourth A and fifth A met on Monday; fourth B

and fifth B on Tuesday; fourth C and fifth C on Wednesday; and fourth D and fifth D on Thursday. According to the music teacher, these students knew basic musical theoretical concepts, such as rhythmic values of quarter, eighth, half, and whole notes and rests, and the concepts of contrast, repetition, melodic contour, dynamics, and phrase. The teacher felt that most subjects were beat competent, could match pitch, and read successfully in either letter names or solfège. Finally, the teacher reported that the subjects had not been exposed to composition or improvisation activities since she was appointed as their music teacher a year and a half prior to the present study.

The Task

The task for the current study was intended to give participant subjects the opportunity to exercise their creativity as much as possible. With that goal in mind, the ideas suggested by Getzels and Csikszentmihalyi (1975) were considered an important foundation for the design of the task. Getzels and Csikszentmihalyi (1975) identified three types of problems. In the first type, the problem is known as well as the solution method. The problem solvers only need to look for the solution. According to Getzels and Csikszentmihalyi, and DeLorenzo (1987) whose findings support this idea, this type of problem does not encourage high levels of creativity because the problem solvers only need to make use of their memory and retrieval. In a second type of problem, the problem is clearly defined, but no method of solution is known by the problem solvers. For this type of problem, the problem solvers need to make use of their reason and rationality. In the third type of problem, the general task is presented but the specific question or problem must be discovered by the problem solvers. According to Getzels and Csikszentmihalyi (1975), the third type of problem requires the problem solvers to

not only make use of their reason and rationality as with the second type of problem, but also use their imagination and creativity so that they can find, articulate, and solve the problem. Since the primary focus of the independent variable of the present study is on creative problem solving skills, the task was designed taking into account the characteristics of the third type of problem, as described by Getzels and Csikszentmihalyi (1975).

The task was also designed to guide the subjects into the exploration of the three stages of creative problem solving as described by Treffinger, Isaksen, and Dorval (1994): Understanding the Problem for the first treatment session, Generating Ideas for the second treatment session, and Planning for Action for the third treatment session. The task was designed in such a way that the subjects were encouraged to explore each of these three stages through questions asked to the subjects during each treatment session. A graphic depiction of the study's design can be found in Appendix C. The researcher and research assistants were only in charge of posing the questions. They did not provide any guidance in order to avoid a teaching-learning situation that would have constituted an extraneous variable to the study.

Independent Variable

The current study explored the effects of creative musical problem solving collectively versus individually on students' compositional products. The independent variable for this study was the condition under which subjects were exposed to problem solve. For treatment A, subjects were asked to solve a creative problem task individually. For treatment B, subjects were asked to solve a creative problem task in teams of four.

Treatment sessions took place in five different locations. For treatment A, one subject was taken to the second grade work room, another to the third grade work room, another to the fifth grade work room, and another to the special education work room. Subjects assigned to treatment B were taken in teams of four to the physical education work room. All of the rooms looked similar. The researcher was in charge of one of the settings, and three other musicians were in charge of the other three settings. The research assistants were instructed to read to the subjects the specific instructions and questions designed for each treatment session. They were asked to avoid giving instructions other than the ones specified, and to answer only the questions related to the specific mechanics of the task. In the case where subjects had questions related to the creative musical product or the creative problem solving process, the researcher and the research assistants would not answer these specific questions. They would repeat the instructions and the specific questions related to the task, and encourage the subjects to think about them and make their own decisions.

It was decided that, since the person in charge of the treatment sessions would basically read the instructions to the subjects, research assistants did not necessarily have to be music educators. The research assistants were a percussionist with a master's degree in jazz; a pianist and a clarinetist at the last stage of their DMA studies; a pianist with a bachelor's degree; and a music educator who was at the early stage of his music education doctoral degree. All of the research assistants, except for the Ph.D. student in music education, were international students. All of them spoke English fluently but with evident foreign accents. The Ph.D. student in music education was male, and was able to assist for only one day. The DMA piano student and the

researcher were females. The other research assistants were males. The ages of the research assistants and the researcher ranged from 25 years old to 36 years old. The researcher attended every day of the study. Once it was determined the days that each of the assistants could attend, the assistants were randomly assigned to the treatment and testing sessions. An example of treatment and testing session assignments is displayed in figure 3.

For treatment sessions 1, 2 and 3, there was a two-octave chromatic song bell RBI RB 2206 and either a tape recorder Tascam™ 424 or a Sony® minidisk MZ-R700, both with a unidirectional microphone, at each of the four different settings. Prior to the study, the recording devices were tested by the researcher and research assistants, all of whom agreed that the produced recordings were very similar in quality. Each chromatic song bell had a red tag marking middle C. The recordings of the treatment sessions were implemented to familiarize subjects with the presence of recording equipment in the room. Subjects also had paper and pencil at hand, but were neither asked nor forbidden to use them. A digital clock was placed where the subjects could clearly see it, so that they could keep track of the time. For treatment session 1, the researcher and research assistants read the following instructions to each subject:

I would like you to compose a piece of music on your bells. It will be a brand new piece of music that no one has ever heard before. There are no right or wrong answers to this task. There are three rules. First, you will only have ten minutes to work on your piece of music. As soon as you start working on your piece of music, I will start this clock, and you will have to stop working when ten minutes have passed. Second, your piece of music must start on middle C, which is marked with the red tag. Third, you will play your piece of music twice for me at the end. The tape recorder will be on the whole time. If you finish your piece of music earlier than the ten minutes, let me know by saying "I am finished." Then, I will ask you to play the piece of music for me twice, and we will record it. Do you have any questions? Before you start working, I would like for you to think about all of your options. What types of melodies can you compose and comfortably

perform? Are there any difficulties that you have to consider? You do not have to tell me your answers. Just think about it before you start working on your composition, and tell me when you are ready to start.

The last questions posed by the researcher were intended to lead the subjects into the problem solving stage that Treffinger, Isaksen, and Dorval (1994) called Understanding the Problem. In this phase, the subjects analyze the situation, consider musical options, technical concerns, and any other circumstances involved in the composition of a final product. At this point, the subjects also find the specific question on which to focus their subsequent efforts. The person in charge was not to discuss subjects' answers to these questions in order to avoid a teaching-learning situation. Since there were four different persons in charge of the treatment sessions, this would have added an extraneous variable to the study.

Treatment sessions 2 and 3 followed the same procedures as those followed for treatment session 1. However, there were changes in the questions posed by the researcher and research assistants. During treatment session 2, the subjects were told:

I would like you to compose a piece of music on your bells. It will be a brand new piece of music that no one has ever heard before. There are no right or wrong answers to this task. There are three rules. First, you will only have ten minutes to work on your piece of music. As soon as you start working on your piece of music, I will start this clock, and you will have to stop working when ten minutes have passed. Second, your piece of music must start on middle C, which is marked with the red tag. Third, you will play your piece of music twice for me at the end. The tape recorder will be on the whole time. If you finish your piece of music earlier than the ten minutes, let me know by saying "I am finished." Then, I will ask you to play the piece of music for me twice, and we will record it. Do you have any questions? Before you start working, I would like for you to think about all of your options. What types of melodies can you compose that are different, but all start on middle C? How can you make the melodies sound like phrases and not just random sounds? Is there anything that you can change besides the melody? You do not have to tell me your answers. Just think about it before you start working on your composition, and tell me when you are ready to start.

The questions added for the second treatment session were intended to familiarize the subjects with the second phase of the creative problem solving process described by Treffinger, Isaksen, and Dorval (1994), Generating Ideas. In this stage, the subjects find specific musical ideas. This could be accomplished by varying the musical elements.

During treatment session 3, the subjects were told:

I would like you to compose a piece of music on your bells. It will be a brand new piece of music that no one has ever heard before. There are no right or wrong answers to this task. There are three rules. First, you will only have ten minutes to work on your piece of music. As soon as you start working on your piece of music, I will start this clock, and you will have to stop working when ten minutes have passed. Second, your piece of music must start on middle C, which is marked with the red tag. Third, you will play your piece of music twice for me at the end. The tape recorder will be on the whole time. If you finish your piece of music earlier than the ten minutes, let me know by saying "I am finished." Then, I will ask you to play the piece of music for me twice, and we will record it. Do you have any questions? Before you start working, I would like for you to remember all of the problems that you found in our previous sessions, and think about the ways in which you could avoid these problems. How can this composition be better than the ones you did in our previous sessions? You do not have to tell me your answers. Just think about it before you start working on your composition, and tell me when you are ready to start.

The questions posed during treatment session 3 were intended to familiarize the subjects with the third and last phase of the creative problem solving process described by Treffinger, Isaksen, and Dorval (1994), Planning for Action. In this stage, the subjects examine promising options to determine the steps needed to achieve the ultimate goal of solving the problem. This stage also involves acceptance-finding, which helps the problem-solver identify ways to make the best possible use of assisters and avoid possible sources of resistance. Acceptance-finding also involves formulating an action plan that describes the specific steps that are taken in order to implement a proposed solution.

Once the directions had been read and all the questions answered, the researcher turned the tape recorder on, and moved to another part of the room. Once the subject indicated that s/he had finished, or the allowed time had passed, the subject played the piece of music. After the subject played the piece of music, the recording was turned off, and the researcher thanked the subject for participating, and walked the subject back to the music classroom.

For treatment B, subjects were asked to solve the same creative problem solving task used for treatment A, but subjects assigned to treatment B solved the problem in teams of four instead of individually. For treatment sessions 1, 2, and 3, the researcher or the research assistants pulled out the four subjects assigned to this treatment from the classroom during the time their music lesson was taking place. The subjects were taken to the physical education work room, where four two-octave chromatic song bells RBI RB 2206 were placed. The bells were set in a circle, so that subjects could face each other. A digital clock was placed in front of each subject. There was either a tape recorder Tascam 424 or a Sony minidisk MZ-R700, both with a unidirectional microphone. As in treatment A, subjects participating in treatment B also had paper and pencil at hand, but were neither asked nor forbidden to use them. There were also differences in the instructions provided during each of the three treatment sessions, which were related to the creative problem solving processes. For treatment session 1, the researcher told the subjects:

I would like you to work together and compose a piece of music on your bells. It will be a brand new piece of music that no one has ever heard before. There are no right or wrong answers to this task. There are three rules. First, you will only have ten minutes to work on your piece of music. As soon as you start working on your piece of music, I will start this clock, and you will have to stop working when ten minutes have passed. Second, your piece of music must start on

middle C, which is marked with the red tag. Third, you will play your piece of music twice for me at the end. The tape recorder will be on the whole time. If you finish your piece of music earlier than the ten minutes, let me know by saying “we are finished.” Then, I will ask the four of you to play the piece of music together for me twice, and we will record it. Do you have any questions? Before you start working, I would like for you to think about and discuss with your partners all of your options. What types of melodies can you compose and comfortably perform? Are there any difficulties that you have to consider?

Once the directions had been read and all the questions answered, the researcher turned the tape recorder on, and moved to another part of the room. Once the subjects indicated that they had finished, or the allowed time had passed, the 4 subjects played the piece of music together. After the subjects played the piece of music, the recording was turned off, and the researcher thanked the subjects for participating, and walked them back to the music classroom.

Treatment sessions 2 and 3 followed the same procedures described above, but the questions addressed collective processes instead of individual processes. For treatment session 2, subjects were told:

I would like you to work together and compose a piece of music on your bells. It will be a brand new piece of music that no one has ever heard before. There are no right or wrong answers for this task. There are three rules. First, you will only have ten minutes to work on your piece of music. As soon as you start working on your piece of music, I will start this clock, and you will have to stop working when ten minutes have passed. Second, your piece of music must start on middle C, which is marked with the red tag. Third, you will play your piece of music twice for me at the end. The tape recorder will be on the whole time. If you finish your piece of music earlier than the ten minutes, let me know by saying “we are finished.” Then, I will ask the four of you to play together the piece of music for me twice, and we will record it. Do you have any questions? Before you start working, I would like for you to think about and discuss with your partners all of your options. What types of melodies can you compose that are different, but all start on middle C? How can you make them sound like phrases and not just random sounds? Is there anything that you can change besides the melody?

For treatment session 3, subjects were told:

I would like you to work together and compose a piece of music on your bells. It will be a brand new piece of music that no one has ever heard before. There are no right or wrong answers to this task. There are three rules. First, you will only have ten minutes to work on your piece of music. As soon as you start working on your piece of music, I will start this clock, and you will have to stop working when ten minutes have passed. Second, your piece of music must start on middle C, which is marked with the red tag. Third, you will play your piece of music twice for me at the end. The tape recorder will be on the whole time. If you finish your piece of music earlier than the ten minutes, let me know by saying “we are finished.” Then, I will ask the four of you to play together the piece of music for me twice, and we will record it. Do you have any questions? Before you start working, I would like for you to remember all of the problems that you found in our previous sessions, and think about the ways in which you could avoid these problems. Discuss with your partners how can this composition be better than the ones you did in our previous sessions.

Each session lasted no more than 25 minutes, including the time required to go from the music classroom to the treatment room and back to the classroom. Since the treatment sessions took place during the 50-minute music class, there was time to meet with one group of 4 subjects working together, and 4 subjects working individually. Subjects were randomly assigned to a person in charge. The researcher and research assistants were randomly assigned to a room, unless they had been assigned the collective condition. In this case, the person in charge was assigned to the physical education work room, and the other persons in charge were randomly assigned to one of the other four locations. Because the collective condition required more equipment, and because there was only enough time for the treatment sessions, it was decided that treatment sessions for the collective condition would always take place at the same location. There were four persons in charge each day, one of whom was assigned to two treatment sessions per class, so that there was enough time for all the subjects to participate in the treatment sessions.

Sampling Procedures

The week of the pretest constituted the first week of the study. All 64 participating subjects took the pretest and posttest individually. The researcher hired and trained five research assistants to help her with the testing and with the treatment sessions, so that there would be one person in charge at each location. The week before the experiment started, the researcher met with the research assistants. During this session, the researcher explained to the assistants the procedures they would follow and the functioning of the equipment to be used. The assistants were told that the researcher would be in charge of one of the locations, and, depending on their schedules and the random assignments, each of the assistants would be in charge of the other three locations. They were also told that they were to memorize the instructions they would give to the subjects, and that they should avoid giving instructions other than the ones specified by the researcher. The researcher and the assistants practiced using the equipment, making sure the equipment was working appropriately and that all of them could manage the equipment efficiently. They also checked that the recordings produced by the tape recorders Tascam 424 and the recordings produced by the Sony minidisks MZ-R700 were qualitatively equivalent.

The weekend before the experiment began, the researcher made the random assignments [by using the computer program SPSS 10.0 for Windows 1998]. First, 4 subjects from each group were randomly assigned to treatment A, and the other 4 subjects were assigned to treatment B. An example of treatment and testing session assignments is displayed in figure 3. Then, the researcher and research assistants were randomly assigned to a different room for each of the days they would participate in the

study. One individual was assigned to the second grade work room, one to the third grade work room, one to the fifth grade work room, and one to the special education work room. The two tape recorders, Tascam 424, and the two Sony minidisks MZ-R700, were randomly assigned to one of the four rooms. In order to avoid logistic complications, the recording equipment was assigned to a specific room, and stayed at that location during the duration of the experiment.

All of these rooms had a very similar look. Therefore, it was decided that the researcher and research assistants would work in the same room during one testing day. In cases in which subjects did not use the entire 10 minutes to compose the piece of music, the person in charge could therefore, proceed to the testing of the next subject, without having to wait until someone else was done with the testing in another room.

Subjects were randomly assigned to a person in charge for each of the days that they would participate in the study. The order in which the subjects were going to participate in the study was also randomly assigned. In summary, the researcher randomly assigned who would be in charge of the pretest, posttests, and treatment sessions of each subject, as well as the order in which each subject would be tested or receive the treatment. An example of treatment and testing session assignments is displayed in figure 3. Every morning during the pretest week, as well as during the rest of the experiment, the research assistants received Comment Sheets containing the names of the subjects that they would be working with, the order at which the subjects would be tested that day, and the location at which they would be working that day. Examples of the Comment Sheets can be seen in figure 4.

Pretest Procedures

There was a tape recorder with a unidirectional microphone placed in front of a two-octave chromatic song bell in each room. A clock was also placed where the subjects could clearly see it. There were paper and pencils on hand, but subjects were neither asked nor forbidden to use them. The person in charge told the subjects:

I would like you to compose a piece of music on your bells. It will be a brand new piece of music that no one has ever heard before. There are no right or wrong answers to this task. There are three rules. First, you will only have ten minutes to work on your piece of music. As soon as you start working on your piece of music, I will start this clock, and you will have to stop working when ten minutes have passed. Second, your piece of music must start on middle C, which is marked with the red tag. Third, you will play your piece of music twice for me at the end. If you finish your piece of music earlier than the ten minutes, let me know by saying "I am finished." Then, I will ask you to play the piece of music for me twice, and we will record it. Do you have any questions?

When the subject finished, or after the ten minutes passed, the person in charge started the recording of the compositional product. Each subject's tape/minidisk was assigned a code number to ensure anonymity. The person in charge started the recording by saying "Subject number x, first time," before the subject performed the composition for the first time, and "second time," before the subject performed the composition for the second time. This allowed for the identification of each of the two performances on the recording. It was decided that the first performance would be used for the analyses because of its proximity in time to the compositional process. It was possible that subjects would more accurately remember the decisions they made during the compositional process right after it had taken place, and that they could introduce new musical ideas during a second performance. Kratus (1989, 1994, 2001) and Henry (1995) followed this procedure.

Once the first subject finished the pretest, the person in charge took the subject back to the music classroom. The next subject was then called and taken to the location assigned to that person in charge, to record the pretest. The same procedures were followed with the eight groups until all 64 subjects had recorded their pretest. The recordings of the pretest took no more than 25 minutes per subject. Since 4 subjects were recording their pretest at the same time, it was possible to record 8 subjects from a fourth grade class, and 8 subjects from a fifth grade class every day, making a total of 16 subjects per day. An example of treatment and testing sessions assignments can be seen on figure 3. At the end of the four days, Monday through Thursday, all 64 subjects had recorded their pretests.

For the pretest, subjects were assigned a code number. Once the 64 compositional product recordings were coded, the researcher copied the first performance of the compositions on a single compact disc, following the subjects' code numbers ordered from 1 to 64. In order to facilitate the identification of each subject's composition, the researcher included in the recording the moment in which the person in charge said the subject's code number before the recording of each composition. Once the 64 pretest compositions were recorded on a single compact disc, the researcher made three copies of this compact disc, so that each judge could receive one copy to work with. For the purpose of making the evaluation process more feasible for the judges, all of the compositions were transcribed from the same compact disc that the judges received. Once transcribed, the compositions were printed by using the computer program Finale 2000 for Windows 1998. Then, the Instructions Letter, the compact discs, the compositions' transcription, and the Judge's Scoring Forms were

distributed among the three judges. The Instructions Letter and the Judge's Scoring Form used in the study are displayed in Appendix D.

Treatment Sessions 1, 2, and 3 Procedures

The week after the pretest took place, the treatment sessions started. On Monday of that week, the research assistants received Comment Sheets containing the names of the subjects that they would be working with, the order in which the subjects would participate that day, and the location at which they would be working for the day. Examples of the Comment Sheets can be seen in figure 4. Treatment sessions lasted no more than 25 minutes, including the time required to go from the classroom to the indicated room. For both conditions, subjects had only 10 minutes to work on their compositions. The rest of the time was spent on explanations and on the performance of the compositions. Once the first subject or team had received the treatment, the person in charge took them back to the music classroom. The next subjects was then called and taken to the location assigned to that person in charge. The same procedure was followed during the next hour with the fourth graders, and the following days with the rest of the groups.

During one music class, there was time for four treatment A sessions and one treatment B session. Eight treatment A sessions and two treatment B sessions (with two teams of 4 subjects each) took place in one day, making a total of 16 subjects participating in the study per day. After four days, Monday through Thursday, all 64 subjects participated in the first treatment session. Monday through Thursday of the following week, the subjects participated in the second treatment session, following the same procedures of the first treatment session. Monday through Thursday of the

following week, the subjects participated in the third treatment session, following the same procedures of the first treatment session. After 3 weeks, each of the 64 subjects had participated in three treatment sessions, with 32 subjects participating in treatment A, and 32 subjects participating in treatment B.

Posttest Procedures

The recording of the posttest took place from Monday through Thursday during the 5th week of the study. The procedures followed for the posttest were exactly the same as those followed for the pretest. As described earlier in this chapter, the weekend before the experiment began, the researcher made the random assignments [by using the computer program SPSS 10.0 for Windows 1998]. The researcher and research assistants were randomly assigned to a different room for each of the days the posttest took place. Subjects, rooms, and the order in which the subjects were going to be tested, were randomly assigned. The research assistants received Comment Sheets containing the names of the subjects that they would be working with, the order in which the subjects would be tested that day, and the location at which they would be working that day. Examples of the Comment Sheets can be seen in figure 4.

As in the pretest, subjects had a maximum of 25 minutes to complete their posttest. The person in charge read the instructions to each subject, and allowed 10 minutes for the subjects to work on the compositions. After this time, the subjects were asked to perform their compositions twice. The compositional products were recorded on individual tapes or minidisks.

Once each subject finished the recording of the posttest, the person in charge brought the subject back to the classroom, and took the second subject assigned to that

particular location. The researcher and the three research assistants followed this procedure until the 8 subjects in that class recorded their posttests. The same procedure took place in each of the fifth and fourth grade classes, until all 64 subjects had recorded their posttests.

For the posttests, subjects were assigned a new code number, so that the judges could not identify one specific pretest recording with the same subject's posttest recording. Once the recordings of the 64 compositional products were coded, the researcher copied the first performance of the compositions on a single compact disc, following the subjects' code numbers ordered from 1 to 64. In order to facilitate the identification of each subject's composition, the researcher included in the recording the moment in which the person in charge said the subject's code number before the recording of each composition. Once the 64 posttest compositions were recorded on a single compact disc, the researcher made three copies of this compact disc, so that each judge could receive one copy to work with. For the purpose of making the evaluation process more feasible for the judges, all of the compositions were transcribed from the same compact disc that the judges received. Once transcribed, the compositions were printed by using the computer program Finale 2000 for Windows 1998. Then, the compact discs and the Judge's Scoring Forms were distributed among the three judges. The Judge's Scoring Form used in the main study is displayed in Appendix D.

Dependent Variable

In this study, the dependent variable was the measures obtained from the compositional products. The dependent variable tests were each labeled as subtests for

the purposes of this study. The measurement process followed the procedures used in Kratus' studies (1989, 1994, 2001), and Webster's study (1977). The selection of the measurement instrument responded to the fact that these authors had systematically explored young students' musical compositions. In addition, Kratus (1994) reported interjudge reliability ratings ranging from .80 to .95. In another study, Kratus (2001) reported interjudge reliability ratings ranging from .68 to .83.

For one of his studies, Kratus (1994) based the selection of his variables on Laycock's study (1993). Laycock explored students' compositions in terms of the tonal and metric structure, cohesiveness, originality, rhythmic complexity, phrasing, and use of developed melodic patterns. Kratus (1994) rated the compositions' cohesiveness, pattern use, and extensiveness based on analytical constructs of children's compositions developed by Laycock (1993) in prior investigations. Webster (1977) explored musical fluency, musical flexibility, musical elaboration, and musical originality, based on Guilford's (1956) tests of general creativity. For the purposes of this study, it was decided that between both Kratus' (1994) and Webster's (1977) measurement instruments, structural and creative aspects of students' musical compositions were taken into account appropriately.

Three independent judges, all of whom were general music educators with extended professional experience, analyzed the audiotapes of each subject's final creative problem solving product. The three judges were females. Judge A had 30 years of experience teaching kindergarten to fifth grade, and held a National OAKE Endorsement as a Certified Kodály Teacher. Judge B had 16 years of experience teaching kindergarten to fifth grade students in public elementary schools, had a

Certification in Orff Schulwerk, was teaching music education courses at a major university in north Texas, and was working toward her Ph. D. in Music Education. Judge C had 5 years of experience teaching kindergarten to fifth grade, and was in the process of finishing her National OAKE Endorsement as a Certified Kodály Teacher as well as her master's degree in music education.

For the purposes of this study, the compositional products were evaluated in terms of their pattern use with four subtests, cohesiveness with two subtests, and creativity with two subtests. The first four subtests were related to melodic and rhythmic pattern use, and were intended to be used in one statistical analysis. The measurement scale used the same variables as Kratus (1994), but measured the variables in a slightly different way. In his study, Kratus (1994) indicated the percentage of the total time the judges believed the subjects spent making use of each type of pattern. To obtain optimal measurement objectivity, the present study had the judges indicate the number of times the subjects made use of repeated melodic patterns, developed melodic patterns, repeated rhythmic patterns, and developed rhythmic patterns. The judges indicated this by writing one tally on the Judge's Scoring Form for each time they heard a recognizable melodic or rhythmic pattern being repeated or developed. The Judge's Scoring Form used in the main study is displayed in Appendix D. At the end, the total number of tallies for each type of pattern was summed in order to obtain the total number of occurrences of each type of pattern during the composition. Kratus (1994) defined the use of melodic and rhythmic patterns as follows:

Melodic Pattern: two to seven pitches that form a distinct pitch pattern that is perceived as a unified whole.

Repeated Melodic Pattern: a melodic pattern that is identical to a previous occurring melodic pattern.

Developed Melodic Pattern: a melodic pattern that is similar to, yet different from, a previously occurring melodic pattern. In most cases, the rhythm is held the same, the melody changes, and the melodic contour is similar.

Rhythmic Pattern: two to seven durations that form a distinct duration pattern that is perceived as a unified whole.

Repeated Rhythmic Pattern: a rhythmic pattern that is identical to a previous occurring rhythmic pattern.

Developed Rhythmic Pattern: a rhythmic pattern that is similar to, yet different from, a previously occurring rhythmic pattern. In most cases, the melody is the same or similar and the rhythm changes. (p. 121)

The measures obtained for tonal and metric cohesiveness, were intended to be used in another statistical analysis. For these subtests, a 7-point Likert-scale was used, as suggested by Kratus (1994, 2001), with a 7 representing very strong cohesiveness, and a 1 indicating no cohesiveness. Kratus (1994) defined these terms as follows:

Tonal Cohesiveness: The degree to which the pitches in a composition are constructed around a tonal center or tonal centers. (7=very strong cohesiveness, 1 = no tonal cohesiveness)

Metric Cohesiveness: The degree to which the durations in a composition are constructed of regularly occurring accented and unaccented beats. (7 = very strong metric cohesiveness, 1 = no metric cohesiveness)

The last two subtests, the creativity measures obtained through a fluency and an originality test, were intended to be used in another statistical analysis. These two subtests, fluency and originality, were central in Webster's suggested measurement instrument designed to assess the musical creativity of students' original compositions (Webster, 1977). Webster defined musical fluency as "the ability to generate a number of clearly defined musical ideas" (p. 137), and musical originality as "the ability to create a composition that contains elements of unique expression" (p. 137). Both variables were measured with a 7-point Likert-scale. In the case of musical fluency, a 1 indicated that the composition could not be reproduced, and a 7 indicated that someone else could reproduce the composition without trouble. In the case of originality, a 1 indicated that the composition had no originality at all, and a 7 indicated that the composition had marked originality. The Judge's Scoring Form used in the main study is displayed in Appendix D.

Statistical Procedures

The statistical procedure indicated for a study with the design that the present study followed was a MANCOVA. This responded to the fact that there was more than one dependent variable subtest and there was at least one covariate. The covariate in the present study was the pretest, and was intended to be used in an attempt to statistically control for pre-existing differences between the groups.

There were eight subtests obtained for the compositional products:

1. Repeated melodic patterns
2. Developed melodic patterns
3. Repeated rhythmic patterns

4. Developed rhythmic patterns
5. Tonal cohesiveness
6. Metric cohesiveness
7. Fluency
8. Originality.

The first four subtests, the measures obtained for the repeated melodic patterns, developed melodic patterns, repeated rhythmic patterns, and developed rhythmic patterns, were grouped and intended to be used in a statistical analysis. Kratus (1994) referred to these subtests as grouped under a general concept of pattern use. The next two subtests, the measures obtained for tonal and metric cohesiveness, were intended to be used in another statistical analysis. Kratus (1994) found tonal and metric cohesiveness functioned as a composite in composition, “rather than as discrete tonal and rhythmic components” (p. 127).

The last two subtests, the measures obtained for fluency and originality, were intended to be used for a third statistical analysis. These subtests were grouped on Webster’s (1977) measurement instrument for the general concept of creativity. In his instrument, Webster (1977) also included a measure for elaboration, which he defined as “the ability to extend, reshape, and refine a musical idea” (p. 137). This description was found to be very similar to that provided by Kratus for the developed rhythmic and melodic patterns. Kratus (1994) considered that both melodic and rhythmic patterns could be repeated when they are identical to a previous occurring pattern, or developed when they are similar to, yet different from, a previously occurring pattern. It was decided that Webster’s concept of elaboration, meaning the extension, reshaping, and

refining of music ideas, was already accounted for by the four variables related to the development of melodic and rhythmic patterns included in Kratus' instrument. Therefore, only the concepts of fluency and originality (Webster, 1977) were held for the assessment of creativity.

Once the judges completed the Judge's Scoring Forms for each subject, the researcher used the computer program SPSS 12.0 for Windows XP to carry out the statistical analyses. The first statistical analysis tested the homogeneity of regression of the covariate. As Henson (1998) pointed out, the homogeneity of regression assumption must be met before [ANCOVA] can be accurately utilized" (p. 4). According to Henson (1998):

...the regression slopes of the covariate and the dependent variable in the individual groups must be the same if the single pooled regression slope can be accurately used with all groups. If the individual regression slopes are notably different, then the pooled regression slope will not effectively represent any of the groups! (p. 10).

The results of the initial analysis showed that the homogeneity of regression assumption was not met. Accordingly, the pretest data could not be used as a covariate, and MANCOVA was not the appropriate procedure. Instead of three MANCOVAs, the appropriate procedure was the exploration of the differences between the two groups' means on the posttest with a one-way ANOVA for each of the eight subtests. For this procedure, the pretest was not used as a covariate, which in a MANCOVA would have controlled for initial differences among the groups. Instead, the eight pretests were explored with individual ANOVAs in order to find if there were significant initial differences between the two groups. Because there were eight ANOVAs, the α level

was adjusted with the Bonferroni technique. That is, the original α level (.05) was divided by the number of tests being run (eight). The new α level was set at .006.

Prior to the analyses of variance, the assumptions of normality and homogeneity of variance were explored. The data obtained for the tonal cohesiveness, metric cohesiveness, fluency, and creativity pretests met the assumptions of normal distribution and homogeneity of variance. Therefore, these subtests were each explored with a one-way ANOVA. Since no initial differences were found, it was possible to test the null hypotheses of no differences between the two groups' posttest means on tonal cohesiveness, metric cohesiveness, fluency, and creativity with one-way ANOVAs.

The data obtained for the use of repeated melodic pattern, developed melodic pattern, repeated rhythmic pattern, and developed rhythmic pattern pretests did not meet the assumptions of normality. Therefore, the square root, the log, and the inverse transformation of the data were explored. The transformed data that provided the best approach to normality for each of the subtests' was used for the analyses of variance. Results showed that there were no significant initial differences between the two groups' transformed pretest scores on the four subtests related to pattern use. Therefore, the null hypotheses of no differences between the two groups' posttest means on the use of repeated melodic pattern, developed melodic pattern, repeated rhythmic pattern, and developed rhythmic pattern were tested with a one-way ANOVA for each subtest. The results obtained from these analyses are reported in the next chapter.

Validity

Three independent judges established the validity of the measurement instrument. These three judges had extended experience in research related to

creativity, and to the assessment of students' compositional products. One of these judges had extended experience in research related to the assessment of musical creativity. The other two judges had experience in the assessment of compositional products by using the Judge's Scoring Form designed by Kratus (1994).

The three judges received the Judge's Scoring Form with a detailed explanation of the study's purposes and methodology via e-mail. The Judge's Scoring Forms used in the pilot study are displayed in figures 8 and 9. They were asked to determine if they believed that (a) this measurement tool gave representative breadth of information about compositional products; (b) there was any way to improve the tool to make it more inclusive of compositional product information; and (c) this measurement instrument was appropriate to assess the compositional products of fourth and fifth grade students. One of the judges suggested the replacement of the word *uniqueness* for the word *originality*. He considered that the word *uniqueness* implied a characteristic that compositions can either have or not, but that could not exist only to a certain degree. However, he said this was a valid measurement instrument. In response to his suggestion, the word *uniqueness* was replaced by the word *originality* on the measurement instrument. The other two judges also considered that, for the purpose of this study, the measurement instrument was a valid one.

Pilot Tests Procedures

The pilot test took place in a different elementary school than the host school, with a population socio-economically equivalent to that of the host school. A consent letter was sent to the parents of 2 fourth and 2 fifth grade classes. The Consent for Project Participation Forms are displayed in Appendix E. In the letter, the parents were

asked to allow their child to participate in a study to determine the interjudge reliability of a measurement instrument designed to assess musical composition. The letter stressed the importance for students to participate in the study, as well as their freedom to drop from the study at any moment they considered there to be a need. The letter also explained that the activities that the students would be involved in for the study were part of the music curriculum of their grade level, and were specified in the National Standards for Arts Education (MENC, 1994). Therefore, the students would not be losing valuable time from their music learning, but continuing with their musical growth while participating in the study. Finally, the letter explained that, because of time and space restrictions, subjects participating in the study would be randomly chosen from each class, and that students who did not participate in the study would participate in similar activities with their music teacher while the randomly chosen students were participating in the study's activities. The University of North Texas Committee for the Protection of Human Subjects approved the consent letters. The approval letter is displayed in Appendix F.

Two research assistants were in charge of collecting the data for the pilot test. One of the assistants was a doctoral student in music education and the other was at the end of her master's degree program in music education. These assistants met individually with the researcher the week before the pilot test took place. During this session, the assistants were trained in the use of the equipment. They were informed about the way to connect the microphone into the tape recorder Tascam 424, and the general functions of the tape recorder. They received a copy of the statement that they were to tell the subjects, and asked to memorize it. They were informed about the way

in which they were to place the two-octave chromatic song bell, the tape recorder, the timer, and the paper and pencil, in front of the subjects. Finally, they were asked to arrive at the host school at least half an hour before the class in order to be ready.

The pilot test took place on a Monday and a Tuesday. The fourth and fifth grade groups attending the music class during these days participated in the study. Consent letters signed by the parents were still being turned in during the days the pilot test took place. Therefore, the selection of subjects participating in the study took place the day of the test. From the letters received, each research assistant pulled out two letters from the first fourth grade group, and two letters from the first fifth grade group. Sampling procedures were the same the following day, making a total of 16 subjects participating in the pilot test. Subjects participated in the pilot test in the order in which their letters were pulled out from the stack of consent letters turned in.

Once subjects were chosen, the research assistants took them to the school's stage. One of them worked in a small area behind the stage, and the other one worked on the stage. Even though it was requested that the study take place at the same location both days, in a last minute situation, the school had to use the stage for a different activity during the second day without giving the researcher previous notice. Therefore, the second day the research assistants were placed in two classrooms, with similar lightning, carpeting, and decoration.

For both of the pilot testing days, the research assistants placed the tape recorders with the unidirectional microphones in front of the two-octave chromatic song bell before taking the subjects to the locations. A timer was also placed where the

subjects could clearly see it. There were paper and pencils on hand, but subjects were neither asked nor forbidden to use them. The research assistants told the subjects:

I would like you to compose a piece of music on your bells. It will be a brand new piece of music that no one has ever heard before. There are no right or wrong answers to this task. There are three rules. First, you will only have ten minutes to work on your piece of music. As soon as you start working on your piece of music, I will set this timer to ring in ten minutes, and you will have to stop working when ten minutes have passed. Second, your piece of music must start on middle C, which is marked with the red tag. Third, you will play your piece of music twice for me at the end. If you finish your piece of music earlier than the ten minutes, let me know by saying "I am finished." Then, I will ask you to play the piece of music for me twice, and we will record it. Do you have any questions?

When the subject finished, or once the ten minutes passed, the research assistants started the recording of the compositional product. Each subject's tape was assigned a code number to ensure anonymity. The person in charge started the recording by saying "Subjects number x, first time," before the subject performed the composition for the first time, and "second time," before the subject performed the composition for the second time. This allowed for the identification of each of the two performances on the recording. It was decided that the first performance would be used for the analyses because of its proximity in time to the compositional process. It was possible that subjects would more accurately remember the decisions they made during the compositional process right after it had taken place, and that they could introduce new musical ideas during a second performance. The same procedures were followed with the two groups participating on Monday and the two groups participating on Tuesday, until all 16 subjects had recorded their tests.

Other than the change of location during the second day of the pilot test, the research assistants did not confront any problems. They reported that most subjects

were able to finish their compositions during the provided 10 minutes or even before. From the 8 fifth grade subjects participating in the pilot test, 7 used the provided paper. From the 8 fourth grade subjects participating in the pilot test, 5 used the provided paper. They also reported that all subjects understood the instructions that they gave them, and seemed comfortable using the two-octave chromatic song bell and recording their compositions.

Once the tapes were returned to the researcher, she listed the numbers 1 to 16 using the computer program SPSS 10.0 for Windows 1998. The computer program was used to randomly choose two of the tapes. Subjects with the code numbers five and nine were chosen as samples to train the judges. The other 14 of the 16 tapes were used to assess interjudge reliability. Once the 14 compositional product tapes were coded, the researcher copied the first performance of the compositions on a single tape, following the subjects' code numbers ordered from 1 to 16. Because they were randomly chosen by the computer as samples to train the judges, the compositions of subjects with code numbers five and nine were missing from the final recording. In order to facilitate the identification of each subject's composition, the researcher said each subject's code number before the recording of each composition. Once the 14 compositions were recorded on a single tape, the researcher made three copies of this tape, so that each judge could receive one copy to work with. The tapes and the Judge's Scoring Sheets were distributed among the three judges.

The Training of the Judges

First session. As described earlier in this chapter, three independent judges were responsible for the analysis of the compositional products. The three independent

judges were already described in the section devoted to describe pretest procedures in this chapter. Prior to the study and to the pilot test, the researcher met individually with each of the judges in order to train them. During these sessions, the researcher explained the definitions of the concepts included in the evaluation forms, such as tonal and metric cohesiveness, use of rhythmic and melodic pattern, and gave examples of fluency and originality. The judges also received these definitions in writing. For the training, the researcher used the two tapes with subjects' compositional products that were randomly chosen from the pilot test sessions.

On the form that the judges received, they were asked to listen to the composition once before starting the scoring. After listening to the first composition, the judges and the researcher tried to come to consensus on what fragments could be considered units of analysis, and which of these constituted an example of each of the concepts to be measured. Once the judges and the researcher seemed to agree on each concept, the judges were asked to fill out a Judge's Scoring Form for each of the two tapes. The Judge's Scoring Form used in the pilot study is displayed in Appendix G. The researcher explained and gave in written form the description of the scoring forms to the judges. The judges were told that the first two scoring items were done in a Likert-scale format.

Tonal and metric cohesiveness were assessed using a 7-points Likert-scale, with 1 indicating a lack of tonal and/or metric cohesiveness, and 7 indicating very strong cohesiveness. The use of repeated and developed rhythmic and melodic patterns was determined by the number of times the judges documented that each type of pattern was used throughout the composition. They wrote a tally in the provided space for each

time they heard a repeated or developed melodic or rhythmic pattern. The level of fluency and originality was to be indicated by a 7-point Likert-scale, with 1 indicating the lack of each of these characteristics, and 7 indicating the highest level possible of each of these characteristics. After the training session, the judges were asked to assess the 14 remaining compositions individually. Judges were asked to listen to each tape once before scoring, and then they could listen to each as many times as they needed.

The three judges agreed that it was easier to assess tonal and metric cohesiveness after they had assessed the use of repeated and developed metric and melodic patterns. Judge A also suggested that the Scoring Form should include the sentence “Very strong tonal cohesiveness, no tonal cohesiveness, Very strong metric cohesiveness, no metric cohesiveness” above the appropriate numbers, so that it would be easier for them to keep in mind what number represented what concept. Judge C suggested that the distribution of the Likert-scale points should be equal for every concept in the scoring form. As a result of these comments, the Judges’ Scoring Form was modified for the main study. The section devoted to assess the use of repeated and developed metric and melodic patterns was placed at the top of the page. The phrases requested by Judge A were placed in the appropriate place for the section devoted to the assessment of metric and melodic cohesiveness, and the 7-points of the Likert-scales were distributed in the page in the same way for each of the concepts. Other than those changes, the judges reported that every concept was explained clearly to them. The final version of the Judge’s Scoring Form is displayed in Appendix D.

Once the judges completed the Judge’s Scoring Forms for the 14 pilot test compositions, the interjudge reliability was determined following the steps described

later in this chapter in the section devoted to the reliability of the measurement instrument. The overall interjudge reliability as well as the interjudge reliability for each of the subtest were low. It was clear that the judges were not in agreement on any of the concepts assessed in the instrument. They did not seem to agree even in what fragments of the pieces could be considered rhythmic and/or melodic patterns. It was decided that there was a need for a second training session.

Second session. For the second training session, the researcher met with the three judges at the same time. During this session, the researcher explained again how each concept was defined for the purpose of the study. The concepts that seemed to be causing the most confusion for the judges were developed rhythmic and melodic patterns. None of them could isolate rhythmic patterns from melodic ones. They agreed on the idea that, in order for them to identify a rhythmic pattern that was being repeated or developed independently for its melody, they would need a transcription of the composition in order to analyze it visually. They found it difficult to isolate rhythm from melody when analyzing aurally. In response to this request, all the pretest and posttest compositions of the study were transcribed by the researcher.

During the second training session, the judges agreed to consider a rhythmic pattern as developed when a melodic contour was reproduced with different rhythm but a similar note sequence. A melodic pattern would be considered as developed when a sequence of notes was slightly modified but was easily identifiable as a close reproduction of the original pattern. If a sequence of notes in a pattern was repeated with different rhythm, the judges agreed on considering it as a repeated melodic pattern, but as a developed rhythmic pattern.

In an attempt to unify the concept of melodic and rhythmic patterns, the judges received the transcription of the two examples used for their training. The transcription of the two compositions used for the training are displayed on figures 11 and 12. In these transcriptions, the researcher marked what she considered to be melodic and rhythmic patterns, and when she considered the patterns to be developed or repeated. After listening to the examples and analyzing the transcriptions, the judges and the researcher discussed their ideas until they reached a consensus. Then the judges were asked to assess the pilot test's compositions for the second time. The data obtained from this second effort were used to establish the interjudge reliability of the measurement instrument.

Reliability

In order to determine the level of interjudge reliability, the proportion of agreement between judges was determined by adding the total number of times the judges agreed on the overall score and for each of the eight subtests. Then, this number was divided by the maximum number of possible agreements or disagreements. Following Boyle and Radocy's (1987) suggestion, it was decided that, because of the subjective nature of concepts such as originality and fluency, the difficulty of identifying rhythmic and melodic patterns aurally, and the ample spectrum provided by a 7-point Likert-scale, scores ± 1 from each other would be considered agreements.

For the overall interjudge reliability, there were eight subtests, 14 subjects, and three judges, making a total of 336 possible agreements. There was a total of 265 agreements, which divided by 336 resulted in an interjudge reliability coefficient of 0.79.

For the interjudge reliability of the tonal and metric cohesiveness sections, there were 2 subtests, 14 subjects, and three judges, making a total of 84 possible agreements. There was a total of 71 agreements, which divided by 84 resulted in an interjudge reliability coefficient of 0.85. For the interjudge reliability of the use of rhythmic and melodic patterns, there were 4 subtests, 14 subjects, and three judges, making a total of 168 possible agreements. There was a total of 129 agreements, which divided by 168 resulted in an interjudge reliability coefficient of 0.77. For the interjudge reliability of the creativity section of the test, there were 2 subtests, 14 subjects, and three judges, making a total of 84 possible agreements. There was a total of 76 agreements, which divided by 84 resulted in an interjudge reliability coefficient of 0.90. The tables included in Appendix H show the coefficients of interjudge reliability obtained by each pair of judges, and for the three judges combined together, for the pilot tests, pretests and posttests data on each of the subtests related to Pattern Use, Cohesiveness, and Creativity.

CHAPTER IV

RESULTS

The results of the present study will be presented in different sections. First, the results obtained for the interjudge reliability test obtained for both the pretest and posttest will be presented. Then, the results obtained for the test of homogeneity of variance will be presented. Finally, the results obtained for each one of the eight subtests will be presented.

Interjudge Reliability

There were three independent judges in charge of assessing each subject's final creative problem solving product. The three female judges were general music educators with extended professional experience. Judge A had 30 years of experience teaching kindergarten to fifth grade, and held a National OAKE Endorsement as a Certified Kodály Teacher. Judge B had 16 years of experience teaching kindergarten to fifth grade students in public elementary schools, had a Certification in Orff Schulwerk, was teaching music education courses at a major university in north Texas, and was working toward her Ph.D. in Music Education. Judge C had 5 years of experience teaching kindergarten to fifth grade, and was in the process of finishing her National OAKE Endorsement as a Certified Kodály Teacher as well as her master's degree in music education.

In addition to the exploration of the interjudge reliability with the data obtained during the pilot test, interjudge reliability was explored for both the pretests and the posttests in the main study. There were eight subtests for the compositional products. The first four subtests, related to pattern use, were the measures obtained for the use of

repeated melodic pattern, developed melodic pattern, repeated rhythmic pattern, and developed rhythmic pattern. The next two subtests, related to cohesiveness, were the measures obtained for tonal and metric cohesiveness. The last two subtests, related to creativity, were the measures obtained for fluency and originality. For the pilot test, the pretests, and the posttests, the proportion of agreement between judges was determined by adding the total number of times the judges agreed overall, the total number of times the judges agreed for pattern use, cohesiveness, and creativity, and the total number of times the judges agreed on each one of the eight subtests. Then, each number was divided by the maximum number of possible agreements for each subtest or group of subtests. Because of the subjective nature of concepts such as originality and fluency, the difficulty of identifying rhythmic and melodic pattern, and the ample spectrum provided by a 7-point Likert scale, scores ± 1 from each other were considered agreements (Boyle & Radocy, 1987).

The overall interjudge reliability was 0.79 for the pilot test, 0.6 for the pretests, and 0.59 for the posttests. Interjudge reliability for pattern use was 0.77 for the pilot test, 0.66 for the pretests, and 0.68 for the posttests. Interjudge reliability for cohesiveness was 0.85 for the pilot test, 0.57 for the pretests, and 0.51 for the posttests. Finally, interjudge reliability for creativity was 0.9 for the pilot test, 0.5 for the pretests, and 0.5 for the posttests. Table 1 shows the overall levels of interjudge reliability found during the pilot test, the pretests, and the posttests, when the three judges were compared.

Table 1

Overall Interjudge Reliability

	Judges A + B + C		
	Pilot Test	Pretest	Posttest
Overall	0.79	0.6	0.59
Pattern Use	0.77	0.66	0.68
Cohesiveness	0.85	0.57	0.51
Creativity	0.9	0.5	0.5

The interjudge reliability obtained for each subtest, and for each possible combination of the judges on the pilot test, the pretests, and the posttests was also explored. The tables included in Appendix H show the coefficients of interjudge reliability obtained by each pair of judges, and for the three judges combined together, for the pilot tests, pretests and posttests data on each of the subtests related to Pattern Use, Cohesiveness, and Creativity.

Homogeneity of Variance

Given the fact that the homogeneity of variance of the covariate is one of the assumptions that must be met before a MANCOVA analysis could take place, this was the first statistical analysis for the present study. In the current study, the covariate was the pretest scores. The homogeneity of variance of the two treatment groups on the three groups of subtests was compared through a Box test. This statistical procedure tests the null hypothesis that the observed covariate matrices of the dependent variables are equal across groups. For the subtests grouped under the concept of pattern use, the Box test reported a significant difference between the two treatment groups ($F = 3.39$), which meant that these subtests might not be appropriate as covariates. Therefore, these variables were explored with one-way ANOVAs. The results of these tests will be reported later in this chapter.

For the subtests grouped under the concept of cohesiveness, the Box test reported that there was not a significant difference between the groups at a .059 level ($F = 2.48$). Because this level was found too close to the established α level of .05, it was decided that these two subtests would also be explored through one-way ANOVAs. The results of these tests will be reported later in this chapter.

Finally, for the subtests grouped under the concept of creativity, the Box tests reported that groups did not differ at a 0.43 level ($F = .91$). However, to maintain consistency with the statistical analyses of the study, and because of the possibility of error due to small sample size, it was decided that these two subtests would also be explored with two one-way ANOVAs. The results of these tests will be reported in the following section.

Subtests

In this study, the dependent variables were the measures obtained from the compositional product posttests. One subject did not attend school the day the posttests were scheduled to take place. Therefore, these scores were missing from the final data. For the purposes of this study, the compositional products were evaluated in terms of their pattern use with four subtests, cohesiveness with two subtests, and creativity with two subtests. The statistical analyses of the subtests are presented in the following section.

Pattern Use

There were four subtests grouped under the concept of pattern use: repeated melodic pattern, developed melodic pattern, repeated rhythmic pattern, and developed rhythmic pattern. The judges indicated the number of times the subjects made use of each type of pattern by writing on the Judge's Scoring Form a tally mark for each time they heard a recognizable melodic or rhythmic pattern being repeated or developed. The Judge's Scoring Form used in the main study is displayed in Appendix D. The total number of tallies for each type of pattern was summed in order to obtain the total number of occurrences of each type of pattern during the composition.

Prior to the hypothesis testing, and in order to check if there were initial differences between the two conditions, the data obtained from the pretests were explored with a one-way ANOVA for each of the four subtests grouped under the concept of pattern use. First, the assumptions of normality and homogeneity of variance were examined for each subtest. As was the case for the other subtests of this study, the assumption of normality was checked through the exploration of skewness and kurtosis. The values obtained for the standard error for skewness and kurtosis were each multiplied by two standard deviations on each side of the mean (± 1.96). If the subtests' obtained statistics for skewness and kurtosis were equal to or below these confidence intervals, it was considered that the assumption of normality was met. Conversely, if the subtests' obtained statistics for skewness and kurtosis were higher than the confidence intervals, it was considered that the assumption of normality was not met. Homogeneity of variance was checked with a Levene test.

Use of Repeated Melodic Pattern. For the use of repeated melodic pattern subtests, the judges indicated the number of times the subjects repeated melodic patterns throughout the composition. In order to determine if the groups differed before the treatment sessions, the data obtained from the use of repeated melodic pattern pretest scores were explored with an ANOVA. Prior to this analysis, the assumptions of homogeneity of variance and normality were checked. The assumption of normality was checked through the exploration of skewness and kurtosis. In the case of the use of repeated melodic pattern pretest scores for the individual condition, the mean was 2.44 ($SD = 4.05$), and the distribution did not meet the assumption of normality; the skewness level of 2.93 did not fall under the established confidence interval (-.81 to

+ .81). The obtained kurtosis level of 10.38 did not fall under the confidence interval (-1.58 to +1.58). In the case of the use of repeated melodic pattern pretests for the collective condition, the mean was 1.18 ($SD = 2.92$), and the distribution did not meet the assumption of normality, with a skewness of 4.38 ($CI = -.81$ to $+.81$), and a kurtosis of 21.83 ($CI = -1.58$ to $+1.58$). Therefore, and as suggested by Huck (2000), the sample data were transformed. The log transformation, the square root transformation, and the inverse transformation were explored. In order to avoid scores of zero, a constant of 1 was added to the log and the inverse transformation. Results indicated that the log transformation values more closely approximated normality. Accordingly, all subsequent statistical tests used log transformed values for the use of repeated melodic pattern.

In the case of the transformed data for the use of repeated melodic pattern pretest scores for the individual condition, the distribution was considered to approach normality. The skewness level of .80 fell under the established confidence interval ($-.81$ to $+.81$). The obtained kurtosis level of -.20 also fell under the confidence interval (-1.58 to +1.58). In the case of the transformed data for the use of repeated melodic pattern pretest scores for the collective condition, the distribution was considered to approach normality closer than the original data did, with a skewness of 1.70 ($CI = -.81$ to $+.81$), and a kurtosis of 3.02 ($CI = -1.58$ to $+1.58$). Despite the fact that these values did not meet the assumption of normality, the obtained values best approached normality of the three transformed data options. Therefore, it was decided to proceed with the exploration of the pretest scores. The Levene test reported that the pretest log transformed values for the use of repeated melodic pattern scores met the assumption

of homogeneity of variance at the .11 level. The log transformed values for the use of repeated melodic pattern pretest scores were then explored with a one-way ANOVA.

Given the fact that there were eight subtests being explored with one-way ANOVAs, the α level was adjusted with the Bonferroni technique ($.05 / 8 = .006$). The new α level was set at the .006 level. The results showed that the groups did not differ significantly on the pretest scores ($F = 3.45, p = .06$). Therefore, it was possible to test the null hypothesis of no difference between the groups' posttest means for the use of repeated melodic pattern scores with a one-way ANOVA.

The range of obtained scores for the individual condition was .00 to 6.33 ($M = .97; SD = 1.46$), and .00 to 17 for the collective condition ($M = 1.75; SD = 3.39$). Prior to running the ANOVA for the use of repeated melodic pattern posttest scores, the assumptions of homogeneity of variance and normality were checked. The distribution of the use of repeated melodic pattern posttest scores for the individual condition ($M = .97; SD = 1.46$) did not meet the assumption of normality, with a skewness of 2.12 ($CI = -.81$ to $+.81$), and a kurtosis of 5.16 ($CI = -1.58$ to $+1.58$). In the case of repeated melodic pattern posttest scores for the collective condition ($M = 1.75; SD = 3.39$), the distribution did not meet the assumption of normality, with a skewness of 3.40 ($CI = -.82$ to $+.82$), and a kurtosis of 13.73 ($CI = -1.60$ to $+1.60$). Therefore, and as suggested by Huck (2000), the sample data were transformed. The log transformation, the square root transformation, and the inverse transformation were explored. In order to avoid scores of zero, a constant of 1 was added to the log and the inverse transformation. Results indicated that the log transformation values more closely approximated

normality. Accordingly, all subsequent statistical tests used log transformed values for the use of repeated melodic pattern.

The range of obtained transformed scores for the individual condition was .00 to .87 ($M = .21$; $SD = .25$), and .00 to .1.26 for the collective condition ($M = .27$; $SD = .34$). Prior to running the ANOVA for the use of repeated melodic pattern posttest scores, the assumptions of homogeneity of variance and normality were checked. In the case of the transformed data for the use of repeated melodic pattern posttest scores for the individual condition, the distribution was considered to approach normality. The skewness level of .94 fell closer to the established confidence interval (-.81 to +.81). The obtained kurtosis level of -.08 fell under the confidence interval (-1.58 to +1.58). In the case of the transformed data for the use of repeated melodic pattern posttest scores for the collective condition, the distribution was considered to approach normality closer than the original data did, with a skewness of 1.26 ($CI = -.81$ to $+.81$), and a kurtosis of .90 ($CI = -1.58$ to $+1.58$). The Levene test reported that the posttest scores for the log transformed values for the use of repeated melodic pattern met the assumption of homogeneity of variance at the .13 level. The log transformed values for the use of repeated melodic pattern posttest scores were then explored with a one-way ANOVA.

Given the fact that there were eight subtests being explored with one-way ANOVAs, the α level was adjusted with the Bonferroni technique ($.05 / 8 = .006$). The new α level was set at the .006 level. According to the statistical analysis, the mean for the subjects in the individual condition ($M = .21$; $SD = .25$), was slightly, although not significantly, lower than the mean for the subjects in the collective condition ($M = .27$; $SD = .34$) (see Table 2 for descriptive statistics).

Table 3 presents the selected output from the SPSS® one-way ANOVA test for the use of repeated melodic pattern log transformed data. The results showed that the groups did not differ significantly ($F = .68$, $p = .41$). The effect size was .01 (SPSS partial η^2), which can be considered a small effect (Huck, 2000). This could mean that, even if the sample size had been larger, the possibility of failing to reject the null hypothesis due to lack of differences between the two conditions, would have been high.

Table 2

*Descriptive Statistics for the Use of Repeated Melodic Pattern Log Transformed
Posttest Data*

	Descriptive Statistics	
	Mean	SD
Individual condition	.21	.25
Collective condition	.27	.34
Total	.24	.30

Table 3

Analysis of Variance for the Use of Repeated Melodic Pattern Log Transformed Posttest

Data

Source	SS	df	MS	F	P
Treatment	.06	1	.06	.68	.41
Error	5.41	61	.09		
Total	9.19	63			

Use of Developed Melodic Patterns. For the use of developed melodic pattern subtests, the judges indicated the number of times the subjects developed melodic patterns throughout the composition. In order to determine if the groups differed before the treatment sessions, the data obtained from the use of developed melodic pattern pretest scores were explored with an ANOVA. Prior to this analysis, the assumptions of homogeneity of variance and normality were checked. The assumption of normality was checked through the exploration of skewness and kurtosis. In the case of the use of developed melodic pattern pretest scores for the individual condition ($M = 2.52$; $SD = 3.23$), the distribution did not meet the assumption of normality; the skewness level of 2.45 did not fall under the established confidence interval (-.81 to +.81). The obtained kurtosis level of 7.00 did not fall under the confidence interval (-1.58 to +1.58). In the case of the use of developed melodic pattern pretest scores for the collective condition ($M = 1.88$; $SD = 1.62$), the distribution did not meet the assumption of normality, with a skewness of 1.92 ($CI = -.81$ to $+.81$), and a kurtosis of 4.44 ($CI = -1.58$ to $+1.58$). Therefore, and as suggested by Huck (2000), the sample data were transformed. The log transformation, the square root transformation, and the inverse transformation were explored. In order to avoid scores of zero, a constant of 1 was added to the log and the inverse transformation. Results indicated that the log transformation values more closely approximated normality. Accordingly, all subsequent statistical tests used log transformed values for the use of developed melodic pattern.

In the case of the transformed data for the use of developed melodic pattern pretest scores for the individual condition, the distribution was considered to approach normality. The skewness level of .57 fell under the established confidence interval (-.81

to +.81). The obtained kurtosis level of -.196 fell under the confidence interval (-1.58 to +1.58). In the case of the transformed data for the use of developed melodic pattern pretest scores for the collective condition, the distribution was considered to approach normality, with a skewness of .56 (CI = -.81 to +.81), and a kurtosis of .28 (CI = -1.58 to +1.58). The Levene test reported that the pretest scores for the log transformed values for the use of developed melodic patterns met the assumption of homogeneity of variance at the .02 level. Therefore, the log transformed values for the use of developed melodic pattern pretest scores were explored with a one-way ANOVA.

Given the fact that there were eight subtests being explored with one-way ANOVAs, the α level was adjusted with the Bonferroni technique ($.05 / 8 = .006$). The new α level was set at the .006 level. The results showed that the groups did not differ significantly on the pretest scores ($F = .06, p = .81$). Therefore, the null hypothesis of no difference between the groups' posttest means for the use of developed melodic pattern scores was tested with a one-way ANOVA.

The range of obtained scores for the individual condition was .33 to 7.33 ($M = 2.03$; $SD = 1.48$), and .00 to 4.67 for the collective condition ($M = 1.85$; $SD = 1.30$). Prior to running the ANOVA for the use of developed melodic pattern posttest scores, the assumptions of homogeneity of variance and normality were checked. The distribution of the use of developed melodic pattern posttest scores for the individual condition ($M = 2.03$; $SD = 1.48$) did not meet the assumption of normality, with a skewness of 1.90 (CI = -.81 to +.81), and a kurtosis of 4.94 (CI = -1.58 to +1.58). In the case of developed melodic pattern use posttest scores for the collective condition ($M = 1.85$; $SD = 1.30$), the distribution did meet the assumption of normality, with a skewness of .54 (CI = -.82

to +.82), and a kurtosis of -.90 (CI = -1.60 to + 1.60). Because the data for the individual condition did not meet the assumption of normality, and as suggested by Huck (2000), the sample data were transformed. The log transformation, the square root transformation, and the inverse transformation were explored. In order to avoid scores of zero, a constant of 1 was added to the log and the inverse transformation. Results indicated that the log transformation values more closely approximated normality. Accordingly, all subsequent statistical tests used log transformed values for the use of developed melodic pattern.

The range of obtained transformed scores for the individual condition was .12 to .92 ($M = .44$; $SD = .19$), and .00 to .75 for the collective condition ($M = .41$; $SD = .20$). Prior to running the ANOVA for the use of developed melodic pattern posttest scores, the assumptions of homogeneity of variance and normality were checked. In the case of the transformed data for the use of developed melodic pattern posttest scores for the individual condition, the distribution was considered to approach normality. The skewness level of .34 fell under the established confidence interval (-.81 to +.81). The obtained kurtosis level of .42 fell under the confidence interval (-1.58 to +1.58). In the case of the log transformed data for the use of developed melodic pattern posttest scores for the collective condition, the distribution was considered to approach normality with a skewness of -.08 (CI = -.81 to +.81), and a kurtosis of -.98 (CI = -1.58 to +1.58). The Levene test reported that the posttest scores for the log transformed values for the use of developed melodic pattern met the assumption of homogeneity of variance at the .32 level. The log transformed values for the use of developed melodic pattern posttest scores were then explored with a one-way ANOVA.

Given the fact that there were eight subtests being explored with one-way ANOVAs, the α level was adjusted with the Bonferroni technique ($.05 / 8$). After the Bonferroni adjustment, the α level was set at the .006 level. According to the statistical analysis, the transformed log mean for the subjects in the individual condition ($M = .44$; $SD = .19$) was slightly, although not significantly, higher than the mean for the subjects in the collective condition ($M = .41$; $SD = .20$) (see Table 4 for descriptive statistics).

Table 5 presents the selected output from the SPSS one-way ANOVA test for the use of developed melodic pattern log transformed data. The results showed that the groups did not differ significantly ($F = .38$, $p = .54$). The effect size was .006 (SPSS partial η^2), which can be considered a small effect (Huck, 2000). This could mean that, even if the sample size had been larger, the possibility of failing to reject the null hypothesis due to lack of differences between the two conditions, would have been high.

Table 4

*Descriptive Statistics for the Use of Developed Melodic Pattern Log Transformed
Posttest Data*

	Descriptive Statistics	
	Mean	SD
Individual condition	.44	.19
Collective condition	.41	.20
Total	.43	.20

Table 5

*Analysis of Variance for the Use of Developed Melodic Pattern Log Transformed**Posttest Data*

Source	SS	df	MS	F	p
Treatment	.02	1	.02	.38	.54
Error	2.36	61	.04		
Total	13.75	63			

Use of Repeated Rhythmic Pattern. For the use of repeated rhythmic pattern subtests, the judges indicated the number of times the subjects repeated rhythmic patterns throughout the composition. In order to determine if the groups differed before the treatment sessions, the data obtained from the use of repeated rhythmic pattern pretest scores were explored with an ANOVA. Prior to this analysis, the assumptions of homogeneity of variance and normality were checked. The assumption of normality was checked through the exploration of skewness and kurtosis. In the case of the use of repeated rhythmic pattern pretest scores for the individual condition ($M = 4.80$; $SD = 7.32$), the distribution did not meet the assumption of normality. The skewness level of 3.00 did not fall under the established confidence interval (-.81 to +.81). The obtained kurtosis level of 10.20 did not fall under the confidence interval (-1.58 to +1.58). In the case of the use of repeated rhythmic pattern pretest scores for the collective condition ($M = 3.94$; $SD = 6.77$), the distribution did not meet the assumption of normality, with a skewness of 4.27 ($CI = -.81$ to $+.81$), and a kurtosis of 20.20 ($CI = -1.58$ to $+1.58$). Therefore, and as suggested by Huck (2000), the sample data were transformed. The log transformation, the square root transformation, and the inverse transformation were explored. In order to avoid scores of zero, a constant of 1 was added to the log and the inverse transformation. Results indicated that the inverse transformation values more closely approximated normality. Accordingly, all subsequent statistical tests used inverse transformed values for the use of repeated rhythmic pattern.

In the case of the inverse transformed data for the use of repeated rhythmic pattern pretest scores for the individual condition, the distribution was considered to approach normality closer than the original data did. The skewness level of .92 fell

closer to the established confidence interval (-.81 to +.81). The obtained kurtosis level of .28 fell under the confidence interval (-1.58 to +1.58). In the case of the transformed data for the use of repeated rhythmic pattern pretest scores for the collective condition, the distribution was considered to approach normality closer than the original data did, with a skewness of .70 (CI = -.81 to +.81), and a kurtosis of -1.69 (CI = -1.58 to +1.58). The Levene test reported that the pretest scores for the inverse transformed values for the use of repeated rhythmic pattern met the assumption of homogeneity of variance at the .07 level. The inverse transformed values for the use of repeated rhythmic pattern pretest scores were then explored with a one-way ANOVA.

Given the fact that there were eight subtests being explored with one-way ANOVAs, the α level was adjusted with the Bonferroni technique ($.05 / 8 = .006$). The new α level was set at the .006 level. The results showed that the groups did not differ significantly on the pretest scores ($F = .15, p = .70$). Therefore, the null hypothesis of no difference between the groups' posttest means for the use of repeated rhythmic pattern scores was tested with a one-way ANOVA.

The range of obtained scores for the individual condition was .00 to 39.33 ($M = 3.88$; $SD = 6.88$), and .00 to 11.33 for the collective condition ($M = 2.63$; $SD = 2.43$). Prior to running the ANOVA for the use of repeated rhythmic pattern posttest scores, the assumptions of homogeneity of variance and normality were checked. The distribution of the use of repeated rhythmic pattern posttest scores for the individual condition ($M = 3.88$; $SD = 6.88$) did not meet the assumption of normality, with a skewness of 4.70 (CI = -.81 to +.81), and a kurtosis of 24.33 (CI = -1.58 to +1.58). In the case of the repeated rhythmic pattern use posttest scores for the collective condition (M

= 2.63; $SD = 2.43$), the distribution did not meet the assumption of normality, with a skewness of 3.40 ($CI = -.82$ to $+.82$), and a kurtosis of 13.73 ($CI = -1.60$ to $+ 1.60$). Because the data did not meet the assumption of normality, and as suggested by Huck (2000), the sample data were transformed. The log transformation, the square root transformation, and the inverse transformation were explored. In order to avoid scores of zero, a constant of 1 was added to the log and the inverse transformation. Results indicated that the inverse transformation values more closely approximated normality. Accordingly, all subsequent statistical tests used inverse transformed values for the use of repeated rhythmic pattern.

The range of obtained transformed scores for the individual condition was .02 to 1.00 ($M = .36$; $SD = .21$), and .08 to 1.00 for the collective condition ($M = .40$; $SD = .24$). Prior to running the ANOVA for the use of repeated rhythmic pattern posttest scores, the assumptions of homogeneity of variance and normality were checked. In the case of the transformed data for the use of repeated rhythmic pattern posttest scores for the individual condition, the distribution was considered to approach normality. The skewness level of 1.14 fell closer to the established confidence interval ($-.81$ to $+.81$). The obtained kurtosis level of 1.90 fell closer to the confidence interval (-1.58 to $+1.58$). In the case of the inverse transformed data for the use of repeated rhythmic pattern posttest scores for the collective condition, the distribution was considered to approach normality with a skewness of 1.11 ($CI = -.81$ to $+.81$), and a kurtosis of .63 ($CI = -1.58$ to $+1.58$). The Levene test reported that the posttest scores for the inverse transformed values for the use of repeated rhythmic pattern met the assumption of homogeneity of

variance at the .27 level. The inverse transformed values for the use of repeated rhythmic pattern posttest scores were then explored with a one-way ANOVA.

Given the fact that there were eight subtests being explored with one-way ANOVAs, the α level was adjusted with the Bonferroni technique ($.05 / 8$). After the Bonferroni adjustment, the α level was set at the .006 level. According to the statistical analysis, the inverse transformed mean for the subjects in the individual condition ($M = .36$; $SD = .21$) was slightly, although not significantly, lower than the mean for the subjects in the collective condition ($M = .40$; $SD = .24$) (see Table 6 for descriptive statistics).

Table 7 presents the selected output from the SPSS one-way ANOVA test for the use of repeated rhythmic pattern inverse transformed data. The results showed that the groups did not differ significantly ($F = .51$, $p = .48$). The effect size was .008 (SPSS partial η^2), which can be considered a small effect (Huck, 2000). This could mean that, even if the sample size had been larger, the possibility of failing to reject the null hypothesis due to lack of differences between the two conditions, would have been high.

Table 6

*Descriptive Statistics for the Use of Repeated Rhythmic Pattern Inverse Transformed
Posttest Data*

	Descriptive Statistics	
	Mean	SD
Individual condition	.36	.21
Collective condition	.40	.24
Total	.38	.23

Table 7

Analysis of Variance for the Use of Repeated Rhythmic Pattern Inverse Transformed

Posttest Data

Source	SS	df	MS	F	p
Treatment	.03	1	.03	.51	.48
Error	3.16	61	.05		
Total	12.10	63			

Use of Developed Rhythmic Pattern. For the use of developed rhythmic pattern subtests, the judges indicated the number of times the subjects developed rhythmic patterns throughout the composition. In order to determine if the groups differed before the treatment sessions, the data obtained from the use of developed rhythmic pattern pretest scores were explored with an ANOVA. Prior to this analysis, the assumptions of homogeneity of variance and normality were checked. The assumption of normality was checked through the exploration of skewness and kurtosis. In the case of the use of developed rhythmic pattern pretest scores for the individual condition ($M = .70$; $SD = .70$), the distribution met the assumption of normality. The skewness level of .69 fell under the established confidence interval (-.81 to +.81). The obtained kurtosis level of -.61 fell under the confidence interval (-1.58 to +1.58). However, in the case of the use of developed rhythmic pattern pretest scores for the collective condition ($M = .76$; $SD = 1.03$), the distribution did not meet the assumption of normality, with a skewness of 2.53 ($CI = -.81$ to $+.81$), and a kurtosis of 8.50 ($CI = -1.58$ to $+1.58$). Therefore, and as suggested by Huck (2000), the sample data were transformed. The log transformation, the square root transformation, and the inverse transformation were explored. In order to avoid scores of zero, a constant of 1 was added to the log and the inverse transformation. Results indicated that the log transformation values more closely approximated normality. Accordingly, all subsequent statistical tests used log transformed values for the use of developed rhythmic pattern.

In the case of the transformed data for the use of developed rhythmic pattern pretest scores for the individual condition, the distribution was considered to approach normality. The skewness level of .25 fell under the established confidence interval (-.81

to +.81). The obtained kurtosis level of -1.32 fell under the confidence interval (-1.58 to +1.58). In the case of the log transformed data for the use of developed rhythmic pattern pretest scores for the collective condition, the distribution was considered to approach normality closer than the original data did, with a skewness of .98 (CI = -.81 to +.81), and a kurtosis of .81 (CI = -1.58 to +1.58). The Levene test reported that the pretest scores for the log transformed values for the use of developed rhythmic pattern met the assumption of homogeneity of variance at the .72 level. The log transformed values for the use of developed rhythmic pattern pretest scores were then explored with a one-way ANOVA.

Given the fact that there were eight subtests being explored with one-way ANOVAs, the α level was adjusted with the Bonferroni technique ($.05 / 8 = .006$). The new α level was set at the .006 level. The results showed that the groups did not differ significantly on the pretest scores ($F = .00, p = 1.00$). Therefore, the null hypothesis of no difference between the groups' posttest means for the use of developed rhythmic pattern was tested with a one-way ANOVA.

The range of obtained scores for the individual condition was .00 to 5 ($M = .90$; $SD = .96$), and .00 to 4 for the collective condition ($M = .92$; $SD = .95$). Prior to running the ANOVA for the use of developed rhythmic pattern posttest scores, the assumptions of homogeneity of variance and normality were checked. The distribution of the use of developed rhythmic pattern posttest scores for the individual condition ($M = .90$; $SD = .96$) did not meet the assumption of normality, with a skewness of 2.92 (-.81 to +.81), and a kurtosis of 10.70 (-1.58 to +1.58). In the case of developed rhythmic pattern use posttest scores for the collective condition ($M = .92$; $SD = .95$), the distribution did not

meet the assumption of normality, with a skewness of 1.39 (-.82 to +.82), and a kurtosis of 2.40 (-1.60 to + 1.60). Because the data did not meet the assumption of normality, and as suggested by Huck (2000), the sample data were transformed. The log transformation, the square root transformation, and the inverse transformation were explored. In order to avoid scores of zero, a constant of 1 was added to the log and the inverse transformation. Results indicated that the square root transformation values more closely approximated normality. Accordingly, all subsequent statistical tests used square root transformed values for the use of developed rhythmic pattern.

The range of obtained transformed scores for the individual condition was .00 to 2.24 ($M = .83$; $SD = .47$), and .00 to 2.00 for the collective condition ($M = .77$; $SD = .58$). Prior to running the ANOVA for the use of developed rhythmic pattern posttest scores, the assumptions of homogeneity of variance and normality were checked. In the case of the square root transformed data for the use of developed rhythmic pattern posttest scores for the individual condition, the distribution was considered to approach normality. The skewness level of .50 fell under the established confidence interval (-.81 to +.81). The obtained kurtosis level of 2.10 fell closer to the confidence interval (-1.58 to +1.58) than the original data did. In the case of the square root transformed data for the use of developed rhythmic pattern posttest scores for the collective condition, the distribution was considered to approach normality with a skewness of -.04 ($CI = -.82$ to +.82), and a kurtosis of -.91 ($CI = -1.61$ to +1.61). The Levene test reported that the posttest scores for the square root transformed values for the use of developed rhythmic pattern met the assumption of homogeneity of variance at the .05 level. The

square root transformed values for the use of developed rhythmic pattern posttest scores were then explored with a one-way ANOVA.

Given the fact that there were eight subtests being explored with one-way ANOVAs, the α level was adjusted with the Bonferroni technique ($.05 / 8$). After the Bonferroni adjustment, the α level was set at the .006 level. According to the statistical analysis, the square root transformed mean for the subjects in the individual condition ($M = .83$; $SD = .47$) was slightly, although not significantly, higher than the mean for the subjects in the collective condition ($M = .77$; $SD = .58$) (see Table 8 for descriptive statistics).

Table 9 presents the selected output from the SPSS one-way ANOVA test for the square root transformed use of developed rhythmic pattern data. The results showed that the groups did not differ significantly ($F = .18$, $p = .68$). The effect size was .003 (SPSS partial η^2), which can be considered a small effect (Huck, 2000). This could mean that, even if the sample size had been larger, the possibility of failing to reject the null hypothesis due to lack of differences between the two conditions, would have been high.

Table 8

Descriptive Statistics for the Use of Developed Rhythmic Pattern Square Root

Transformed Posttest Data

	Descriptive Statistics	
	Mean	SD
Individual condition	.83	.47
Collective condition	.77	.58
Total	.80	.53

Table 9

*Analysis of Variance for the Use of Developed Rhythmic Pattern Square Root**Transformed Posttest Data*

Source	SS	df	MS	F	p
Treatment	.05	1	.05	.18	.68
Error	17.06	61	.28		
Total	57.33	63			

Cohesiveness

There were two subtests grouped under the concept of cohesiveness: tonal cohesiveness and metric cohesiveness. Both subtests were assessed with a 7-point Likert scale. The results obtained on these two subtests were each explored with a one-way ANOVA.

Tonal Cohesiveness. For tonal cohesiveness, the judges indicated their assessment on a 7-point Likert scale, with a 7 representing very strong tonal cohesiveness, and a 1 indicating no cohesiveness. In order to determine if the groups differed before the treatment sessions, the data obtained from the tonal cohesiveness pretest scores were explored with an ANOVA. Prior to this analysis, the assumptions of homogeneity of variance and normality were checked. In the case of the tonal cohesiveness pretest scores for the individual condition, the mean was 3.49 ($SD = 1.85$), and the distribution was considered to approach normality. The skewness level of .233 fell under the established confidence interval (-.81 to +.81). The obtained kurtosis level also fell under the confidence interval of -1.29 ($CI = -1.58$ to $+1.58$). In the case of tonal cohesiveness pretest scores for the collective condition, the mean was 3.84 ($SD = 1.81$), and the distribution was considered to approach normality, with a skewness of .026 ($CI = -.81$ to $+.81$), and a kurtosis of -1.26 ($CI = -1.58$ to $+1.58$). The Levene test reported that the pretest scores for tonal cohesiveness met the assumption of homogeneity of variance at the .78 level. Because the assumptions of normality and homogeneity of variance were met, the tonal cohesiveness pretest scores were explored with a one-way ANOVA.

Given the fact that there were eight subtests being explored with one-way ANOVAs, the α level was adjusted with the Bonferroni technique ($.05 / 8 = .006$). The new α level was set at the .006 level. The results showed that the groups did not differ significantly on the pretest scores ($F = .60, p = .44$). Therefore, the null hypothesis of no difference between the groups' posttest means for tonal cohesiveness was tested with a one-way ANOVA.

The original range of possible scores was 1.00 to 7.00. Once the scores assigned by the three judges were averaged, the range of obtained scores was from 1.33 to 7.00 for the individual condition ($M = 4.69; SD = 1.82$), and 1.33 to 6.67 for the collective condition ($M = 4.30; SD = 1.50$). Prior to running the ANOVA for the tonal cohesiveness posttest scores, the assumptions of homogeneity of variance and normality were checked. The distribution of the tonal cohesiveness posttest scores for the individual condition was considered to approach normality, with a skewness of .41 ($CI = -.81$ to $+.81$), and a kurtosis of -1.13 ($CI = -1.58$ to $+1.58$). In the case of tonal cohesiveness posttest scores for the collective condition, the distribution was considered to approach normality, with a skewness of .107 ($CI = -.82$ to $+.82$), and a kurtosis of -1.004 ($CI = -1.60$ to $+1.60$). The Levene test reported that the posttest scores for tonal cohesiveness met the assumption of homogeneity of variance at the .14 level. Because the assumptions of normality and homogeneity of variance were met, tonal cohesiveness was explored with a one-way ANOVA.

Given the fact that there were eight subtests being explored with one-way ANOVAs, the α level was adjusted with the Bonferroni technique ($.05 / 8$). After the Bonferroni adjustment, the α level was set at the .006 level. According to the statistical

analysis, the mean for the subjects in the individual condition ($M = 4.69$; $SD = 1.82$) was slightly, although not significantly, higher than the mean for the subjects in the collective condition ($M = 4.30$; $SD = 1.50$) (see Table 10 for descriptive statistics).

Table 11 presents the selected output from the SPSS one-way ANOVA test for tonal cohesiveness. The results showed that the groups did not differ significantly ($F = .84$, $p = .36$). The effect size was .01 (SPSS partial η^2), which can be considered a small effect (Huck, 2000). This could mean that, even if the sample size had been larger, the possibility of failing to reject the null hypothesis due to lack of differences between the two conditions, would have been high.

Table 10

Descriptive Statistics for the Tonal Cohesiveness Posttest Data

	Descriptive Statistics	
	Mean	<i>SD</i>
Individual condition	4.68	1.82
Collective condition	4.30	1.50
Total	4.49	1.67

Table 11

Analysis of Variance for the Tonal Cohesiveness Posttest Data

Source	SS	df	MS	F	p
Treatment	2.35	1	2.35	.84	.36
Error	170.73	61	2.79		
Total	173.08	62			

Metric cohesiveness. Metric cohesiveness was assessed with a 7-point Likert scale, with a 7 representing very strong metric cohesiveness, and a 1 indicating no metric cohesiveness. In order to determine if the groups' metric cohesiveness differed before the treatment sessions, the data obtained from the pretest scores were explored with an ANOVA. Prior to this analysis, the assumptions of homogeneity of variance and normality for the pretest scores were checked. In the case of the metric cohesiveness pretest scores for the individual condition, the mean was 3.57 ($SD = 1.66$), and the distribution was considered to approach normality. The skewness level of $-.16$ fell within the established confidence interval ($-.81$ to $+.81$), as well as the kurtosis level of -1.14 (-1.58 to $+1.58$). In the case of the tonal cohesiveness pretest scores for the collective condition, the mean was 3.75 ($SD = 1.77$), and the distribution was considered to approach normality, with a skewness of $.08$ ($CI = -.81$ to $+.81$), and a kurtosis of -1.21 ($CI = -1.58$ to $+1.58$). The Levene test reported that the pretest scores for metric cohesiveness met the assumption of homogeneity of variance at the $.77$ level. Because the assumptions of normality and homogeneity of variance were met, the tonal cohesiveness pretest scores were explored with a one-way ANOVA.

Given the fact that there were eight subtests being explored with one-way ANOVAs, the α level was adjusted with the Bonferroni technique ($.05 / 8 = .006$). The new α level was set at the $.006$ level. The results showed that the groups did not differ significantly on the pretest scores ($F = .19, p = .66$). Therefore, the null hypothesis of no difference between the groups' posttest means for metric cohesiveness was tested with a one-way ANOVA.

The original range of possible scores was 1.00 to 7.00. Once the scores assigned by the three judges were averaged, the range of obtained scores was from 1.67 to 7.00 for the individual condition ($M = 4.76$; $SD = 1.30$), and 1.33 to 6.67 for the collective condition ($M = 4.44$; $SD = 1.36$). Prior to running the ANOVA for the metric cohesiveness posttest scores, the assumptions of homogeneity of variance and normality were checked. In the case of the metric cohesiveness posttest scores for the individual condition, the distribution was considered to approach normality, with a skewness of $-.66$ ($CI = -.81$ to $+.81$), and a kurtosis of $-.09$ ($CI = -1.58$ to $+1.58$). In the case of the metric cohesiveness posttest scores for the collective condition, the distribution was considered to approach normality, with a skewness of $-.14$ ($CI = -.82$ to $+.82$), and a kurtosis of $-.34$ ($CI = -1.60$ to $+1.60$). The Levene test reported that the posttest scores for metric cohesiveness met the assumption of homogeneity of variance at the .80 level. Because the assumptions of normality and homogeneity of variance were met, metric cohesiveness posttest scores were explored with a one-way ANOVA.

Given the fact that there were eight subtests being explored with one-way ANOVAs, the α level was adjusted with the Bonferroni technique ($.05 / 8$). After the Bonferroni adjustment, the α level was set at the .006 level. According to the statistical analysis, the mean for the subjects in the individual condition ($M = 4.76$; $SD = 1.30$) was slightly, although not significantly, higher than the mean for the subjects in the collective condition ($M = 4.44$; $SD = 1.36$) (see Table 12 for descriptive statistics).

Table 13 presents the selected output from the SPSS one-way ANOVA test for metric cohesiveness. The results showed that the groups did not differ significantly ($F = .91$, $p = .34$). The effect size was .01 (SPSS partial η^2), which can be considered a small effect

(Huck, 2000). This could mean that, even if the sample size had been larger, the possibility of failing to reject the null hypothesis due to lack of differences between the two conditions, would have been high.

Table 12

Descriptive Statistics for the Metric Cohesiveness Posttest Data

	Descriptive Statistics	
	Mean	<i>SD</i>
Individual condition	4.76	1.30
Collective condition	4.44	1.36
Total	4.60	1.32

Table 13

Analysis of Variance for the Metric Cohesiveness Posttest Data

Source	SS	df	MS	F	p
Treatment	1.61	1	1.61	.91	.34
Error	107.69	61	1.77		
Total	109.30	62			

Creativity

There were two subtests grouped under the concept of creativity: fluency and originality. Both subtests were assessed with a 7-point Likert scale. The results obtained on these two subtests were each explored with a one-way ANOVA.

Fluency. Fluency was assessed with a 7-point Likert scale, with a 1 indicating that the composition could not be reproduced, and a 7 indicating that someone else could reproduce the composition without trouble. In order to determine if the groups differed before the treatment sessions, the data obtained from the fluency pretest scores were explored with an ANOVA. Prior to this analysis, the assumptions of homogeneity of variance and normality were checked. In the case of the fluency pretest scores for the individual condition, ($M = 3.76$; $SD = 1.73$), the distribution met the assumption of normality. The skewness level of .13 fell within the established confidence interval (-.81 to +.81). The obtained kurtosis level of -1.45 fell under the confidence interval (-1.58 to +1.58). In the case of the fluency pretest scores for the collective condition ($M = 3.86$; $SD = 1.83$), the distribution met the assumption of normality, with a skewness of .01 ($CI = -.81$ to $+.81$), and a kurtosis of -1.01 ($CI = -1.58$ to $+1.58$). The Levene test reported that the pretest scores for fluency met the assumption of homogeneity of variance at the .74 level. Because the assumptions of normality and homogeneity of variance were met, the fluency pretest scores were explored with a one-way ANOVA.

Given the fact that there were eight subtests being explored with one-way ANOVAs, the α level was adjusted with the Bonferroni technique ($.05 / 8 = .006$). The new α level was set at the .006 level. The results showed that the groups did not differ

significantly on the pretest scores ($F = .05$, $p = .81$). Therefore, the null hypothesis of no difference between the groups' posttest means for fluency was tested with a one-way ANOVA.

The original range of possible scores was 1.00 to 7.00. Once the scores assigned by the three judges were averaged, the range of obtained scores was from 1.67 to 7.00 for the individual condition ($M = 4.91$; $SD = 1.49$), and 2.33 to 6.67 for the collective condition ($M = 4.57$; $SD = 1.51$). Prior to running the ANOVA for the fluency posttest scores, the assumptions of homogeneity of variance and normality were checked. In the case of the fluency posttest scores for the individual condition, the distribution was considered to approach normality, with a skewness of $-.52$ ($CI = -.81$ to $+.81$), and a kurtosis of $-.82$ ($CI = -1.58$ to $+1.58$). In the case of the fluency posttest scores for the collective condition, the distribution was considered to approach normality, with a skewness of $-.14$ ($CI = -.82$ to $+.82$), and a kurtosis of -1.36 ($CI = -1.60$ to $+1.60$). The Levene test reported that the posttest scores for fluency met the assumption of homogeneity of variance at the .90 level. Because the assumptions of normality and homogeneity of variance were met, fluency was explored with a one-way ANOVA.

Given the fact that there were eight subtests being explored with one-way ANOVAs, the α level was adjusted with the Bonferroni technique ($.05 / 8$). After the Bonferroni adjustment, the α level was set at the .006 level. According to the statistical analysis, the mean for the subjects in the individual condition ($M = 4.91$; $SD = 1.49$), was slightly, although not significantly, higher than the mean for the subjects in the collective condition ($M = 4.57$; $SD = 1.51$) (see table 14 for descriptive statistics).

Table 15 presents the selected output from the SPSS one-way ANOVA for fluency. The results showed that the groups did not differ significantly ($F = .79, p = .38$). The effect size was .01 (SPSS partial η^2), which can be considered a small effect (Huck, 2000). This could mean that, even if the sample size had been larger, the possibility of failing to reject the null hypothesis due to lack of differences between the two conditions, would have been high.

Table 14

Descriptive Statistics for the Fluency Posttest Data

	Descriptive Statistics	
	Mean	<i>SD</i>
Individual condition	4.91	1.49
Collective condition	4.57	1.51
Total	4.74	1.50

Table 15

Analysis of Variance for the Fluency Posttest Data

Source	SS	df	MS	F	p
Treatment	1.78	1	1.78	.79	.38
Error	137.87	61	2.26		
Total	139.65	62			

Originality. Originality was assessed with a 7-point Likert scale, with a 1 indicating that the composition had no originality at all, and a 7 indicating that the composition had marked originality. In order to determine if the groups differed before the treatment sessions, the data obtained from the originality pretest scores were explored with an ANOVA. Prior to this analysis, the assumptions of homogeneity of variance and normality were checked. In the case of the originality pretest scores for the individual condition ($M = 3.47$; $SD = 1.18$), the distribution met the assumption of normality, with a skewness of .33 ($CI = -.81$ to $+.81$), and a kurtosis level of $-.31$ ($CI = -1.58$ to $+1.58$). In the case of the originality pretest scores for the collective condition ($M = 3.59$; $SD = 1.11$), the distribution met the assumption of normality, with a skewness of .79 ($CI = -.81$ to $+.81$), and a kurtosis of .48 ($CI = -1.58$ to $+1.58$). The Levene test reported that the pretest scores for originality met the assumption of homogeneity of variance at the .53 level. Because the assumptions of normality and homogeneity of variance were met, the originality pretest scores were explored with a one-way ANOVA.

Given the fact that there were eight subtests being explored with one-way ANOVAs, the α level was adjusted with the Bonferroni technique ($.05 / 8 = .006$). The new α level was set at the .006 level. The results showed that the groups did not differ significantly on the pretest scores ($F = .26$, $p = .61$). Therefore, the null hypothesis of no difference between the groups' posttest means for originality was tested with a one-way ANOVA.

The original range of possible scores was 1.00 to 7.00. Once the scores assigned by the three judges were averaged, the range of obtained scores was from 2.33 to 7.00 for the individual condition ($M = 4.48$; $SD = 1.49$), and 2.00 to 6.00 for the

collective condition ($M = 4.25$; $SD = .93$). Prior to running the ANOVA for the originality posttest scores, the assumptions of homogeneity of variance and normality were checked. In the case of the originality posttest scores for the individual condition, the distribution met the assumption of normality, with a skewness of .33 ($CI = -.81$ to $+.81$), and a kurtosis of $-.06$ ($CI = -1.58$ to $+1.58$). In the case of the originality posttest scores for the collective condition, the distribution met the assumption of normality, with a skewness of $-.09$ ($CI = -.82$ to $+.82$), and a kurtosis of $.16$ ($CI = -1.60$ to $+1.60$). The Levene test reported that the posttest scores for originality met the assumption of homogeneity of variance at the $.16$ level ($\alpha = .006$). Because the assumptions of normality and homogeneity of variance were met, originality was explored with a one-way ANOVA.

Given the fact that there were eight subtests being explored with one-way ANOVAs, the α level was adjusted with the Bonferroni technique ($.05 / 8$). After the Bonferroni adjustment, the α level was set at the $.006$ level. According to the statistical analysis, the mean for the subjects in the individual condition ($M = 4.48$; $SD = 1.15$), was slightly, although not significantly, higher than the mean for the subjects in the collective condition ($M = 4.25$; $SD = .93$) (see Table 16 for descriptive statistics). Table 17 presents the selected output from the SPSS one-way ANOVA for originality. The results showed that the groups did not differ significantly ($F = .77$, $p = .38$). The effect size was $.01$ (SPSS partial η^2), which can be considered a small effect (Huck, 2000). This could mean that, even if the sample size had been larger, the possibility of failing to reject the null hypothesis due to lack of differences between the two conditions, would have been high.

Table 16

Descriptive Statistics for the Originality Posttest Data

	Descriptive Statistics	
	Mean	<i>SD</i>
Individual condition	4.48	1.15
Collective condition	4.25	.93
Total	4.36	1.04

Table 17

Analysis of Variance for the Originality Posttest Data

Source	SS	df	MS	F	p
Treatment	.84	1	.84	.77	.38
Error	66.86	61	1.09		
Total	67.71	62			

CHAPTER V

SUMMARY, DISCUSSION, AND RECOMMENDATIONS

Summary

Purpose of the Study

The purpose of the study was to explore the effect that individual versus collective structured creative musical problem solving tasks had on students' compositional products. The null hypothesis stated that there would be no difference on scores obtained from students trained under collective condition with those obtained by students trained under individual condition. The original research questions were:

1. Is there a significant difference between the level of tonal and metric cohesiveness identified on the compositional products of fourth- and fifth-grade students who participated in individual creative problem solving activities, and the level of tonal and metric cohesiveness identified in the compositional products of students who participated in collective creative problem solving activities?
2. Is there a significant difference between the use of melodic and rhythmic patterns identified on the compositional products of fourth- and fifth- grade students who participated in individual creative problem solving activities, and the use of melodic and rhythmic patterns identified on the compositional products of fourth- and fifth- grade students who participated in collective creative problem solving activities?
3. Is there a significant difference between the level of fluency and of originality found in the compositions of fourth- and fifth- grade students

who participated in individual creative problem solving activities from that found on the compositions of fourth- and fifth- grade students who participated in collective creative problem solving activities?

Summary of Methodology

Subjects. The school chosen for the study was located in a suburban area in northern Texas. There were 4 fourth grade groups, and 4 fifth grade groups at the school. There were 32 fourth graders and 32 fifth graders participating in the study, making a total of 64 subjects. A graphic description of the sample is displayed in figure 2. Except for 1 subject assigned to the collective condition, who did not attend the posttest session, all subjects were present during the 5 weeks the experiment lasted.

Design. This experimental study used the nonequivalent control group design. The independent variable for this study was the condition under which students were exposed to solve a creative musical problem. For treatment A, students were asked to solve a creative musical problem task individually. For treatment B, students were asked to solve a creative musical problem task in groups of four. Subjects were randomly assigned to the treatment condition.

Treatment sessions. There were three treatment sessions. During these sessions, students were asked to compose an original piece of music. The task was designed as an open-ended problem. Subjects were asked to start their composition on middle C, but they did not receive any other musical instruction. During each treatment session, subjects were asked questions designed to guide them into the three stages of the creative problem solving process as described by Treffinger, Isaksen, and Dorval (1994): Understanding the Problem, Generating Ideas, and Planning for Action.

For treatment A, individual creative problem solving, there was a two-octave chromatic song bell at the location. For treatment B, collective creative problem solving, there were four two-octave chromatic song bells set in a circle, so that the four students would face each other while composing. Subjects in both conditions had paper and pencil at hand, but were neither asked nor forbidden to use them.

The researcher or research assistant in charge read the set of instructions, including the specific questions designed for each treatment session. Subjects assigned to the individual condition were asked to think about their answers. Subjects assigned to the collective condition were asked to discuss the questions until they had reached a common answer. Subjects indicated when they were ready to start working. After the 10-minutes allowed for the compositional process, or when the subjects indicated they were done, the person in charge recorded the compositional product twice. Subjects assigned to the collective condition were asked to perform the composition together. Treatment session products were recorded in order to familiarize subjects with the procedure, but the recordings were not used for any analysis.

Pretest and posttest procedures. The pretests took place a week before the treatment sessions. The posttests were completed the week after the treatment sessions. Scores obtained for the pretest were intended to be used to control for initial differences between the two groups. Scores obtained for the posttest were the dependent variable of the study. Subjects were tested individually for both the pretest and the posttest. Pretest and posttest procedures were basically the same as those that were followed for the three treatment sessions. Subjects were asked to compose an original piece of music on a two-octave chromatic song bell with a red tag marking

middle C. They were asked to start the composition on middle C, and were allowed 10 minutes to work on their compositions. The only differences between the treatment sessions and the pretest and posttest were that for the pretest and posttest all subjects worked on their compositions individually, and that there were no questions related to the problem solving process added to the instructions that were read to the subjects. For the pretest and posttest, subjects were asked to perform their final composition twice. However, only the first performances of the compositions were used for the analysis.

Subtests. There were eight subtests. Four subtests were grouped under the concept of pattern use. For the purpose of this study, the judges indicated the number of times the subjects made use of repeated melodic patterns, developed melodic patterns, repeated rhythmic patterns, and developed rhythmic patterns.

There were two subtests grouped under the concept of cohesiveness: tonal and metric cohesiveness. For these subtests, a 7-point Likert-scale was used, with a 7 representing very strong cohesiveness, and a 1 indicating no cohesiveness. There were two subtests grouped under the concept of creativity: fluency and creativity. Both variables were measured with a 7-point Likert-scale. In the case of musical fluency, a 1 indicated that the composition could not be reproduced, and a 7 indicated that someone else could reproduce the composition without trouble. In the case of originality, a 1 indicated that the composition had no originality at all, and a 7 indicated that the composition had marked originality.

Pilot Test. The pilot test took place in a different elementary school than the host school, with a population socio-economically equivalent to that of the host school. Two

research assistants were in charge of collecting the data for the pilot test. There were 16 subjects participating in the pilot test. Subjects were tested individually. Two of the tapes were randomly chosen to be used for the training of the three judges. The other 14 tapes were used to assess interjudge reliability.

The researcher met individually with each of the three judges in order to train them. During these sessions, the researcher explained the definitions of the concepts included in the evaluation forms. The judges also received these definitions in writing. The instructional letter that the judges received is displayed on Appendix G. The researcher and the judges assessed the two compositions randomly chosen for this purpose together, to ensure that there was a common criterion. After the training sessions, the judges were asked to assess the 14 compositions individually.

Because the levels of interjudge reliability were low after the first analysis, the researcher met once again with the three judges. For the second training session, the researcher met with the three judges at the same time, and discussed every concept included in the measurement instrument until all of them were in agreement. Then, the judges were asked to assess the pilot test's compositions for a second time. The data obtained from this second effort were used to establish the interjudge reliability of the measurement instrument.

Reliability. Interjudge reliability was determined prior to the study, using the data obtained during the pilot test. In order to determine the level of interjudge reliability, the proportion of agreement between judges was determined by adding the total number of times the judges agreed on the overall score, as well as for each of the eight subtests. Then, these numbers were divided by the maximum number of possible agreements or

disagreements. Scores ± 1 from each other were considered agreements. The resulting overall coefficient of interjudge reliability for the pilot test was 0.79, and it ranged from 0.77 to 0.9 for the three groups of subtests: pattern use, cohesiveness, and creativity. These levels were considered acceptable to proceed with the experiment.

The interjudge reliability was tested again for the pretest and the posttest. For the pretest, the overall level of interjudge reliability was 0.6, and it ranged from 0.5 to 0.66 for the three groups of subtests. For the posttest, the overall level of interjudge reliability was 0.59, and it ranged from 0.5 to 0.68. The level of interjudge reliability, therefore, decreased at each of the three measurements of the study, from the pilot test to the posttest.

Results

The homogeneity of variance of the two treatment groups' scores on the three groups of subtests was compared through a Box test in order to determine if the pretest scores could be used as covariates. Scores for the subtests grouped under the concept of pattern use, and for the subtests grouped under the concept of cohesiveness did not meet the assumption of homogeneity of variance. Therefore, it was decided that MANCOVA was not the appropriate statistical procedure, and that the posttest scores would be explored with individual ANOVAs. Scores for the subtests grouped under the concept of creativity met the assumption of homogeneity of variance. However, to maintain consistency with the statistical analyses of the study it was decided that these two subtests would also be explored with two one-way ANOVAs.

With the new statistical procedure, there were eight null hypotheses that were tested:

1. There would be no difference between the groups' posttest means for the use of repeated melodic patterns.
2. There would be no difference between the groups' posttest means for the use of developed melodic patterns.
3. There would be no difference between the groups' posttest means for the use of repeated rhythmic patterns.
4. There would be no difference between the groups' posttest means for the use of developed rhythmic patterns.
5. There would be no difference between the groups' posttest means for tonal cohesiveness.
6. There would be no difference between the groups' posttest means for metric cohesiveness.
7. There would be no difference between the groups' posttest means for fluency.
8. There would be no difference between the groups' posttest means for originality.

Hypotheses tests. Prior to the hypothesis testing, and in order to check if there were initial differences between the two conditions, the data obtained from the pretest were explored with a one-way ANOVA for each of the eight subtests. Prior to these analyses, the assumptions of normality and homogeneity of variance were examined for each subtest. The scores for the four subtests grouped under the concept of pattern use did not meet the assumption of normality. Therefore, and as suggested by Huck (2000), the sample data were transformed. The log transformation, the square root

transformation, and the inverse transformation were explored for each of the four subtests. The obtained values that more closely approximated normality were the ones used for all subsequent statistical tests.

The Levene test reported that the pretest transformed values for the subtests grouped under the concept of pattern use met the assumption of homogeneity. The transformed values were then explored with one-way ANOVAs. In the case of the other four subtests, tonal cohesiveness, metric cohesiveness, fluency, and originality, the assumptions of normality and homogeneity of variance were met. Therefore, these subtests' data were also explored with one-way ANOVAs.

Given the fact that there were eight subtests being explored with one-way ANOVAs, the α level was adjusted with the Bonferroni technique ($.05 / 8 = .006$). The new α level was set at the .006 level. The results showed that the groups did not differ significantly on any of the eight pretest scores. Therefore, it was now possible to test the null hypotheses of no difference between the groups' posttest means for the use of repeated melodic pattern, for the use of developed melodic pattern, for the use of repeated rhythmic pattern, for the use of developed rhythmic pattern, for tonal cohesiveness, for metric cohesiveness, for fluency, and for originality, with eight one-way ANOVAs.

Prior to running the ANOVAs on the posttest data, the assumptions of homogeneity of variance and normality were checked. Once again, the scores of the four subtests grouped under the concept of pattern use did not meet the assumption of normality. Therefore, and as suggested by Huck (2000), the sample data were

transformed. The obtained values that more closely approximated normality were the ones used for all subsequent statistical tests.

The Levene test reported that the posttest transformed values for the subtests grouped under the concept of pattern use met the assumption of homogeneity. The transformed values were then explored with one-way ANOVAs. The other four subtests, tonal cohesiveness, metric cohesiveness, fluency, and originality, met the assumptions of normality and homogeneity of variance. Therefore, these subtests' data were also explored with one-way ANOVAs. Since there were eight subtests being explored with one-way ANOVAs, the α level was adjusted with the Bonferroni technique. The new α level was set at the .006 level.

According to the results obtained from the statistical analyses, none of the eight null hypotheses were rejected. Hence, there were no statistically significant differences between scores obtained by subjects exposed to the collective condition and scores obtained by the subjects exposed to the individual condition on any of the eight subtests. In the case of the use of repeated melodic patterns, the range of obtained transformed scores for the individual condition was .00 to .87 ($M = .21$; $SD = .25$), and .00 to .1.26 for the collective condition ($M = .27$; $SD = .34$). In the case of the use of developed melodic patterns, the range of obtained transformed scores for the individual condition was .12 to .92 ($M = .44$; $SD = .19$), and .00 to .75 for the collective condition ($M = .41$; $SD = .20$).

In the case of the use of repeated rhythmic patterns, the range of obtained transformed scores for the individual condition was .02 to 1.00 ($M = .36$; $SD = .21$), and .08 to 1.00 for the collective condition ($M = .40$; $SD = .24$). The range of obtained

transformed scores for the use of developed rhythmic patterns for the individual condition was .00 to 2.24 ($M = .83$; $SD = .47$), and .00 to 2.00 for the collective condition ($M = .77$; $SD = .58$). In the case of tonal cohesiveness, the possible scores were 1.00 to 7.00. The range of obtained scores was from 1.33 to 7.00 for the individual condition ($M = 4.69$; $SD = 1.82$), and 1.33 to 6.67 for the collective condition ($M = 4.30$; $SD = 1.50$). In the case of metric cohesiveness, the possible scores were 1.00 to 7.00. The range of obtained scores was from 1.67 to 7.00 for the individual condition ($M = 4.76$; $SD = 1.30$), and 1.67 to 7.00 for the collective condition ($M = 4.44$; $SD = 1.36$).

In the case of fluency, the original range of possible scores was 1.00 to 7.00. The range of obtained scores was from 1.67 to 7.00 for the individual condition ($M = 4.91$; $SD = 1.49$), and 2.33 to 6.67 for the collective condition ($M = 4.57$; $SD = 1.51$). In the case of originality, the original range of possible scores was 1.00 to 7.00. The range of obtained scores was from 2.33 to 7.00 for the individual condition ($M = 4.48$; $SD = 1.15$), and 2.00 to 6.00 for the collective condition ($M = 4.25$; $SD = .93$).

Discussion

The results from the current study indicate that participating in collective creative problem solving experiences versus individual creative problem solving experiences did not have a significant effect on fourth- and fifth-grade students' compositional products. Subjects exposed to the individual condition obtained higher scores than subjects exposed to the collective condition on six of the eight subtests that were explored (use of developed melodic and rhythmic patterns, tonal and metric cohesiveness, and fluency and originality), but these differences were not significant. These results contrast with those obtained by Chiara, Schuster, Bell, and Wolery (1995), and Larey

and Paulus (1999) who found that, when the problem solving processes took place in collective settings, the resulting products tended to be scored significantly lower than when the products were created by subjects working individually. A quantitative study in the area of music education exploring individual versus creative problem solving experiences, against which to compare the current study's findings, was not found during the development of the current study.

The lack of difference found between the individual and the collective condition on the compositional products may have resulted from various factors. First, it is possible that exposing children to collective creative problem solving experiences as opposed to individual creative problem solving experiences has in fact no effect on children's compositional products. The effect sizes obtained in the eight analyses of variance carried out in the current study were small, which indicates that, even if the sample size had been larger, it is possible that the results would have been the same. The implications that these results could have for music education will be discussed later in this chapter.

The second factor that may have contributed to the lack of difference found between the individual and the collective condition on the compositional products was the compositional task. The task for the current study was designed to guide the subjects into the exploration of the three stages of creative problem solving as described by Treffinger, Isaksen, and Dorval (1994): Understanding the Problem for the first treatment session, Generating Ideas for the second treatment session, and Planning for Action for the third treatment session. One of the assumptions made for the current study was that it was possible to familiarize subjects with these stages of the

creative problem solving process by asking subjects specific questions, and without an open discussion between the subjects and the person in charge of the treatment sessions. In fact, the researcher observed indications that this was possible during compositional experiences with fourth- and fifth grade students that took place prior to the pilot test of the current study. However, it is possible that subjects participating in the current study were in reality not going through each of the three stages of creative problem solving (Treffinger et al., 1994) that were suggested for the treatment sessions. In any case, there were no means to guarantee they were.

In the case of the individual condition, it was almost impossible to know what the subjects were thinking after the questions were posed. There were instances in which the subjects started their musical explorations immediately after the questions were posed, which may suggest that they did not think about each question. In the case of the collective condition, there were instances in which the subjects started their musical explorations without having discussed each question until they had reached a consensus. It is possible that, if the subjects participating in the current study, who were assigned to the collective condition, had been asked to answer the questions, they would have been forced to interact, negotiating and compromising, with all of the benefits and/or conflicts that this may imply. This could possibly have yielded different results.

One circumstance that might have had an impact on the way the subjects reacted to the task that was presented to them was the presence of a stranger at every testing and treatment session. Subjects were randomly assigned to a person in charge for each of the five sessions during the study. Since there were six persons in charge of

the testing and treatment sessions, it was possible for the subjects to have a different person in charge of each session. While this procedure may have avoided a teaching-learning situation, which could have become a confounding variable, it could also have been a challenge for the current study. The subjects might have felt unease while composing and discussing the posed questions in front of a stranger. A different procedure may have produced different results.

A third factor that may have contributed to the lack of difference found between the individual and the collective condition on the compositional products was the length of the study. It is possible that three weeks of treatment were not enough to originate a difference between the two conditions. The length of the current study was determined not only by the requirements of the host school, but also by the characteristics required for a problem solving process to take place. As it was said earlier in this paper, once we know how to solve a problem, it is not a problem anymore. Treffinger, Isaksen, and Dorval (1994) stressed the idea that creative problem solving “might be termed a problem if, and only if, the person does not already know [the] solution” (p. 226). Therefore, it is possible that a longitudinal study, with more than three weeks of treatment, may allow for the exploration of the effect that individual versus collective creative tasks have on subjects’ compositions; but the task may not constitute a problem solving situation any more. The students would be too familiarized with the task for it to be considered a problem to be found and solved. Thus, while it may be possible that a longitudinal study may have produced significant differences between the two conditions, the task may have been a creative task, but probably not a problem solving task anymore.

Another factor that may have contributed to the lack of significant difference between the two conditions explored in the current study may be the measurement instrument. It is clear that overcoming problems associated with the measurement instrument was a challenge for the current study. For instance, the judges had to be trained two times before acceptable levels of reliability could be reached. Once acceptable levels of reliability were reached, they were not maintained over time. The reliability coefficients went down from the pilot test to the pretest, and then to the posttest.

One possible explanation for the problems associated with the measurement instrument is that the measurement instrument was not appropriate for the assessment of compositions that resulted from open-ended tasks. The open-ended task used in the current study may have left room for subjects to create compositions that lacked a definite structure that could be objectively assessed. The task for the current study was designed using Getzels and Csikszentmihalyi's (1975) description of the creative problem solving experience. Accordingly, a general task was presented, but the problem solvers were expected to discover the specific question or problem. According to Getzels and Csikszentmihalyi, this type of task would encourage problem solvers to not only make use of their reasoning, but also to use their imagination and creativity so that they can find, articulate, and solve the problem. For this type of open-ended compositional task, there are no set rules or limits. Subjects are expected to use what they already know, both intuitively and consciously, about music. As described by McCoy (1999), "rather than following a specific scope and sequence of activities en

route to a prespecified end, students would be allowed to frame the problem as it appears to them and then follow it to its most satisfactory resolution” (p. 16).

Open-ended compositional tasks, though, must not be mistaken with improvisational tasks. For the compositional tasks, subjects are allowed time to plan and rehearse their musical ideas before they have to perform their compositions on request. In contrast, improvisational tasks entail that the subjects plan their musical ideas *during* the performance. There is no time for planning ahead or for rehearsal. Subjects participating in the current study were allowed time to plan and to rehearse their compositions. Accordingly, the creative problem solving products of the current study were considered compositions.

Kratus (1989) found that children without formal training in composition, and with the same age range as that of the subjects participating in the current study, were capable of producing compositions with high levels of tonal and metric cohesiveness, using melodic and rhythmic patterns. Accordingly, there were two expectations for the current study. First, that subjects participating in the current study would be able to generate compositions with evident levels of tonal and metric cohesiveness, using melodic and rhythmic patterns. Second, that the chosen measurement instrument would allow for an objective exploration of tangible structural characteristics of the creative musical problem solving products, such as cohesiveness and pattern use, as well as for more subjective characteristics, such as originality and fluency. However, it seems that the outcome of the open-ended compositional task used in the current study may have been compositions with no common structural characteristics, such as melodic and rhythmic patterns. Hence, the assessment of the creative problem solving products of

the current study may have depended on the judges' subjective criteria, instead of the rubric criteria that was provided for them.

When interviewing the judges individually, both during and after they had finished the assessment of the compositions, they reported the subtests grouped under the concept of pattern use as the hardest ones to evaluate. They found it to be very difficult and time consuming to determine what could be considered a melodic or rhythmic pattern, and then what could be considered a development or repetition of that pattern. For a music educator trained under the Orff methodology, as was the case of Judge B, a pattern needed to have a contrast for it to be considered a pattern. For example, a group of four quarter notes could not be considered a rhythmic pattern. It needed to have a contrasting rhythmic figure, such as two eighth-notes in place of one of the quarter notes, for it to be considered a pattern. In contrast, for a music educator trained under the Kodály methodology, as was the case for judges A and C, the same group of 4 quarter notes could be considered a pattern, as long as they perceived it as such, mainly because of the melodic combination attached to it.

Possibly as a result of the discrepancies between the judges' criteria, the levels of interjudge reliability for the current study were not maintained. It seemed as if the further in time the judges were from the training sessions, the more they went back to their personal definitions and criteria. Pilsbury and Alston (1996) observed the difficulty of achieving unanimity between experienced assessors' evaluation of children's compositions, even when they were asked to work with common criteria. The tendency for judges to use their own set of criteria when judging children's compositions, even

when there were criteria intended to be common between them, seems to have been a likely phenomenon in the current study.

Although there was an interest in the current study on the exploration of tangible structural characteristics of the creative musical problem solving products, it is possible that the use of Amabile's consensual technique (1982, 1996) for the assessment of creative products could have been a more appropriate measurement instrument. As opposed to some quantitative measurement instruments for creative products, Amabile's technique (1983) requires judges to assess creative products by using their own subjective set of criteria. Hence, Amabile's consensual technique would take advantage of the judges' tendency to use their own set of subjective criteria when judging children's compositions. According to Byrne, Macdonald, and Carlton (2003), this could be done while not necessarily sacrificing the exploration of more objective criteria. In a study examining the link between the concept of flow (Csikszentmihalyi, 1992) and the creative output of college students' compositions, Byrne et al. (2003) found a significant correlation between standard criteria scores and the creativity scores produced with Amabile's consensual technique. The authors suggested that,

[The] Consensual Assessment Technique has the potential to be used as an alternative to the more problematic criteria-based approaches.... For example, studies, including the present one, suggest that experienced raters may take into account the many factors of criteria-based methods while subjectively rating musical compositions on the single dimension of creativity. (p. 286)

Unfortunately, the findings of Byrne et al. were published after the current study had already taken place, and therefore Amabile's consensual technique was not considered an option that would have allowed for the exploration of tangible structural characteristics of the creative musical problem solving products.

It is also possible that the measurement instrument used in the current study might be more appropriate and reliable if the creative problem solving task were not as open-ended. If children were asked to compose a piece of music using patterns, both rhythmic and melodic, with a specific length, it might be easier for the judges to perceive them as patterns. This could be done by determining the harmonic sequence over which the subjects were to compose a melody, a rhythmic pattern for which the subjects were to compose a melody, a melody for which the subjects were to compose a rhythmic pattern, or even the number of pitches that the subjects were to use. This would make it easier and clearer for the judges to assess if the subjects were developing or repeating a pattern at some point in the composition. But with the characteristics of an open-ended task as the one used for the current study, it became too complicated for the judges to agree on the identification of patterns, let alone the assessment of them as repetitions or developments.

While having children composing with predefined musical concepts might be an important music education task when children are being taught those specific musical concepts, it has been documented as being important to let children explore and express their musical ideas without rigid restrictions (Burnard, 2002; Wiggins, 1999).

Kaschub (1999) suggested that,

Music educators should consider designing project guidelines that engage students with particular music concepts, ideas, and experiences, and should give greater freedom to students engaging in compositional activities solely for the purpose of exploring and developing their personal compositional styles. (p. 250)

When providing students with open-ended compositional tasks, it is also important to keep in mind that, even when children have not yet mastered the knowledge of basic musical concepts, they have been acculturated into a music

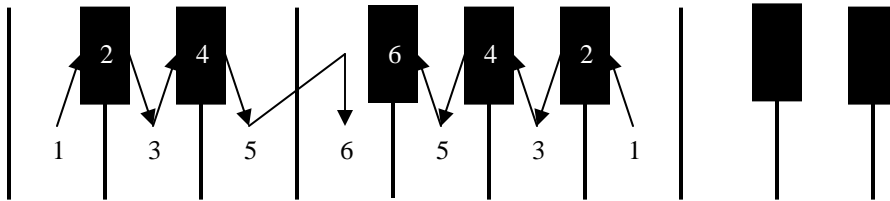
tradition that has already created musical ideas in their minds. As Wiggins (1999) pointed out,

We have operated on the assumption that it is the teacher who is the expert and that the students have little or nothing to bring to the situation. In actuality, students bring to our classrooms an enormous amount of musical knowledge of our culture gleaned from living within that culture....Because of our long tradition of teacher control and our assumptions that students do not really know very much about music, teachers sometimes construct creative assignments in ways that not only fail to promote creative thought but may actually hamper it...compositional assignments with restrictive parameters can cause students to focus on the extramusical, nonexpressive aspects of a project, and this can hamper rather than enable or promote the creative process. (pp. 30-31)

We must keep in mind, however, that subjects with different levels of musical skills may produce compositional products through open-ended tasks that have different levels of cohesiveness and pattern use. For instance, it is possible that subjects with higher level of musical skills than the ones of the subjects participating in the current study, may have produced compositions with more clearly defined rhythmic and melodic patterns. When music educators or researchers expect to obtain compositions with high levels of cohesiveness and pattern use through an open-ended task, it is possible that they would have better results the higher the subjects' musical skills levels are.

Another characteristic of the compositional task that became noticeable as the current study progressed, not only for the researcher and research assistants, but also for the judges who were not even present during the testing and treatment sessions, was a visual component of the musical instrument. It seemed that some subjects were relying on the visual aspect of the task more than on the sound aspect of it. Some subjects based their compositions on the shapes that the melody would create on the keyboard, regardless of the tonal and/or metric result. Figure 1 illustrates an example of this type of melody.

Figure 1. Example of a melodic sequence based on the visual context. The subject played two notes simultaneously, starting with the major 7th notated as number 1.



In the example illustrated in figure 1, there is no sense of tonality. It seemed that the subject focused on the combination of white and black keys. It is possible that the compositional products would be different if the subjects were asked to compose on an instrument that allowed for chromatic exploration, but that did not have the visual component that the keyboard had. The students could have focused on the sound without being distracted by the visual characteristics of the instrument. Kratus (2001) investigated the effect that different melodic configurations on an Orff xylophone had on children's compositional processes and products. Even though Kratus covered the letter names of the bars with dark tape, he was not exploring the visual component of the instrument. He was interested in exploring the differences between compositional processes and products obtained by subjects distributed among four different conditions: (1) pentatonic with 10 bars; (2) pentatonic with 5 bars; (3) melodic minor with 10 bars; and (4) melodic minor with 5 bars. No research study was found which addressed the question of whether the visual characteristics of a chromatic instrument affect children's compositional processes and products.

Recommendations for Music Educators

If, as the results of the current study may indicate, exposing children to either collective or individual creative problem solving tasks does not have an effect on the resulting individual subjects' compositional products, music educators interested in observing specific characteristics of students' compositional products could expose students to either collective or individual condition according to their individual circumstances and interests. While having students composing under one of these two conditions, music educators could explore the resultant compositional products knowing that it might be possible that the compositional products would have had similar levels of cohesiveness, pattern use, and creativity, if they had been composed in the other condition.

The reasons for the design of either collective or individual creative problem solving tasks may vary, from classroom size to the interest on the different benefits that each of these compositional processes could provide for the students. For instance, it might be difficult to create a truly individual condition in a small classroom in which students can not be isolated from each others' musical ideas. The true individual condition might only be possible in a large classroom, or in a setting where students can go into different rooms to be isolated and where they are not being influenced by others' musical ideas or comments. When dealing with a situation where students can not be isolated, music educators interested in the exploration of individual students' compositional products might design a collective compositional activity, while still being able to observe the characteristics of the compositions, knowing that it is possible that

these characteristics would have been similar if they had been produced under a truly individual condition.

In another instance, music educators could be interested in the different benefits that either the collective or the individual compositional processes could provide for the students. Both Christensen (1992) and Kaschub (1999) reported that the individual and the collective compositional process contributed in different ways to the students' musical experiences. More specifically, Kaschub (1999) suggested that "students should be exposed to both collaborative and individual compositional experiences so that they may enjoy the support of their peers and also experience the freedoms afforded by individual work" (p. 250). Music educators could also be interested in observing the levels of cohesiveness, creativity, and/or the ways in which students make or do not make use of patterns. According to the current study's results, music educators could observe these specific characteristics of individual students' compositional products regardless of the condition under which the compositional product was created.

Caution should be taken, however, when generalizing the results of the current study, first and foremost, because the sample of the current study was not selected randomly. It was a convenience sample, and therefore the results can not be generalized to other fourth- and fifth-grade students. Additionally to that, caution should also be taken when generalizing the results of the current study because of the problems that were faced, and which have already been described earlier in this chapter. Particularly, music educators need to make sure that they use measurement

instruments that have been tested and reported as being valid and reliable for the specific type of compositional task that they are planning to use.

Recommendation for Future Research

The first recommendation for future research is related to the three different problem solving stages suggested for the compositional task by Treffinger et al. (1994). It would be necessary to explore in future research if asking students to go through the three stages of the compositional process (Understanding the Problem, Generating Ideas, and Planning for Action), in an overt way, as compared to a covert way, would have significantly different effects on students' compositional products. This would also allow for the exploration of the ways in which these stages take place in the creative musical problem solving context, both at the individual level as well as at the collective level. One option for the exploration of this variable could be to have the person in charge of the treatment sessions listening to the subjects' answers without discussing them, thereby avoiding a teaching-learning situation, which might become a confounding variable. Another option would be to have only one person in charge of the treatment sessions, so that there would be no possibility of varying teachers' characteristics intervening on the creative problem solving process.

Another recommendation for future research is related to the measurement instrument. There is a need for studies exploring two possibilities. First, there have been some studies exploring children's musical creativity (Bangs, 1992; Brinkman, 1999; Daignault, 1997; Hickey, 1995, 1996, 2001) that have made use of Amabile's Consensual Assessment Technique (1982, 1996), and that have reported acceptable levels of reliability. All of these studies have used open-ended compositional tasks.

Nevertheless, it would still be necessary to explore the use of Amabile's consensual technique in future research in order to find out if in fact experienced raters "take into account the many factors of criteria-based methods while subjectively rating musical compositions on the single dimension of creativity" (Byrne, et al., 2003, p. 286). It would also be necessary to replay the current study using Amabile's Consensual Assessment Technique in order to find out, first and foremost, if consistent levels of interjudge reliability could be reached, and then if this measurement instrument would produce results similar to the ones found in the current study.

The second recommendation for future research related to the measurement instrument would involve the use of the measurement instrument of the current study with closed-ended tasks. There is a need to explore if the measurement instrument of the current study would produce higher levels of interjudge reliability when exploring individual versus collective creative problem solving with a closed-ended task. In that case, subjects could be asked to compose a piece of music using a specific harmonic sequence over which they would compose a melody, a rhythmic pattern for which the subjects were to compose a melody, a melody for which the subjects were to compose a rhythmic pattern, or even determining the number of pitches that the subjects were to use.

There were other variables in the study that became apparent as the treatment sessions progressed, and that would need further investigation. The first variables that became noticeable for the researcher and research assistants were gender and ethnicity. All of the assistants commented that, when there was one girl grouped with boys in a team of 4 subjects, the girl would tend to remain quiet, and would not share

her suggestions with the boys. The same would happen when a subject pertaining to a minority ethnic group was grouped with a majority of white students. Research assistants and the researcher herself noticed that sometimes these individuals were in fact creating melodies while the rest of the group was discussing or working on the composition, however, they would not speak up. However, these variables were not explored systematically in the current study. The role that the variables of gender and ethnicity could play in the collective compositional processes and products needs further exploration.

Another variable that would need further exploration through systematic inquiry is the visual and sound characteristics of the musical instruments in which the students compose. It would be important for future research to explore if the use of different musical instruments, with different visual and sound characteristics, results in contrasting compositional processes and products. While it is important to recognize the efforts that recent studies have made in exploring the role of computers on children's compositional processes and products (Daignault, 1996; Emmons, 1998), there is still a need for further exploration focusing on the effect that the use of musical instruments usually available in the general music classrooms, musical instruments with different visual characteristics, and musical instruments with different sound characteristics, could have on students' compositions.

However, before the effect that any variable could have on students' compositional products can be explored, appropriate measurement instruments need to be specifically designed for the compositional tasks at hand. The reliability and validity of these measurement instruments should be tested before the effect of any other

variable on children's compositional products can be explored. The experiments designed to test these measurement instruments need to be replicated, so that the measurement instruments are tested under a variety of circumstances. Then, the results must be reported with honesty and clarity, so that more advancement of this field is possible.

APPENDIX A

CONSENT FOR PROJECT PARTICIPATION FORMS (MAIN STUDY)

Dear Parent:

APPROVED BY THE UNT IRB
FROM 5/18/03 TO 5/18/04
DB

This is to inform parents of a music education research project which will soon be conducted at the Sam Houston Elementary School that could involve your child. A description of the project follows. Additional information regarding the project, as well as a copy of any measurement instrument which might be used, will be available in the school's main office. Participation in this project is voluntary and no consequences will follow from declining to participate or withdrawing from the study once it has begun. Also, the results of the tests will not affect the students' grades or any other academic standing. The music teacher at Sam Houston Elementary School, Ms. Jennifer Galbreath, will participate in this project only by allowing the selected students to leave the music classes when required by the study and by providing the required space. However, she will not be directly involved in the project. Please, if you and your child agree to participate in the study, sign the Consent To Project Participation form and return it to the school's music teacher, Ms. Jennifer Galbreath. If you agree to have your child participate in the research described here but later change your mind, you have the option of withdrawing your child from the project at any moment. Return the signed form promptly to Ms. Jennifer Galbreath. Note that given time and space limitations, not all children will participate in the project. Children participating in the project will be randomly selected. Therefore, signing the Consent To Project Participation form does not guarantee your child's participation in the project.

If you have further questions please feel free to contact the school's principal, the school's music teacher Mrs. Laura Kinnaird, or me.

Sincerely,

Beatriz E. Aguilar, doctoral candidate
University of North Texas
College of Music
bea0005@unt.edu
Phone: 940-891-4251

Dr. Debbie Rohwer, Main Advisor
University of North Texas
College of Music
Coordinator of the Ph.D. program in Music
Education
Phone: 940-369-7538

RESEARCH CONSENT FORM

Subject Name: _____

Date: _____

Title of Study: The Effect Creative Musical Problem Solving in Team versus Individually on 4th and 5th Grade Students' Compositional Products. (Main Study)

Main Investigator: Beatriz E. Aguilar

Co-investigators: Rodrigo Villanueva

Before agreeing to participate in this research study, it is important that you read and understand the following explanation of the proposed procedures. It describes the procedures, benefits, risks, and discomforts of the study. It also describes your right to withdraw from the study at any time. It is important for you to understand that no guarantees or assurances can be made as to the results of the study. **This study has been reviewed and approved by the UNT committee for the Protection of Human Subjects (940-565-3940).**

Purpose of the study and how long it will last:

The purpose of the study is to explore the effects that training students to solve creative musical problems in teams versus individually could have on their compositional products. Thirty two 4th graders and Thirty two 5th grade students attending Sam Houston Elementary School will participate in the study. Given time and space limitations, children participating on this pilot study will be randomly selected from those whose parents sign the Consent To Project Participation. The selected students will be taken out of their music class for 15 minutes on _____, 2003, on _____, 2003, on _____, 2003, on _____, 2003, and once again on _____, 2003.

Description of the study including the procedures to be used:

For pre and posttests, students participating in this pilot study will be taken individually into a room, where a two-octave xylophone will be placed. There will also be a clock placed where students can clearly see it, and a tape recorder with a unidirectional microphone. Once in the room, the researcher will ask the student to solve a compositional task. This compositional task does not have a correct or wrong answer. Once the directions have been read and all the questions answered, the researcher will turn the tape recorder on, and move to another part of the room. The student will have ten minutes to work on the task. Once the student indicates s/he has finished, or the allowed time has passed, the researcher will ask the student to perform the compositional product and the recording will be turned on. After the student has played the song, the recording will be turned off, and the student will be taken back to the music classroom. For treatment sessions, some students will work individually on similar tasks than the one already described, and some of them will work on teams of four. The audiotapes of each student's final creative problem solving products will be assigned with a code number to ensure anonymity. Once coded, all the audiotapes will be recorded in a single compact disk, which will be analyzed by two independent judges, both of them general music educators with extended professional experience.

Research Consent Form -Page 1 of 3 _____ Participant's initials

RESEARCH CONSENT FORM (Continued)

Description of procedures/elements that may result in discomfort or inconvenience:

The only inconvenience foreseen in this project is the student's absence from his/her music lesson for fifteen minutes. However, it is important to note that the task required in this project is an activity designed to provide students with music creative opportunity appropriate for their age, and suggested by the TAKs. It is also important to note that the study is designed to develop student's creative musical problem solving skills, and that it is expected that all participants will learn some creative problem solving strategies.

Description of the procedures/elements that are associated with foreseeable risks:

There are not procedures or elements involved in this research project that we foresee could constitute a physical or emotional risk for the participants.

Benefits to the subjects or others:

Encouraging the development of critical thinking has become a constant concern for North American educators. In Goals 2000: Educate America Act, it is stressed that the percentage of all students who demonstrate the ability to reason, solve problems, apply knowledge should increase substantially. Since the establishment of these Goals, it has become imperative to favor the development of effective thinking skills, of critical thinking skills in the schools. However, research on critical thinking and problem solving skills is relatively new in music education. Clearly, music educators could highly benefit from research that explores the creative problem solving products as they take place in the music classroom. This research project will explore the effects that training students to solve creative musical problems in teams versus individually could have on their compositional products. Students participating in this study will have the opportunity to participate in a creative problem solving activities.

Confidentiality of research records:

Children participating in this project will be randomly selected from those whose parents sign the Consent To Project Participation. Once the participants have been selected, they will be assigned a code number. Their identity will not be released to a third party at any time. The pretest and posttest audiotapes will be assigned a code number, and the identity of the participant will be known only by the main researcher. The final compositions will be recorded from the audiotapes to a single compact disk, in which the compositions will be identified as "subject 1," "subject 2," and so on. In this way, the identity of the students will be unknown to the judges or any other third party. Once transferred into the compact disk, the audiotapes will be storage by the main researcher for her records. The results of the tests will be used to establish the reliability of the measurement instrument, and will not affect the students' grades or any other academic standing.

Research Consent Form -Page 2 of 3 _____ Participant's initials

CONSENT TO PROJECT PARTICIPATION

Sam Houston Elementary School

You are making the decision about whether or not have your child participate in this study. Your signature indicates that you have decided to allow your child to participate, that you have read (or have had read to you the information provided in this Consent Form and that you have received a copy of it. You should also be aware that participation in this project is voluntary and no consequences will follow from declining to participate or withdrawing from the study once it has begun, and that the results of the tests will not affect the students' grades or any other academic standing. Finally, you should know that **this study has been reviewed and approved by the UNT committee for the Protection of Human Subjects (940-565-3940).**

Student's Parent's or Guardian's Signature

Date

Investigator's Signature

Date

ASSENT OF CHILD

_____(name of child) has agreed to participate in the research project The Effect of Creative Musical Problem Solving in Team versus Individually on 4th and 5th Grade Students' Compositional Products.

Student's Signature

Date

APPROVED BY THE UNT IRB
FROM 5/19/03 TO 5/18/04
[Signature]

Research Consent Form -Page 3 of 3 _____ Participant's initials

APPENDIX B

SAMPLING AND TREATMENT ASSIGNMENT

Figure 2. Sample size and treatment assignment

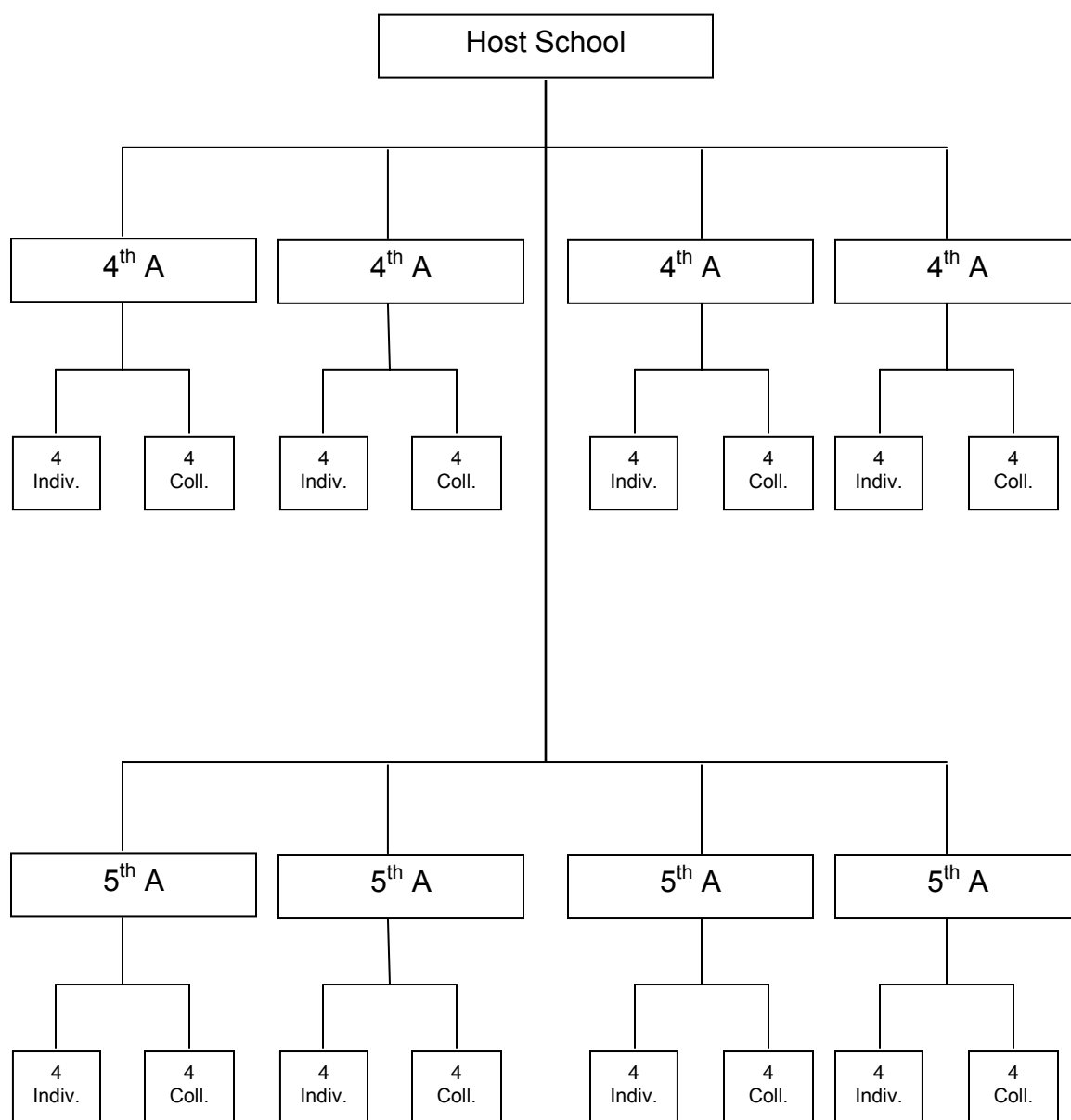


Figure 3. Example of treatment and testing sessions assignments

4A (Monday)								
Name	Pretest Code #	Posttest Code #	Treatment assignment	Pretest Tester	Posttest Tester	Treatment 1 Tester	Treatment 2 Tester	Treatment 3 Tester
XXX	1	9	Individual	Beatriz (2 nd)	Rodrigo (1 st) 5WR	Dimitri SEWR	Carlos 5WR	Rodrigo 3WR
XXX	4	12	Individual	Beatriz (1 st) 3WR	Dimitri (2 nd)	Carlos 5WR	Dimitri (2 nd) 3WR	Dimitri 5WR
XXX	5	13	Individual	Rodrigo (1 st) SEWR	Carlos (2 nd)	Beatriz (1 st) 3WR	Rodrigo 2WR	Carlos (1 st) 2WR
XXX	7	15	Individual	Rodrigo (2 nd)	Beatriz (2 nd)	Rodrigo 2WR	Beatriz SEWR	Beatriz SEWR
XXX	2	10	Group	Carlos (2 nd)	Beatriz (1 st) 3WR			
XXX	3	11	Group	Joe (2 nd)	Rodrigo (2 nd)			
XXX	6	14	Group	Carlos (1 st) 2WR	Dimitri (1 st) 2WR	Beatriz 2 nd PEWR	Dimitri 1 st PEWR	Carlos 2 nd PEWR
XXX	8	16	Group	Joe (1 st) 5WR	Carlos (1 st) SEWR			

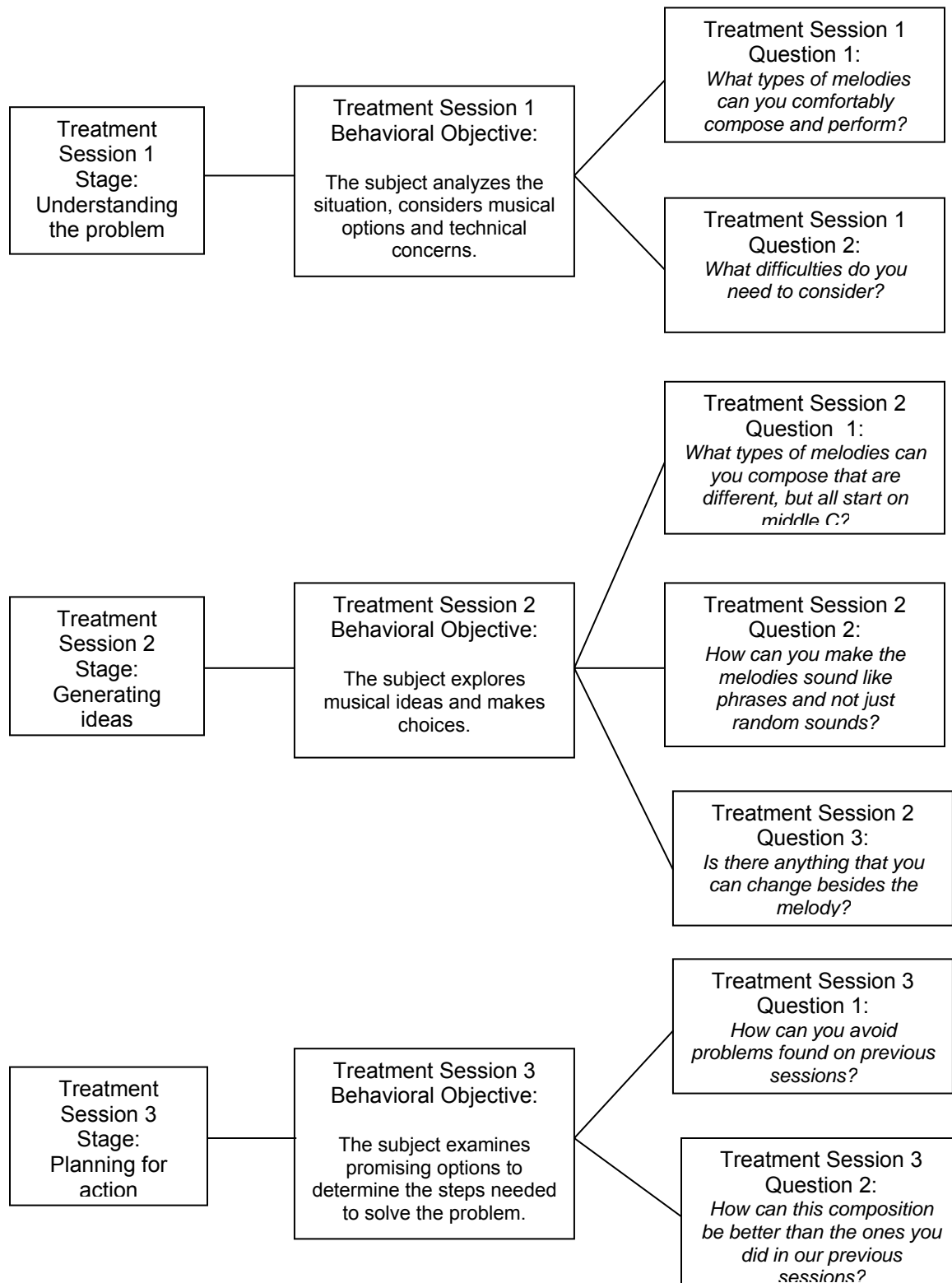
Figure 4. Examples of Comment Sheets

Pretest	Subject Name: <u>X, X (Beatriz, 4th Grade, 1st Subject, 2WR)</u>		Subject's Code Number: <u>1</u>	
	1. Time spent on the compositional process:			
	10 Minutes		Other: _____	
	2. Did the subject use the paper and pencil to write down the composition?			
	Yes		No	
	3. Comments:			

Treatment Session 2	Subject Name: <u>X,X; X,X; X,X; X,X (Beatriz, 4th Grade, PEWR)</u>		Team's Code Number: <u>1</u>	
	1. Time spent on the compositional process:			
	10 Minutes		Other: _____	
	2. Did the subject use the paper and pencil to write down the composition?			
	Yes		No	
	3. Comments:			

APPENDIX C
TREATMENT DESIGN

Figure 5. Treatment Design



APPENDIX D

JUDGES' INSTRUCTIONS LETTER AND SCORING FORMS FOR MAIN STUDY

MUSICAL CREATIVE PROBLEM SOLVING: DOCTORAL DISSERTATION PROJECT

Dear judges:

Thank you very much for participating in this study. The time and work that you are going to invest on it are of great importance and are highly appreciated. The experimental stage has already taken place, and it is now the moment to assess the children's compositions. Along with this Instructions Form, you are receiving four CDs, the transcriptions of the compositions that you will assess, and enough Judges' Scoring Forms for the pretests and for the posttests.

The CDs you are receiving contain 127 compositions, 64 are the pretests and 63 are the posttests. The two CDs with the pretests contain 32 compositions each. The other two CDs with the posttests contain 32 compositions the first one, and 31 the second one. The posttest of one subject is missing. To make this process easier for you, each composition is contained in a single track. Compositions were recorded in the order of the subjects' code number. Before each composition, you will hear the subject's number.

Please, before start scoring the composition, write down your name and the subject's number on the space provided in the Judges' Scoring Form. **(If you are working with the pretests, make sure you are using Judges' Scoring Forms that indicate *Pretest*. Similarly, if you are working with the posttests, make sure you are using Judges' Scoring Forms that indicate *Posttest*).** Then, listen to the composition once without making any marks on the Scoring Form. Use this first listening to become familiar with the composition. Then, you can listen to the compositions as many times as you need in order to score them.

All the compositions have been transcribed for you. The transcriptions are intended to help you during the assessment process. They are not, however, the only possible written representation of the compositions. Therefore, take into consideration the following:

1. In order to ensure anonymity and unbiased assessment of the compositions, subjects were assigned two different code numbers. The code numbers used for the pretest are not the same as the subjects' code number used for the posttest. That is, subject number 1 in the pretest **is not** subject number 1 in the posttest.
2. All of the compositions have been transcribed, but it was not an easy process.
 - a. In some instances, some compositions lacked a clear sense of beat, which made difficult to determine notes values. In these cases, all of the notes were written as quarter notes.
 - b. Unless the performance of the composition indicated very clearly a time signature, all compositions were written in 4/4, when the beat was subdivided in two, and on 6/8 when the beat was divided in three.
 - c. The computer program was set so that each line included four measures for most of the compositions. There were some instances where, for example, the composition lasted

five measures. In cases like this, the five measures were inserted in a single line. The important issue to stress here is that the visual distribution of the measures was not intended to illustrate the structure of the composition.

- d. In some examples, subjects appeared to have some performance mistakes, for example, hitting the wood part of the bells instead of the key. In most cases, it was possible to write down the note that the subject intended to play, either because some subjects wrote down their compositions, or because it seemed clear that the subject played the intended note immediately after the mistake. However, when the subject did not write the composition down, or the correction of the mistake was not clear, the mistake has been ignored and was not included in the transcription.
3. There is not posttest for the subject number 61. In the CD containing posttest compositions by subjects 33 to 64, you will not find subject 61's composition.
4. **If you find that a mistake was made on the compositions' transcription, and that the correction of it would have an impact on your assessment of the composition, please get in touch with me and let me know of any change.** I will inform the other judges if any change has been made. It is really important that all of us are looking at the same score.

It is also important to remind you that, for the purposes of this study, the compositions are going to be evaluated in terms of their use of patterns with four subtests, cohesiveness with two subtests, and creativity with two subtests. For the first four subtests, related to melodic and rhythmic pattern use, you will indicate the number of times the subjects made use of repeated melodic patterns, developed melodic patterns, repeated rhythmic patterns, and developed rhythmic patterns. You will indicate this by writing on the Judges' Scoring Form one tally for each time you heard a recognizable melodic or rhythmic pattern being repeated or developed. At the end, the total number of tallies for each type of pattern will be summed in order to obtain the total number of occurrences of each type of pattern during the composition. Repeated melodic patterns, developed melodic patterns, repeated rhythmic patterns, and developed rhythmic patterns are defined as follows:

Melodic Pattern: two to seven pitches that form a distinct pitch pattern that is perceived as a unified whole.

Repeated Melodic Pattern: a melodic pattern that is identical to a previous occurring melodic pattern.

Developed Melodic pattern: a melodic pattern that is similar to, yet different from, a previously occurring melodic pattern. In most cases, the rhythm is held the same, the melody changes, and the melodic contour is similar.

Rhythmic Pattern: two to seven durations that form a distinct duration pattern that is perceived as a unified whole.

Repeated Rhythmic Pattern: a rhythmic pattern that is identical to a previous occurring rhythmic pattern.

Developed Rhythmic Pattern: a rhythmic pattern that is similar to, yet different from, a previously occurring rhythmic pattern. In most cases, the melody is the same or similar and the rhythm changes.

For the next two subtests, tonal and metric cohesiveness are defined as follows:

Tonal Cohesiveness: The degree to which the pitches in a composition are constructed around a tonal center or tonal centers. (1 = no tonal cohesiveness, 7 = very strong tonal cohesiveness)

Metric Cohesiveness: The degree to which the durations in a composition are constructed of regularly occurring accented and unaccented beats. (1 = no metric cohesiveness, 7 = very strong metric cohesiveness)

The last two subtests are a fluency and an originality test. Fluency is understood as the ability to generate a number of clearly defined musical ideas, which would facilitate its reproduction by a person other than the composer. In the case of musical fluency, a one indicates that the composition could not be reproduced, and a seven indicates that someone else could reproduce the composition without trouble. You can mark any number of this scale that better expresses your assessment of the composition's level of fluency. Musical originality is defined as the ability to create a composition that contains elements of unique expression. In the case of originality, a one indicates that the composition had no uniqueness at all, and a seven indicated that the composition had marked uniqueness.

If you have any question, feel free to contact me. During the month of December 2003, it will only be possible to contact me via e-mail. After that, you can reach me at home also. Let me know about any problem or concern you might find during this process. Your comments and suggestions are highly valued and appreciated. I hope you find this project as interesting as I do. Thank you very much!

Sincerely,

Beatriz E. Aguilar

Phone: (940) 300-8782

E-mail: betina68@yahoo.com

Judge's Scoring Form

Subject Code Number: _____

Type of pattern	Tallies for the number of occurrences	Total
Repeated Melodic Patterns		
Developed Melodic Patterns		
Repeated Rhythmic Patterns		
Developed Rhythmic Patterns		

1	2	3	4	5	6	7
No metric cohesiveness						Very strong metric cohesiveness

1	2	3	4	5	6	7
No Originality						Marked Originality

APPENDIX E

CONSENT FOR PROJECT PARTICIPATION FORMS (PILOT TEST)

Dear Parent:

APPROVED BY THE UNT IRB
FROM 5/19/03 TO 5/18/04

This is to inform parents of a music education research project which will soon be conducted at the Town Center Elementary School that could involve your child. A description of the project follows. Additional information regarding the project, as well as a copy of any measurement instrument which might be used, will be available in the school's main office. Participation in this project is voluntary and no consequences will follow from declining to participate or withdrawing from the study once it has begun. Also, the results of the tests will not affect the students' grades or any other academic standing. The music teacher at Town Center Elementary School, Mrs. Laura Kinnaird, will participate in this project only by allowing the selected students to leave the music classes when required by the study and by providing the required space. However, she will not be directly involved in the project. Please, if you and your child agree to participate in the study, sign the Consent To Project Participation form and return it to the school's music teacher, Mrs. Laura Kinnaird. If you agree to have your child participate in the research described here but later change your mind, you have the option of withdrawing your child from the project at any moment. Return the signed form promptly to Mrs. Laura Kinnaird. Note that given time and space limitations, not all children will participate in the project. Children participating in the project will be randomly selected. Therefore, signing the Consent To Project Participation form does not guarantee your child's participation in the project.

If you have further questions please feel free to contact the school's principal, Ms. Judy Sparkman, the school's music teacher Mrs. Laura Kinnaird, or me.

Sincerely,

Beatriz E. Aguilar, doctoral candidate
University of North Texas
College of Music
bea0005@unt.edu
Phone: 940-891-4251

Dr. Debbie Rohwer, Main Advisor
University of North Texas
College of Music
Coordinator of the Ph.D. program in Music
Education
Phone: 940-369-7538

RESEARCH CONSENT FORM

Subject Name: _____

Date: _____

Title of Study: The Effect Creative Musical Problem Solving in Team versus Individually on 4th and 5th Grade Students' Compositional Products. (Pilot Test)

Main Investigator: Beatriz E. Aguilar

Co-investigators: Rodrigo Villanueva

Before agreeing to participate in this research study, it is important that you read and understand the following explanation of the proposed procedures. It describes the procedures, benefits, risks, and discomforts of the study. It also describes your right to withdraw from the study at any time. It is important for you to understand that no guarantees or assurances can be made as to the results of the study. **This study has been reviewed and approved by the UNT committee for the Protection of Human Subjects (940-565-3940).**

Purpose of the study and how long it will last:

The purpose of the pilot study is to explore the reliability of the measurement instrument to be used on a research study exploring the effects that training students to solve creative musical problems in teams versus individually could have on their compositional products. The measurement instrument will be tested on eight 4th grade and eight 5th grade students attending Town Center Elementary School. Given time and space limitations, children participating on this pilot study will be randomly selected from those whose parents sign the Consent To Project Participation. The selected students will be taken out of their music class for 15 minutes on May 19th, 2003, or May 20th, 2003.

Description of the study including the procedures to be used:

Students participating in this pilot study will be taken individually into a room, where a two-octave xylophone will be placed. There will also be a clock placed where students can clearly see it, and a tape recorder with a unidirectional microphone. Once in the room, the researcher will ask the student to solve a compositional task. This compositional task does not have a correct or wrong answer. Once the directions have been read and all the questions answered, the researcher will move to another part of the room. The student will have ten minutes to work on the task. Once the student indicates s/he has finished, or the allowed time has passed, the researcher will ask the student to perform the compositional product and the recording will be turned on. After the student has played the song, the recording will be turned off, and the student will be taken back to the music classroom. The audiotapes of each student's final creative problem solving products will be assigned with a code number to ensure anonymity. Once coded, all the audiotapes will be recorded in a single compact disk, which will be analyzed by two independent judges, both of them general music educators with extended professional experience.

RESEARCH CONSENT FORM (Continued)

Description of procedures/elements that may result in discomfort or inconvenience:

The only inconvenience foreseen in this project is the student's absence from his/her music lesson for fifteen minutes. However, it is important to note that the task required in this project is an activity designed to provide students with music creative opportunity appropriate for their age, and suggested by the TAKs. It is also important to note that the study is designed to develop student's creative musical problem solving skills, and that it is expected that all participants will learn some creative problem solving strategies.

Description of the procedures/elements that are associated with foreseeable risks:

There are no procedures or elements involved in this research project that we foresee could constitute a physical or emotional risk for the participants.

Benefits to the subjects or others:

Encouraging the development of critical thinking has become a constant concern for North American educators. In Goals 2000: Educate America Act, it is stressed that the percentage of all students who demonstrate the ability to reason, solve problems, apply knowledge should increase substantially. Since the establishment of these Goals, it has become imperative to favor the development of effective thinking skills, of critical thinking skills in the schools. However, research on critical thinking and problem solving skills is relatively new in music education. Clearly, music educators could highly benefit from research that explores the creative problem solving processes and products as they take place in the music classroom. This pilot test will allow the assessment of the reliability of a particular measurement instrument to be used on a research project exploring the effects that training students to solve creative musical problems in teams versus individually could have on their compositional products. Students participating in this pilot test will have the opportunity to participate in a creative problem solving activity and learn creative problem solving strategies.

Confidentiality of research records:

Children participating in this project will be randomly selected from those whose parents sign the Consent To Project Participation. Once the participants have been selected, they will be assigned a code number. Their identity will not be released to a third party at any time. The audiotapes will be assigned a code number, and the identity of the participant will be known only by the main researcher. The final compositions will be recorded from the audiotapes to a single compact disk, in which the compositions will be identified as "subject 1," "subject 2," and so on. In this way, the identity of the students will be unknown to the judges or any other third party. Once transferred into the compact disk, the audiotapes will be stored by the main researcher for her records. The results of the tests will be used to establish the reliability of the measurement instrument, and will not affect the students' grades or any other academic standing.

CONSENT TO PROJECT PARTICIPATION

Town Center Elementary School

You are making the decision about whether or not have your child participate in this study. Your signature indicates that you have decided to allow your child to participate, that you have read (or have had read to you the information provided in this Consent Form and that you have received a copy of it. You should also be aware that participation in this project is voluntary and no consequences will follow from declining to participate or withdrawing from the study once it has begun, and that the results of the tests will not affect the students' grades or any other academic standing. Finally, you should know that this study has been reviewed and approved by the UNT committee for the Protection of Human Subjects.

Student's Parent's or Guardian's Signature

Date

Investigator's Signature

Date

ASSENT OF CHILD

_____(name of child) has agreed to participate in the research project The Effect of Creative Musical Problem Solving in Team versus Individually on 4th and 5th Grade Students' Compositional Products.

Student's Signature

Date

APPROVED BY THE UNT IRB
FROM 5/19/03 TO 5/18/04
118

Research Consent Form -Page 3 of 3 _____ Participant's initials

APPENDIX F
LETTERS OF APPROVAL

UNIVERSITY of NORTH TEXAS

Office of Research Services

May 19, 2003

Beatriz E. Aguilar
1800 Jason Dr., Apt. 126
Denton, TX 76205

RE: Human Subjects Application No. 03-157

Dear Ms. Aguilar:

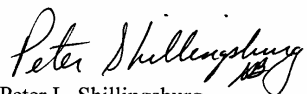
On April 25, 2003, the University of North Texas Institutional Review Board reviewed your project titled "The Effect of Creative Musical Problem Solving in Team Versus Individually on 4th and 5th Grade Students' Compositional Products." The Board agrees that the risks inherent in this research are minimal, and the potential benefits to the subjects outweigh those risks. Your study is hereby approved for the use of human subjects. **Federal policy 45 CFR 46.109(e) stipulates that IRB approval is for one year only. Approval date for this study is May 19, 2003 through May 18, 2004.**

Enclosed is the consent document with stamped IRB approval. Please copy and **use this form only** for your study subjects.

U.S. Department of Health and Human Services regulations require that you submit annual and terminal progress reports to the UNT Institutional Review Board.

Further, the UNT IRB must re-review this project annually and/or prior to any modifications you make in the approved project. Please contact me if you wish to make such changes or need additional information.

Sincerely,



Peter L. Shillingsburg
Chair
Institutional Review Board

PS:sb

P.O. Box 305250 • Denton, Texas 76203-5250 • (940) 565-3940
Fax (940) 565-4277 • TTY (800) RELAY TX • www.unt.edu



Denton Independent School District

Dr. Roger D. Rutherford

Assistant Superintendent

Academic Programs

1307 North Locust

Denton, Texas 76201

(940) 369-0133

Fax (940) 369-4983

Ms. Beatriz E. Aquilar
1800 Jason Dr. #126
Denton, TX 76205

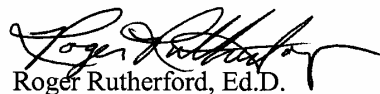
August 26, 2003

Dear Ms. Aquilar:

I am pleased to inform you that your research proposal: "The Effect of Individual vs. Group Problem Solving Training on 4th and 5th Grade Student's Compositional Products" has been approved by Mrs. Teresa Andress (369.2900), Principal at Sam Houston Elementary School and the Instructional Services Division.

Please contact the principal to initiate your research activities.

Sincerely yours,



Roger Rutherford, Ed.D.

Assistant Superintendent

Instructional Services Division

RR/ml

cc: Mrs. Teresa Andress

APPENDIX G

JUDGES' TRAINING MATERIAL AND SCORING FORMS FOR PILOT TEST

MUSICAL CREATIVE PROBLEM SOLVING: DOCTORAL DISSERTATION PROJECT

Dear judges:

Thank you very much for participating in this study. The time and work that you are going to invest on it are of great importance and are highly appreciated. Today's session is intended to unify our understanding about the terminology and procedures that need to be followed during the scoring process, and to solve any problems you might have found the first time you worked with this form. You will receive, along with this Instructions Form, an audiotape containing all the compositions that you will score, and enough Judges' Scoring Forms.

The audiotape you are receiving contains 14 compositions. Before each composition, you will hear the subject number. You will notice that Subjects number 5 and 9 are missing. These two compositions were randomly chosen to be used on today's training session. Before start scoring the composition, write down your name and the subject's number on the space provided in the Judges' Scoring Form. Then, listen to the composition once without making any marks on the Scoring Form. Use this first listening to become familiar with the composition. Then, you can listen to the compositions as many times as you need in order to score them.

For the purposes of this study, the compositions are going to be evaluated in terms of their cohesiveness with two subtests, pattern use with four subtests, and creativity with two subtests. For the first two subtests, tonal and metric cohesiveness are defined as follows:

Tonal Cohesiveness: The degree to which the pitches in a composition are constructed around a tonal center or tonal centers. (1 = very strong tonal cohesiveness, 7 = no tonal cohesiveness)

Metric Cohesiveness: The degree to which the durations in a composition are constructed of regularly occurring accented and unaccented beats. (1 = very strong metric cohesiveness, 7 = no metric cohesiveness)

The next four subtests are related to melodic and rhythmic pattern use. For the purposes of this study, you will indicate the number of times the subjects made use of repeated melodic patterns, developed melodic patterns, repeated rhythmic patterns, and developed rhythmic patterns. You will indicate this by writing on the Judges' Scoring Form one tally for each time you heard a recognizable melodic or rhythmic pattern being repeated or developed. At the end, the total number of tallies for each type of pattern will be summed in order to obtain the total number of occurrences of each type of pattern during the composition. Repeated melodic patterns, developed melodic patterns, repeated rhythmic patterns, and developed rhythmic patterns are defined as follows:

Melodic Pattern: two to seven pitches that form a distinct pitch pattern that is perceived as a unified whole.

Repeated Melodic Pattern: a melodic pattern that is identical to a previous occurring melodic pattern.

Developed Melodic pattern: a melodic pattern that is similar to, yet different from, a previously occurring melodic pattern. In most cases, the rhythm is held the same, the melody changes, and the melodic contour is similar.

Rhythmic Pattern: two to seven durations that form a distinct duration pattern that is perceived as a unified whole.

Judges' Information Letter for Pilot Test (Continues)

Repeated Rhythmic Pattern: a rhythmic pattern that is identical to a previous occurring rhythmic pattern.

Developed Rhythmic Pattern: a rhythmic pattern that is similar to, yet different from, a previously occurring rhythmic pattern. In most cases, the melody is the same or similar and the rhythm changes. (p. 121)

The last two subtests are a fluency and an originality test. Fluency is understood as the ability to generate a number of clearly defined musical ideas. In the case of musical fluency, a one indicates that the composition could not be reproduced, and a seven indicates that someone else could reproduce the composition without trouble. You can mark any number of this scale that better expresses your assessment of the composition's level of fluency. Musical originality is defined as the ability to create a composition that contains elements of unique expression. In the case of originality, a one indicates that the composition had no uniqueness at all, and a seven indicated that the composition had marked uniqueness.

If you have any questions, we can solve them today. If you have any question once that you have started working on the scoring of the compositions, feel free to contact me. Make a note of any problem or concern you might find during this task. Your comments and suggestions are highly valued and appreciated. I hope you find this project as interesting as I do. Thank you very much!

Sincerely,

Beatriz E. Aguilar

Phone: (940) 300-8782

E-mail: betina68@yahoo.com

MUSICAL CREATIVE PROBLEM SOLVING PRODUCT

Judge's Scoring Form

Judge: _____

Subject Code Number: _____

Cohesiveness

Tonal Cohesiveness	1	2	3	4	5	6	7
Metric Cohesiveness	1	2	3	4	5	6	7

Use of patterns

Type of pattern	Tallies for the number of occurrences	Total
Repeated Melodic Patterns		
Developed Melodic Patterns		
Repeated Rhythmic Patterns		
Developed Rhythmic Patterns		

Creativity

Fluency

Rate, as a whole, the quality of delineation (how well could someone else reproduce the composition)

1	2	3	4	5	6	7
It can not be reproduced						Very clear with no trouble at all

Originality

Rate, as a whole, the composition in regard to its originality of expression

1	2	3	4	5	6	7
No Originality						Marked Originality

Figure 6. Transcription of Composition Example A for Judges Training Session

MUSICAL CREATIVE PROBLEM SOLVING: DOCTORAL DISSERTATION PROJECT

Transcription of Examples for Training Session

M = Melodic pattern

R = Rhythmic pattern

Subject # 5

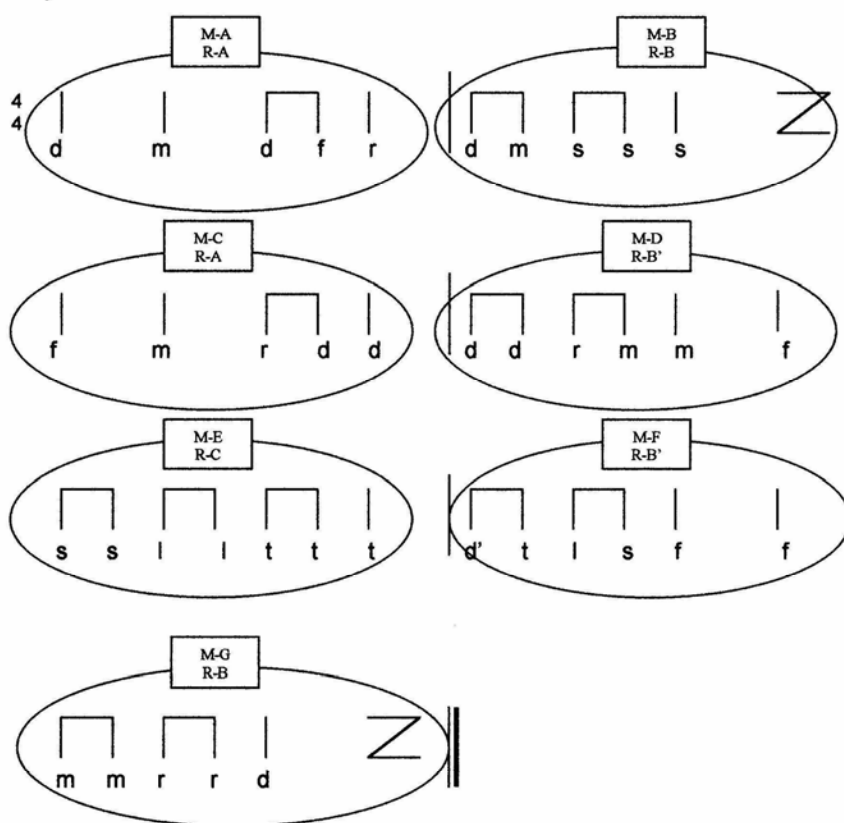


Figure 7. Transcription of Composition Example B for Judges Training Session

[illegible]

APPENDIX H
INTERJUDGE RELIABILITY

Table 18

Overall Interjudge Reliability, Pilot Test

Judges' combinations	Overall Interjudge Reliability
Judges A + B + C	0.79
Judges A + C	0.84
Judges A + B	0.78
Judges B + C	0.70

Table 19

Interjudge Reliability for Pattern Use, Pilot Test

Judges' combinations	<i>Pattern Use</i>				Overall
	Repeated	Developed	Repeated	Developed	
	Melodic Pattern	Melodic Pattern	Rhythmic Pattern	Rhythmic Pattern	
Judges A + B + C	0.86	0.64	0.71	0.86	0.77
Judges A + C	0.93	0.78	0.64	0.79	0.79
Judges A + B	0.78	0.64	0.85	0.78	0.76
Judges B + C	0.85	0.57	0.64	1.00	0.76

Note: Scores \pm from each other were considered agreements.

Table 20

Interjudge Reliability for Cohesiveness, Pilot Test

<i>Judges'</i> <i>combinations</i>	Cohesiveness		
	Tonal	Metric	
	Cohesiveness	Cohesiveness	Overall
Judges A + B + C	0.90	0.79	0.85
Judges A + C	0.93	0.86	0.89
Judges A + B	0.86	0.86	0.86
Judges B + C	0.93	0.64	0.79

Note: Scores \pm from each other were considered agreements.

Table 21

Interjudge Reliability for Creativity, Pilot Test

<i>Judges' combinations</i>	Creativity		
	Fluency	Originality	Overall
Judges A + B + C	0.93	0.88	0.90
Judges A + C	0.93	0.86	0.89
Judges A + B	0.92	0.85	0.89
Judges B + C	0.93	0.64	0.79

Note: Scores \pm from each other were considered agreements.

Table 22

Interjudge Reliability for Pattern Use

Subtests	<i>Judges A + B + C</i>		
	<i>Pattern Use</i>		
	Pilot Test	Pretest	Posttest
Repeated Melodic Pattern	0.86	0.72	0.79
Developed Melodic Pattern	0.64	0.66	0.64
Repeated Rhythmic Pattern	0.71	0.45	0.5
Developed Rhythmic Pattern	0.86	0.81	0.79
Overall Pattern Use	0.77	0.66	0.68

Table 23

Interjudge Reliability for Pattern Use by Each Pair of Judges

Subtests	<i>Pattern Use</i>								
	<i>Judges A + B</i>			<i>Judges B + C</i>			<i>Judges A + C</i>		
	Pilot	Pretest	Posttest	Pilot	Pretest	Posttest	Pilot	Pretest	Posttest
	Test			Test			Test		
Repeated									
Melodic	0.78	0.76	0.76	0.85	0.7	0.81	0.93	0.7	0.85
Pattern									
Developed									
Melodic	0.64	0.7	0.6	0.57	0.73	0.6	0.78	0.7	0.71
Pattern									
Repeated									
Rhythmic	0.85	0.46	0.6	0.64	0.57	0.48	0.64	0.46	0.4
Pattern									
Developed									
Rhythmic	0.78	0.79	0.89	1.00	0.84	0.67	0.79	0.79	0.82
Pattern									

Table 24

Interjudge Reliability for Cohesiveness

Subtests	<i>Judges A + B + C</i>		
	<i>Cohesiveness</i>		
	Pilot Test	Pretest	Posttest
Tonal Cohesiveness	0.90	0.57	0.55
Metric Cohesiveness	0.79	0.56	0.48
Overall Cohesiveness	0.85	0.57	0.51

Table 25

Interjudge Reliability for Cohesiveness by Each Pair of Judges

Subtests	<i>Cohesiveness</i>								
	<i>Judges A + B</i>			<i>Judges B + C</i>			<i>Judges A + C</i>		
	Pilot	Pretest	Posttest	Pilot	Pretest	Posttest	Pilot	Pretest	Posttest
	Test			Test			Test		
Tonal Cohesiveness	0.86	0.48	0.35	0.93	0.56	0.53	0.93	0.7	0.75
Metric Cohesiveness	0.86	0.48	0.31	0.64	0.5	0.4	0.86	0.71	0.75

Table 26

Interjudge Reliability for Creativity

Subtests	<i>Judges A + B + C</i>		
	<i>Creativity</i>		
	Pilot Test	Pretest	Posttest
Fluency	0.93	0.55	0.51
Originality	0.88	0.46	0.48
Overall Creativity	0.90	0.5	0.5

Table 27

Interjudge Reliability for Creativity by Each Pair of Judges

Judges' combinations	<i>Creativity</i>								
	<i>Judges A + B</i>			<i>Judges B + C</i>			<i>Judges A + C</i>		
	Pilot	Pretest	Posttest	Pilot	Pretest	Posttest	Pilot	Pretest	Posttest
	Test			Test			Test		
Fluency	0.93	0.6	0.45	0.93	0.45	0.34	0.93	0.6	0.76
Originality	0.86	0.43	0.29	0.64	0.7	0.7	0.86	0.43	0.6

REFERENCES

- Amabile, T. M. (1982). Social psychology of creativity: A consensual assessment technique. *Journal of Personality and Social Psychology*, 43, 997-1013.
- Amabile, T.M. (1983). *The social psychology of creativity*. New York: Springer-Verlag.
- Amabile, T.M. (1996). *Creativity in context*. Boulder, CO: Westview Press.
- Azmitia, M. (1988). Peer interaction and problem solving: When are two heads better than one? *Child Development*, 59, 87-96.
- Bangs, R. L. (1992). An application of Amabile's model of creativity to music instruction: A comparison of motivational stages. *Dissertation Abstracts International*, 53 (12), 4243A. (UMI No. 9239649)
- Bloom, B. S. (Ed.). (1956). *Taxonomy of educational objectives: The classification of educational goals*. New York: David McKay.
- Boardman, E. (Ed.). (1989). *Dimensions of musical thinking*. Reston, VA: Music Educators National Conference.
- Boyle, D. J., & Radocy, R. E. (1987). *Measurement and evaluation of musical experiences*. New York: Schirmer Books.
- Byrne, C., Macdonald, R., & Carlton, L. (2003). Assessing creativity in musical compositions: flow as an assessment tool. *British Journal of Music Education*, 23, 277-290.
- Brinkman, D. (1999). Problem finding, creativity style and musical compositions of high school students. *Journal of Creative Behavior*, 33, 62-68.
- Bruner, J. S. (1960). *The process of education*. New York: Random House.

- Burnard, P. (2000). Making a piece you don't play and forget: children composing and the role of context. *Australian Journal of Music Education*, 1, 30-39.
- Campbell, D. T. & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Boston: Houghton Mifflin.
- Campbell, P. S. (1998). *Songs in their heads: Music and its meaning in children's lives*. New York: Oxford University Press.
- Chiara, L., Schuster, J., Bell, J. K., & Wolery, M. (1995). Small-group massed-trial and individually-distributed-trial instruction with preschoolers. *Journal of Early Intervention*, 19, 203-217.
- Christensen, C. B. (1992). Music composition, invented notation and reflection: Tools for music learning and assessment. *Dissertation Abstracts International*, 53 (06), 1834A. (UMI No. 9231370)
- Claire, L. (1993). The social psychology of creativity: The importance of peer social processes for students' academic and artistic activity in classroom context. *Bulletin of the Council of Research in Music Education*, 119, 21-28.
- Colwell, R. (1969). *Music Achievement Test*. Chicago, Ill: Follet Educational Report.
- Cope, D. E. & Murphy, A. J. (1981). The value of strategies in problem solving. *The Journal of Psychology*, 107, 11-16.
- Csikszentmihalyi, M. (1992). *Flow: The psychology of happiness*. London: Random House Ltd.
- Dagnault, L. (1997). Children's creative musical thinking within the context of a computer-supported improvisational approach to composition. *Dissertation Abstracts International*, 57 (11), 4681A. (UMI No. 9714572)

- DeLorenzo, L. C. (1987). An exploratory field study of sixth grade students' creative music problem solving processes in the general music class. *Dissertation Abstracts International*, 48 (07), 1689A. (UMI No. 8721099)
- DeLorenzo, L. C. (1989). A field study of sixth-grade students' creative music problem solving processes. *Journal of Research in Music Education*, 37, 188-200.
- DeTurk, M. (1989). Critical and creative musical thinking. In Boardman, E. (Ed.). *Dimensions of Musical Thinking*. Reston, VA: Music Educators National Conference.
- Dewey, J. (1910). *How we think*. Buffalo, NY: Prometheus Books.
- Duran, R., & Gauvain, M. (1993). *The role of age versus expertise in peer collaboration during joint planning*. *Journal of Experimental Child Psychology*, 55, 227-242.
- Eisman, L. W. (1975). The formulation and testing of a problem-solving approach to the development of perceptive listening skills in selected college music appreciation class. *Dissertation Abstracts International*, 36 (08), 4837A. (UMI No. 7601774)
- Elliot, D. (1995). *Music matters: A new philosophy of music education*. New York: Oxford University Press.
- Emmons, S. E. (1998). Analysis of music creativity in middle school students through composition using computer assisted instruction: A multiple case study. *Dissertation Abstracts International*, 59 (02), 441A. (UMI No. 9825697).
- Garton, A., & Pratt, C. (2001). Peer assistance in children's problem solving. *British Journal of Developmental Psychology*, 19, 307-318.

- Getzels, J. W. & Csikszentmihalyi, M. (1975). From problem solving to problem finding. In I. A. Taylor & J. W. Getzels (Eds.), *Perspectives in creativity* (pp.90-116). Chicago: Aldine.
- Guilford, J. P. (1956). The structure of intellect. *Psychological Bulletin*, 53, 267-293.
- Gordon, E. (1982). A longitudinal predictive validity study of the intermediate measures of music audiation. *Bulletin of the Council of Research in Music Education*, 78, 1-23.
- Henry, W. (1995). The effects of pattern instruction, repeated composing opportunities, and musical aptitude on the compositional processes and products of fourth-grade students. *Dissertation Abstracts International*, 56 (07), 2560A. (UMI No. 9537223)
- Henson, R. K. (1998). *ANCOVA with intact groups: Don't do it!* Paper presented at the annual meeting of the Mid-South Educational Research Association, New Orleans. (ERIC Document Reproduction Service No. ED 426086).
- Hickey, M. (1995). Qualitative and quantitative relationship between children's creative musical thinking processes and products. *Dissertation Abstracts International*, 57 (01), 145A. (UMI No. 9614754)
- Hickey, M. (1996). *Consensual assessment of children's musical compositions*. Paper presented at the New York State School Music Association Convention, NY.
- Hickey, M. (2001). An application of Amabile's Consensual Assessment Technique for rating the creativity of children's musical compositions. *Journal of Research in Music Education*, 49, 234-244.

- Huck, S. W. (2000). *Reading statistics and research*. NY: Longman.
- Kaschub, M. E. (1999). Sixth grade student's descriptions of their individual and collaborative music composition processes and products initiated from prompted and unprompted task structures. *Dissertation Abstracts International*, 60 (06), 1955A. (UMI No. 9932187)
- Kay, S. (1994). A method for investigating the creative thought process. In M. A. Runco (Ed.), *Problem finding, problem solving, and creativity* (pp. 116-129). Norwood, NJ: Ablex.
- Kratus, J. (1989). A time analysis of the compositional processes used by children ages 7 to 11. *Journal of Research in Music Education*, 37, 5-20.
- Kratus, J. (1994). Relationships among children's music audiation and their compositional processes and products. *Journal of Research in Music Education*, 42, 115-130.
- Kratus, J. (2001). Effect of available tonality and pitch options on children's compositional processes and products. *Journal of Research in Music Education*, 49, 294-306.
- Larey, T. S. & Paulus, P. B. (1999). Group preferences and convergent tendencies in small groups: A content analysis of group brainstorming performance. *Creativity Research Journal*, 12, 175-184.
- Laycock, R. P. (1992). The relationship of musical experience, music aptitude, self-concept, age, and academic achievement to the musical problem solving abilities of high school students. *Dissertation Abstracts International*, 53 (8), 2728A. (UMI No. 9236315)

- MacDonald, R. A. R. & Miell, D. (2000). Creativity and music education: The impact of social variables. *International Journal of Music Education*, 36, 58-68.
- Marzano, R. J., Brandt, R. S., Hughes, C. S., Jones, B. F., Presseisen, B. Z., Rankin, S. C., et al. (1988). *Dimensions of thinking: A framework for curriculum and instruction*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Marzano, R. J., Pickering, D. J., Arredondo, D. E., Blackburn, G. J., Brandt, R. S., Moffett, C. A., et al. (1997). *Dimensions of learning: Teacher's manual*. Alexandria, VA: Association for Supervision and Curriculum Development.
- McCarthy, J. F. (1979). Individualized Instruction, student achievement, and dropout in an urban elementary instrumental music program. *Journal of Research in Music Education*, 28, 59-69.
- McCoy, P. M. (1999). Effects of variable task structuring and guided self-reflection on compositional quality, self-assessments, and attitudes of novice student composers. *Dissertation Abstracts International*, 60 (12), 4359A. (UMI No. 9953339)
- Mead, G. H. (1970). *Mind, self, and society: From the standpoint of a social behaviorist*. Chicago: The University of Chicago Press.
- Mumford, M. D, Feldman, J. M., Hein, M. B, & Nagao, D. J. (2001). Tradeoffs between ideas and structure: Individual versus group performance in creative problem solving. *Journal of Creative Behavior*, 35, 1-23.
- Mumford, M. D., & Connelly, M. S. (1994). Creativity and problem solving: Cognition, adaptability, and wisdom. *Roper Review*, 16, 241-247. (Electronic text)

- Music Educators National Conference. (1994). National Standards for Arts Education. Reston, VA: Music Educators National Conference.
- Peel, S. (1998). An innovative problem-solving assessment for groups of first-year medical undergraduates - Think Tanks. *Medical Education* 32, 35-39.
- Perkins, D. N. (1984). Creativity by design. *Educational Leadership*, 43, 18-25.
- Piaget, J. (1967). *Six psychological studies*. New York: Random House.
- Pilsbury, C. & Alston, P. (1996). Too fine a net to catch the fish? An investigation of the assessment of composition in GSCE music. *British Journal of Music Education*, 13, 243-258.
- Richardson, C. P. & Whitaker, N. L. (1992). *Critical thinking and music education*. In Richard Colwell (Ed.). *Handbook of research on music teaching and learning*. New York: Schirmer Books.
- Robertson, S. I. (2001). *Problem solving*. Philadelphia, PA: Taylor and Francis.
- Sloboda, J. A. (1985). *The musical mind: The cognitive psychology of Music*. Oxford: Claredon Press.
- Tomic, W. (1995). Training in inductive reasoning and problem solving. *Contemporary Educational Psychology*, 20, 483-490.
- Treffinger, D. J., Isaksen, S. G., & Dorval, B. (1994). Creative problem solving: An overview. In M. A. Runco (Ed.), *Problem finding, problem solving, and creativity* (pp. 223-236). Norwood, NJ: Ablex.
- U.S. Congress. (1989). *Linking for learning: A new course for education, OTA-SET-430*. Washington: US Government Printing Office.

- U.S. Congress. (1993). *Goals 2000: Educate America Act*. Washington: US Government Printing Office.
- Vold, J. N. (1986). A study of musical problem solving behavior in kindergarten children and a comparison with other aspects of creative behavior. *Dissertation Abstracts International*, 47 (09), 3352A. (UMI No. 8628183)
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Webster, P. (1977). A factor of intellect approach to creative thinking in music. *Dissertation Abstracts International*, 38 (06), 3136A. (UMI No. 7726619)
- Webster, P. (1987). Refinement of a measure of creative thinking in music. In Madsen, C. and Prickett, C. (Eds.), *Applications of research in music behavior* (pp. 257-271). Tuscaloosa: University of Alabama Press.
- Webster, P. R. & Hickey, M. M. (1995). Rating scales and their use in assessing children's music compositions. *The Quarterly Journal of Music Teaching and Learning*, 6, 28-44.
- Whitaker, N. L. (1989). *Reflective thinking as exemplified in musical decision-making*. *Dissertation Abstracts International*, 50 (11), 3510A. (UMI No. 9011074)
- Whitaker, N. L. (1996). A theoretical model of the musical problem solving and decision making of performers, arrangers, conductors, and composers. *Bulletin of the Council of Research in Music Education*, 128, 1-14.
- Wiggins, J. H. (1990). *Composition in the classroom: A tool for teaching*. Reston, VA: Music Educators National Conference.

Wiggins, J. H. (1999). Teacher control and creativity. *Music Educators Journal*, 85, 30-35.

Wiggins, J. H. (2000). The nature of shared musical understandings and its role in empowering independent musical thinking. *Bulletin of the Council of Research in Music Education*, 143, 65-90.