EFFECTS OF VISUAL PRESENTATION ON AURAL MEMORY FOR MELODIES

A Dissertation
Submitted to
the Temple University Graduate Board

In Partial Fulfillment
of the Requirements for the Degree
DOCTOR OF PHILOSOPHY

By
Nathan O. Buonviri
May 2010

Examining Committee Members:

Deborah Sheldon, Advisory Chair, Music Education
Alison Reynolds, Music Education
Michael Klein, Music Studies, Theory
Richard Brodhead, Music Studies, Composition

ABSTRACT

The purpose of this study was to determine how pitch and rhythm aspects of melodic memory are affected by aural distractions when melodic stimuli are presented both visually and aurally, as compared to aurally only. The rationale for this research is centered on the need for improved melodic memory skills of students taking melodic dictation, and the possibility that temporary visual imagery storage of target melodies might enhance those skills.

The participants in this study were undergraduate and graduate music majors (*n*=41) at a large northeastern university. All participants had successfully completed the first two semesters of college-level music theory, and none had perfect pitch. Participants progressed through two self-contained experimental tests at the computer. Identical target melodies were presented: 1) aurally only on one test; and 2) aurally, with visual presentation of the matching notation, on the other test. After the target melody, a distraction melody sounded, during which time participants were to maintain the original target melody in memory. Participants then chose which of two aural options matched the original target, with a third choice of "neither." The incorrect answer choice in each item contained either a pitch or rhythm discrepancy.

The 2x2 factorial design of this experiment was based on independent variables of test presentation format and answer discrepancy type. The dependent variable was experimental test scores. Each participant took both parts of both tests, yielding 164 total observations. Additional data were collected for exploratory analysis: the order in which each participant took the tests, the major instrument of each participant, and the educational status of each participant (undergraduate or graduate).

Results of a 2x2 ANOVA revealed no significant differences in test scores, based on either test format or answer discrepancy type, and no interaction between the factors. The exploratory analyses revealed no significant differences in test scores, based on test order, major instrument, or student status.

Results suggest that visual reinforcement of melodies does not affect aural memory for those melodies, in terms of either pitch or rhythm. Suggestions for further research include an aural-visual melodic memory test paired with a learning modalities survey, a longitudinal study of visual imagery applied to aural skills study, and a detailed survey of strategies used by successful and unsuccessful dictation students.

ACKNOWLEDGEMENTS

With great love and gratitude, I acknowledge the outstanding contribution my wife Maria and sons Carlo, Paolo, and Matteo have made to this dissertation. Your patience, strength, and support have inspired me beyond words.

Heartfelt thanks to my teachers, beginning with the earliest, Mom and Dad.

I extend deep appreciation to Dr. Deborah Sheldon, my dissertation advisor, for being an outstanding model of professionalism in music, education, and research.

Many thanks to the members of my dissertation advisory committee: Dr. Alison Reynolds and Dr. Michael Klein. Your expertise, helpfulness, and positive attitude have sustained me throughout this process. Thanks also to Professor Richard Brodhead for document review, to Dr. Cynthia Folio for review of experimental test melodies, and to Andrew Buonviri for technical support.

Finally, my sincere appreciation to the participants in the experimental work of this study, and to those who helped me establish contact with them. This research was made possible by you.

This dissertation is dedicated to my wife Maria; these pages represent your hard work every bit as much as mine.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Music theory coursework is an integral part of university music curricula. The aural skills component of these courses typically comprises a variety of related tasks, including sight singing, error detection, and dictation. Thorough study of any one of these tasks may shed light on requisite skills shared by the others, because all three involve visual and aural information in combination.

Dictation requires students to notate what they have heard. Common musical materials for dictation are rhythm sequences, harmonic progressions, and melodies. Rhythmic dictation is simplified by the absence of pitch changes. Harmonic dictation is often simplified by the absence of complicated rhythms. Melodic dictation, however, typically combines active rhythms and changing pitch contour.

Melodic dictation can be a daunting task for both high school and college music students, yet formation of the requisite skills for successful dictation can begin early in a child's musical development. Melodic dictation can be defined, after all, as a simple combination of correct perception of aural information and accurate notation of that information. A barrier to success for those taking dictation may lie in the transference of aural stimuli into visual output.

Students often perceive a number of challenges to dictation success, including lack of time, lack of repetitions of the melody, and lack of a silent environment in which to process percepts. They must work quickly and accurately, and develop strategies to overcome the interference of aural distractions.

In light of these considerations, the transference process is complicated by two human limitations:

1. the inability to hold an entire melody in aural memory (Miller, 1956), and
2. the inability to maintain even a section of the melody in aural memory in the midst of ensuing melodic sections.

It seems that if the effect of these two limitations could be decreased, the problems of lack of time and lack of repetitions would automatically decrease as well.

Miller (1956) suggested that the human mind is capable of holding seven (plus or minus two) bits of information in memory. He also suggested that a small group of bits can be encoded into a larger chunk, for memory storage. Students of dictation may be able to use this "chunking" while processing melodic information. However, they must combat aural distractions while carrying out this task. How can they best secure the information cognitively, for quick and accurate retrieval?

Segal and Fusella (1970) found that aural images are more likely to be confused by external aural stimuli, and visual images by visual stimuli, than by cross-modal combinations. In light of Miller's suggestions, a visual "snapshot" of a specific chunk of melody held in visual imagery may be an effective strategy for overcoming aural distractions when taking melodic dictation. While it is difficult to test whether subjects are visualizing a melody, and to what extent (Burton, 2003), it is feasible and worthwhile to explore effects of overt visual presentation of melodies on aural memory for those melodies.

Why is it necessary to develop skills in melodic dictation? Even a cursory review of objectives and requirements for undergraduate degree programs at most universities would reveal a focus on development of aural acuity. Dictation represents a simple, yet crucial, musical skill, especially for musicians of the Western world. Musicians are quite likely to work in musical situations that employ both aural and visual skills, separately and in combination. Through dictation, musicians demonstrate that they have accurately heard and processed a given melody, in context of the standardized system of notation used in Western music.

Although a clean notated copy of dictation is not entirely essential, it is the equivalent of a typed or hand-written paper in any other subject, proving that the student did indeed employ the necessary cognitive faculties. It is the culminating step, reflecting the successful acquisition and coordination of internal musical skills. Edwin Gordon (2007) explains: "…reading and writing are considered important in such a [professional] curriculum if for no other reason than that they serve as objective and efficient means for measuring and evaluating students' potential to audiate" (pp. 43-44).

Fluency in connections among the visual, aural, and kinesthetic aspects of music are an important underpinning of musical performance as well. Cumming (2000) suggests that the important sense of "gesture" in music-making is dependent upon three intertwined elements: the performance of a pattern or patterns, the notation of those patterns, and the idiomatic or inherently stylistic suggestiveness of the patterns. A central goal in music education is helping students draw tangible connections across musical skills. Melodic dictation affords students the opportunity to demonstrate achievement of that goal, and improved performance on dictation tasks represents improved integration of the cognitive skills required to carry out those tasks. The purpose of this study was to determine how pitch and rhythm aspects of melodic memory are affected by aural distractions when melodic stimuli are presented both visually and aurally, as compared to aurally only.

CHAPTER 2

REVIEW OF RELATED LITERATURE

Goodman (1978) described relationships between sight and sound in artistic endeavors:

> For the forms and feelings of music are by no means all confined to sound; many patterns and emotions, shapes, contrasts, rhymes and rhythms are common to the auditory and the visual and often to the tactual and the kinesthetic as well….In these days of experimentation with the combination of media in the performing arts, nothing is clearer than that music affects seeing, that pictures affect hearing… (p.160)

Goodman was drawing important connections across sensory experiences, creating a fuller picture of artistic understandings and skills. As developing musicians tackle musical tasks incorporating sight and sound, strengthening these connections is important, as evidenced by several authors' descriptions.

Telesco (1991) described the development of aural analytical skills as "the wedding of the aural experience to abstract musical constructs" (p. 179). This definition may provide an important clue in solving the problem of retaining aural stimuli for transfer to visual output during dictation. Even though students perceive stimuli accurately, they may lack the well-oiled mental music theory concepts needed to finish the transaction. Covington (1992) suggested that students must complete three steps in acquiring these skills: experience musical stimuli, construct a mental image of the function of the stimuli, and represent the stimuli and their function symbolically. The first step is common in virtually all music learning situations. The second step relates to mental tools that students already should have developed going into a dictation situation. The third step describes dictation itself.

The studies of aural perception and imagery presented in this chapter concern listening to sound from external sources and from within the mind. The visual perception and imagery studies are drawn from other fields such as reading comprehension. Studies involving cognitive connections between the two domains offer insights into the channels of the mind that may be used in processing auditory and visual percepts, both separately and concurrently. These reports begin to lay the foundation for concepts to be explored in the current study. The skilled imagery studies give the reader a glimpse into mental structures used by blindfold chess players, expert abacus operators, and successful mental mathematicians. Finally, the melodic perception and musical applications sections outline studies that connect uses of imagery to specific musical tasks in aural skills and sight-singing.

Aural Perception and Imagery

Aural perception is defined here as mental registry of aural stimuli that are physically present. Aural imagery refers to "inner hearing" of sound that is not physically present. Authors on musical learning have coined more specific terms to connect these concepts to specifically musical tasks. Gary Karpinski (2000) described "auralization" as the "process of hearing music mentally in the absence of the physical sound" (p. 49). While auralization refers simply to musical inner hearing, other notable music educators have tied both "outer" and "inner" hearing to grasping musical meaning. David Elliott (1995) described "audition," in which the listening mind is engaged in both listening *to* music and listening *for* music (i.e., making musical predictions. Edwin Gordon emphasized the importance of comprehending musical function and context in the development of skill in "audiation," both for sounds that are, and are not, physically present. He states: "Audiation is to music what thought is to language" (2007a, p. 2).

For the purpose of this discussion, "aural perception" and "aural imagery" were chosen to afford the greatest breadth and flexibility in understanding the cognitive tasks specific to the dictation process. Without the basic skills described by these terms, none of the more complex skills just described could be approached. Studies of aural perception and imagery help us understand how the mind initially registers, and subsequently processes, melodic information during dictation.

Margulis (2005) suggested that aural attention to melodic events be explored both forward and backward in time. She explained a detailed model of melodic expectancy, highlighting the effects of confirmed and unconfirmed aural expectancies on listeners' decisions about musical meaning (see also Meyer, 1968). Margulis described four key attributes of melodic events that contribute to listeners' overall melodic expectation satisfaction ratings: stability, proximity, direction, and mobility.

Margulis related the relationships of these attributes to different categories of melodic tension, the most pertinent of which are expectancy-tension and denial-tension. Expectancy-tension refers to forward-looking melodic tension suggested by current melodic events; it causes listeners to focus sharper attention on the events to come. Denial-tension refers to backward-looking melodic tension suggested by current events that deny the expectations of the listener; it causes listeners to focus sharper attention on the events that preceded thwarted expectations. These two tension types could be used to study and improve melodic memory, because salient moments within the perception of the melody can serve as a foundation for structural understanding and coding of the information.

Reading music prompts the dual expectations just described as well. For instance, in visually scanning a melody, several F-sharps or B-flats within a group of running eighth notes may suggest a move away from C major. We expect to see the melody outline G major or F major, respectively, in the ensuing melodic events. We have entertained expectancy-tension by looking forward, and we scan forward with heightened awareness. Conversely, if we see an F-sharp under a fermata at a major cadence, we recognize that our normal expectation of a member of the C major or G major triads has been undermined. We are experiencing denial-tension and we scan backward with heightened awareness. These experiences of tension can occur based solely on visual understanding of notation organization, or on a combination of visual perception of notation and concurrent connection to the sound it represents.

Both aurally and visually, melodies can be evaluated forward and backward, assuming that the music (sound or notation) is present or is remembered accurately. Consideration of Margulis' expectation possibilities provides a foundation on which to build concepts of hierarchical melodic attention, factors in melodic tension and memory, and relationships between melodic events in normal and reverse order. If these concepts are learned and practiced in parallel form aurally and visually, they may be internalized more deeply, and prove more useful in focused listening situations.

The ability to form useful, efficient visual snapshots of melodic stimuli likely depends on fluency in basic music theory concepts and notation. Long (1977) examined effects of subjects' background in music perception and understanding of melodic contour, tonality, and length, on the ability to remember pitches within melodies. She formed three groups of undergraduate and graduate students based on music perception ability levels. Long played recorded melodic examples of either tonal or atonal nature, of either v-shaped or w-shaped contour, and of varying lengths (7, 11, and 15 pitches), followed by a two-second pause and a 1-second probe tone. Subjects were to decide whether the tone had been part of the melody and to what extent they were sure.

Long found that only the main effects of tonality/atonality and perception ability were statistically significant. While these two effects may initially seem obvious, they highlight an important consideration for future research. Long explains:

> The results of the perception ability factor support the conjecture that memory is dependent upon learned systems….It has been suggested that Western tonality, a learned system, aids memory. (p. 280)

Therefore, access to a pliable imaged structure of basic music theory constructs may be helpful in processing melodic dictation percepts.

The usefulness of such a manipulable mental construct is suggested by the work of Povel and Jansen (2002). Subjects' "goodness" ratings for melodic stimuli were found to be highest for melodic segments that outlined common chord progressions. For segments containing chromatics, subjects scored melodies higher when the chromatic was anchored to a subsequent resolution tone or surrounded by a chromatic run. Segments containing a chromatic connected to neither an anchor nor a run were rated lowest. To make these evaluations, subjects would have to interpret the melodic stimulus in context of harmonic and structural knowledge and experience. It appears that when subjects were able to place the stimuli into existing familiar schema, they rated the melodies higher than when they could not contextualize the information. The formation of efficient skill with contextual schema can occur in various domains (e.g., aural, visual, tactile), and be used to organize, memorize, and manipulate melodic information during dictation.

Melodic memory may indeed be directly aided by experience with music theory concepts. Wilson and Saling (2008) examined relationships of the right and left temporal lobes to accuracy of melodic memory. The researchers studied a group of patients with temporal lobe epilepsy and a group of musicians of varying ability (from “no training” to “moderate”). The musician group was used to establish “normal” results, against which to compare the results of the patients. Findings from the musician group are most pertinent to the current study.

Subjects listened to target pairs of three-note melodic segments, each followed by three more pairs from which they were to choose which matched the target. Subjects completed three trials each of tonal examples and non-tonal examples. The researchers found a significant interaction between stimulus type and trial and a significant main effect for group across the three learning trials of the non-tonal (difficult) pairs.

They suggested that the interaction of stimulus type and trial resulted from subjects improving over all three non-tonal trials, while reaching a plateau of learning at the second trial of tonal examples. The significant main effect for group across non-tonal pairs was attributed to experienced musicians' greater perceptual acuity and increased exposure to non-tonal musical materials, as compared to "no training" musicians. This suggests that general musical and, perhaps, theoretical, experience contributes to the ability to accurately perceive and remember diverse musical ideas.

The initial stages of aural perception may be affected by musical training as well. Kizkin, Karlidag, Ozcan, and Ozisik (2006) measured musicians' and non-musicians' brainwave responses to pairs of audible clicks. They were specifically looking for differences between the two groups in sensory "gating"—suppressing certain signals in the brain in order to focus on other signals—in relation to P50 and N100 waves that occur during the very early stages of aural perception. They found a significant difference in P50 suppression, with musicians suppressing less than non-musicians. This means that musicians are less inclined to ignore auditory signals than non-musicians. No significant difference was found in N100 suppression, the stage just milliseconds after P50. It was concluded that auditory perception processing happens in multiple stages, and that the N100 wave is sheltered or protected by the P50.

While musicians may show more attention to auditory signals, humans are generally incapable of attending to the entirety of a melody during dictation (Miller, 1956). Heightened awareness of real-time auditory stimuli must be balanced with memorization of stimuli immediately preceeding. In other words, heightened aural attention can actually work against itself in the presence of too much information processed simultaneously (Segal & Fusella, 1970).

Perceiving and organizing melodic information is complex, involving, at the very least, pitches and rhythms. Boisen (1981) presented complete and incomplete rhythms to 7th, 9th, and 11th grade students under three circumstances: on a single pitch; in conjunction with a matched tonally complete or tonally incomplete pitch sequence; and in conjunction with an unmatched tonally complete or tonally incomplete pitch sequence. The labels "complete" and "incomplete" are rather subjective, but Boisen distinguished between the two based on whether each stimulus sounded like a finished or unfinished musical idea. He tested whether the musical (tonal) context established by the pitch sequence in an example would make a difference in students' assessment of the rhythm sequence. He formulated two hypotheses:

> 1. Students display greater accuracy in perceiving rhythmic completeness and incompleteness in melodies whose pitch sequences match the completeness or incompleteness of their rhythmic unit (matching items) than they do in melodies whose rhythmic units occur on only one pitch (single-pitch items).
>
> 2. Students display less accuracy in perceiving rhythmic completeness or incompleteness in melodies whose pitch sequences do not match the completeness or incompleteness of their rhythmic units (non-matching items) than they do in single-pitch melodies. (p. 168)

Boisen found that, while all three differences in condition were statistically significant, only the differences between unmatched sequences and each of the other conditions were practically significant. He concluded that only unmatched pitch sequences made a practical (lower) difference in test scores, and that tonal context does affect rhythmic perception. Therefore, the importance of studying and imagining music as a whole, such that component parts might reinforce each other, might be considered in systematic investigation. Students taking dictation might benefit from mentally representing pitch and rhythmic aspects of a passage simultaneously. If so, processing limitations may require mental division of the melody into sections for short-term memory and analysis.

Memorization of short sections of a melody might be aided by understanding the relationship of each section to the whole. The ability to aurally imagine the component sections of the melody helps in understanding the function of each, and in manipulating those parts in meaningful ways (Gordon, 2007). This approach yields dual results: students improve melodic memory skills, and incorporate those skills into larger musical contexts.

Hubbard and Stoeckig (1988) investigated how auditory imagery functions in processing single notes as compared to chords, in both close and distant harmonic relationships. They hypothesized that chords would require longer processing times than single notes, and that distant harmonic relationships would require longer processing times than close harmonic relationships.

Psychology students, representing a wide range of musical training, listened to either a note or chord cue and formed an aural image of a note or chord a whole step above the cue. When the image was formed they pressed a button which sounded another note or chord. The subjects were to decide if the second sound matched their image.

Unfortunately, several problems undermine the usefulness of the study. Because the students ranged in musical training from “none” to “twenty-one years,” the results of this experiment are not very useful, and are not reported here. That all subjects were told what “tones,” “chords,” and “whole steps” are does not assume that they were able to imagine vertical sound and compare it to aural percepts. The authors also discarded a lot of information in this study, perhaps due to this faulty assumption. For example, seven of the original twenty-eight subjects’ data were thrown out because of “failure to follow instructions” (p. 659), and 3% of the remaining data were discarded because subjects recorded an image formation time of zero.

Although their study should be carried out again under better circumstances, their hypotheses do inspire thought regarding the present discussion. In both aural and visual imagery of musical materials, “hearing” or “seeing” a chord versus “hearing” or “seeing” a single note may require more time. Notes and chords in distant harmonic relationships may require more processing time than those in close harmonic relationships. The extent of differences in processing time may depend on listeners’ familiarity and fluency with music theory concepts, and their efficiency at scanning mental images.

Halpern (1988) conducted four experiments involving aural scanning of known melodies. The first investigation (the most relevant to this discussion) was devised to determine whether the maintenance of spatial distance that has been observed in visual imagery scanning experiments would prove useful in aural imagery studies as well. Forty-one undergraduate students were divided into two groups to take a true/false quiz on familiar songs. The students were shown the title of the song and one word of the lyrics. When shown a second word from the lyrics, subjects from the control group were to choose as quickly as possible whether it belonged to the lyrics, while those of the experimental group were to imagine the intervening part of the song and indicate that the word belonged to the lyrics only when they had arrived at that point in the song.

Halpern found that the control group responded faster than the imagery group. She also found that reaction times for all subjects grew longer with the increase in the amount of intervening beats between the two words given. Overall, subjects required more time to make judgments about lyrics that were temporally farther apart in the song. All subjects were using some degree of aural scanning to accomplish the tasks, a conclusion supported by questionnaires that asked whether or not students were imagining the song and to what extent. Halpern concluded that students made use of aural imagery with or without instruction to do so.

Halpern made no specific references to the possibility that the aural imagery she was trying to evaluate might be related to visual imagery (i.e., images of the words themselves), although she does mention scanning lists of the lyrics as a possible alternative explanation for results. Some combination of aural and visual imagery may have been at work in this experiment because the music used contained words.

The studies of aural perception and imagery presented thus far suggest that aural perception in music depends on prior musical experiences, understanding of common systems underlying musical organization, and mental conditioning resulting from musical training. It appears that aural percepts and images can be scanned freely and manipulated, based on listeners' musical expectations and whether those expectations are satisfied. Perception and cognition of melodies is a complex process, involving pitch and rhythmic information in combination. The length of a perceived melody appears to affect processing times during cognition. Although the holistic nature of complete melodies is important, accurate perception and cognition of melodies may require division into smaller sections.

Visual Perception and Imagery

Studies drawn from reading and reading comprehension shed light on the process of visualizing music notation, a skill possibly useful when transferring aural stimuli to written form. Weber, Kelley, and Little (1972) observed undergraduate students to see whether verbal imagery—subvocal verbalization—of each of a string of letters is necessary before those letters can be visually imaged. They conducted a series of three experiments, the last of which has implications for the current study. In that last experiment, participants were given a mixture of 4-letter words and non-word strings of letters in alphabetical order. They were asked to scan those letters in one of two ways, out loud and silently. Then they judged the relative vertical sizes of target letters as “large” or “small,” (e.g., “k” or “s,” respectively), as quickly as possible.

The most pertinent result for music teachers is a significant interaction of scan and string, suggesting that, for those who scanned silently, scan times for words were significantly faster than scan times for strings of letters. The researchers conclude that

> The visual imaging of individual letters in an alphabet string is under verbal control in the sense that each letter is implicitly spoken prior to imagining it. In contrast, the results for word strings are not equivalent for spoken and nonspoken scanning and are therefore inconsistent with the notion that [a subject] implicitly speaks each letter of a word when using a nonspoken scan mode. Verbal control over the sequential generation of visual images of the letters is not needed with the word letter-strings used in this study. Instead, on the basis of a word's spoken name, [a subject] may visually generate the word as a whole, or at least with more than one letter at a time simultaneously represented in visual imagination. (pp. 361-362)

Students of melodic dictation may find that musical motives operate in much the same way as words. Groups of tones that occur in readily recognizable patterns (e.g., parts of an arpeggio or scale) may be understood and visually scanned all at once instead of note by note (Madsen, 1983; Miller, 1956). This should facilitate faster scanning and more efficient use of time, enabling optimum cognition.

Weber and Harnish (1974) performed two experiments investigating the retrieval and creation of words in visual imagery. They aimed to dispute the theory that a word in mental imagery can be seen only in left-to-right letter order, and not scanned freely like other visual percepts. In the first experiment subjects were given a mixture of visual percepts and visual imagery prompts for three- and five-letter words. Subjects were instructed to answer "yes" or "no" to probes of particular letters and whether they appeared in a given visual percept or image. Word length made a significant difference in response time results, but word presentation (percept vs. image) did not.

In the second experiment the probe letter was presented *before* the word (in both percept and image prompt forms) to measure response times related to image generation. In this case the interaction between length and format was statistically significant, suggesting that word length and image capacity affect the ability to generate and manipulate images. The two experiments present an interesting perspective on the ability to create and scan meaningful images, which is exactly what is needed in the dictation process. Unfortunately, experiments that employ "yes/no" responses to measure imagery activity risk high participant error, because of the potential for participant guessing. Indeed, the researchers report that "Medians were used because of the substantial variability encountered" (p. 411). Although they corrected for the variability as much as possible in their analyses, an alternative experimental design to test similar hypotheses would be appropriate.

Rinck and Denis (2004) investigated effects of Euclidean distance and categorical distance on visual imagery of spatial movement through a series of rooms and hallways. Euclidean distance refers to standard linear distance while categorical distance refers to the number of units that have to be passed (e.g., number of rooms, in this experiment). They gave subjects a drawing of a building layout. Each room in the drawing contained the name of an artist, representing a painting hanging on a wall of that room. The participants studied the layout and self-tested on the information until they knew it completely. They were then given imagery instructions one-by-one on a computer screen. These instructions included imaginary movements from room to room and identification test items in which the participants had to decide where certain paintings were located in relation to their current imagined position.

The purpose of the study was to determine if mental scanning of the visual imagery used by the students would be affected by both Euclidean distance and categorical distance. The participants were asked to imagine themselves actually moving about through the rooms based on the instructions. The paths outlined in the instructions varied in two ways to examine the effects of the two types of distance under scrutiny: long versus short “Euclidean paths,” and one-room versus two-room “categorical paths.”

Rinck and Denis found that scan times increased with longer paths, and with two-room paths. These findings are important to those who may use, learn, and teach visual imagery for musical tasks. Accurate and efficient processing of visualized notation imagery requires understanding both Euclidean distance—the linear distance from one note to another—and categorical distance—the number of semitones traversed regardless of the spelling of the interval.

Visualization of information, ideas, and systems is common in math, science, and language arts. Its usefulness in the classroom has been tested in experimental studies. Center, Freeman, Robertson, and Outhred (1999) conducted a study to determine whether visualizing items and concepts would boost listening comprehension scores more than traditional approaches that did not explicitly encourage visualization. Sixty-six low-scoring students in “Year 2” (mean age: 7 years, 8 months) from four schools were matched and placed randomly into two groups. Both groups were given instruction in the same approaches to listening comprehension, but the experimental group was also given instruction in representational visual imagery, while the control group received extra instruction in discussing and describing the details of stories.

Initial selection of subjects was based on a version of the *Neale Analysis of Reading Ability* (1999) that had been adapted for listening comprehension. Once students had been assigned to groups, they took three pre-tests: the reading accuracy and reading comprehension portions of the *Neale Analysis*, and an individual storytelling measure. The latter was graded on a 4-point scale based on coherence, sequencing of main events, and inclusion of a climax and resolution.

All students then received the 20-minute listening comprehension sessions three times per week for four weeks. Two teachers trained in approaches to listening comprehension and visual imagery each taught control group students and experimental group students. All students received the general training, which included discussion of narrative elements, background knowledge, questions about example stories, and plotting of events in sequence. Only the experimental group students received the representational visual imagery training, which involved visualization of common objects, teacher models of sentence visualization, and students' own sentence visualization. The control group students received extra practice in description and discussion during that time.

At the end of the treatment, all students were given the following post-tests: parallel forms of the accuracy and reading comprehension portions of the *Neale Analysis*, the *Neale* (adapted) listening comprehension measure (Form 2), the *Byrne Listening Comprehension* (1995) measure, and a story event structure test. Although no significant difference in mean scores was found on the *Neale Listening Comprehension* test, the researchers found that the experimental group outscored the control group on the *Byrne Listening Comprehension* test, the comprehension portion of the *Neale Analysis*, and the story event structure test.

Although studies of comprehension and visualization would benefit from more standardized and validity-tested listening comprehension measures, Center et al. found encouraging support for the use of visual imagery intervention to improve listening and reading comprehension. One of their conclusions was

> What visual imagery instruction may have provided for poor comprehenders is a non-verbal conceptual peg, on which associated information is hooked for storage and retrieval and which can compensate to some extent for weaker phonological integration skills. (p. 251)

The model suggested by this conclusion readily translates to mental visual imagery used in melodic dictation exercises. Visual imagery may bolster the perceptual and memory capacities for aural stimuli during dictation.

The studies of visual perception and imagery outlined in this section suggest that meaningful strings of visual information can be scanned faster than random strings. Both Euclidian length and categorical length appear to affect processing times for visual imagery tasks. Visual imagery may be able to be used to aid comprehension and memory in conjunction with other skills, such as verbalization, reading, and listening.

Relationships Between Aural and Visual Perception and Imagery

Taking melodic dictation requires transferring aural stimuli into visual output. The end product is a written melody that accurately represents the original stimulus. It seems a simple task, but how do students actually make the mental transfer, and produce the written product? Why do some students have such difficulty completing the task? The following section examines relationships between the aural and visual domains, providing a foundation for exploration of both in combination during dictation.

Segal and Fusella (1970) observed relationships between images and percepts in the aural and visual domains simultaneously. Subjects were instructed to create various aural and visual images. As soon as the subject indicated that the image had been generated, the researchers introduced either an aural or visual signal at a very low intensity (low enough to guarantee at least some errors, as tested in pilot studies). The subjects were then asked whether they had perceived a signal, and, if so, which.

The researchers found that accuracy of answers was much higher—taking into account correct answers versus false alarms in each case—when cross-modal relations occurred between percept and image. In their words:

> The [subjects] apparently confused auditory images with auditory signals, visual images with visual signals; e.g., (sic) when [a subject] was imaging a mental picture, he was significantly more likely to miss a visual signal and also more likely to guess that a visual signal was there when there was either nothing or an auditory signal. (p. 463)

Confusion of simultaneous auditory images and percepts is the central problem of the current study. If students can use cross-modal relations during dictation, they might avoid aural confusion. This lends support to the idea that visual imagery may be a viable approach to short-term memory in melodic dictation, because visual images may be less affected than aural images by additional aural stimuli.

In the first of two experiments, Iwamiya (1994) used laser discs to observe the influence of audio and visual tracks on subjects' perception. Subjects viewed various audiovisual materials (music videos, music with video scenery, and cartoons with audio) in their original state, and with mismatched audio and visual tracks. They rated the “tightness,” “evaluation,” “cleanness,” “brightness,” and “uniqueness” of sound and sight with scales such as “loose----tight,” “powerful----powerless,” and “pleasant----unpleasant.” Unfortunately, these terms are rather open to interpretation; they may mean very different things to different people.

Iwamiya ran correlations between unimodality perception scores and unimodal versus bimodal perception condition scores to see how each modality affected the other. Statistically significant differences for “tightness,” “cleanness,” and “brightness” were found between audio scores overall and the difference between visual scores in the unimodal and bimodal conditions. He concludes “It appears that the audio systematically influences the visual-factor but not vice versa. These results show the existence of consonance [a relationship between the two modalities] and the direction of influence from auditory to visual processing” (p. 141). This supports the possibility that auditory perception can shape and enhance visual imagery. Although nebulous terminology detracts from the strength of the results, this report does provide an interesting complement to Segal and Fusella’s (1970) aforementioned research, laying a foundation for the current study.

While tests of isolated single images are helpful, the ability to create and manage manipulable images within a given context is perhaps the most important clue to the current research problem. Guttman, Levin, and Pressley (1977) explored the usefulness of complete and partial pictures in helping young students (K-3) form visual images for comprehension of oral sentences and stories. They note that as early as third grade, participants were able to create images on their own without partial- or full-picture help. This finding supports the possibility that students might use a visual representation of aurally-perceived stimuli for comprehension of those stimuli during dictation.

Guttman, Levin, and Pressley also reported that “Subjects presented with dynamic partial pictures recalled substantially more propositions [subject-verb pairs] than those presented with static partial pictures.” (p. 479). Dynamic pictures show the action in a situation, while static pictures simply show the main protagonist or object of interest. These results suggest that “imagery in motion” may be more effective than static images for comprehension of dynamic (e.g., melodic) information.

Dynamic imagery can be cultivated using dynamic tangibles. Cassidy (2001) tested the ability of undergraduate students (music education majors and non-music majors) to follow listening maps, which are learning tools typically used in elementary school music education. Three listening maps corresponding to three musical examples were used to observe three levels of ability in matching iconic representations to the music: note-by-note, measure-by-measure, and section-by-section. Specific target icons from the listening maps were highlighted in red. Students were to push the “1” key on the computer each time the music arrived at one of these red icons.

Cassidy’s purpose was to see how user-friendly the listening maps are, especially for non-music majors. Because these subjects had little music experience beyond elementary school themselves, they were seen as representative of many elementary school teachers, given the task of teaching music in the classroom without much musical education of their own.

Cassidy determined that while there was no significant difference in results for music majors across iconic representation formats, there was a significant difference in results of non-music majors. For them, accuracy decreased as tasks became more abstract; they were more accurate with notes, than with measures, than with sections. It is also important for this discussion that music majors performed extremely well on the nine icons of the note-for-note test. It seems that visual representation of melodic sequences does not hinder ability to listen critically and, in fact, may enhance that ability. Indeed, “Nearly half of the music education majors and nearly a third of the non-music majors cited the [note-for-note] listening map as the reason they preferred the music [over the other musical examples]” (p. 19). Cassidy suggests that listening maps seem to be quite useful in general, but should certainly be as clear and practical as possible, especially for relatively untrained teachers. The visual domain may be used to aid and enhance the aural domain during dictation, but students’ skills must be sharpened carefully and accurately to do so.

Parallels are frequently drawn between music mastery and language mastery. For instance, both systems, in most Western cultures, involve visual symbols that represent auditory information. Forgeard, Schlaug, Norton, Rosam, Iyengar, and Winner (2008) examined the relationship between musical aural processing ability and reading ability, in children with dyslexia and in normal-reading children. They compared children with instrumental musical training to those without, both at baseline and throughout their four experiments.

The results of Experiment 1 showed that in normal-reading children, phonological skill and musical pitch processing are related, and that the relationship is stronger for those receiving instrumental training. Experiment 2 revealed the same relationship patterns for reading ability and aural musical skills. Experiment 3 found subjects diagnosed with dyslexia performing below normal on Edwin Gordon's *Intermediate Measures of Musical Aptitude* (1986). Experiment 4 showed musically-trained normal-readers outperforming non-trained normal-readers in melodic discrimination, and both groups of normal readers outperforming subjects with dyslexia in both melodic and rhythmic discrimination.

There are several specific problems with this study. Samples were very small. Baseline musical training was not clearly defined or considered, and results are somewhat inconclusive. However, the main two findings, supported by several citations of extant research in the authors' report, appear sound: "Music training enhances phonemic awareness….phonemic awareness predicts reading abilities" (p. 388).

While the researchers' interest was in reading and language ability, the results bear implications for music educators as well. If both music and language involve visual symbols that represent aural information, both music and language might be studied in similar fashion. "Phonemic awareness" in musicians may predict musical reading ability as well, which would support the "sound before sight" approach to musical learning. It would also suggest that educators should encourage students to make strong connections between the aural and visual domains, as early as students are ready.

Language teachers help students make these connections using flash cards, word games and songs, and rhyming words. Eventually, the connections are made in the imagination, rather than with physical manipulatives, spoken words, and songs. Music teachers can help students learn to connect aural and visual information in similar fashion.

Lucas and Gromko (2007) examined relationships between music pattern discrimination and phonemic awareness in reading. Specifically, they looked for matching trends in scores on tonal and rhythm pattern discrimination in comparison to scores on phoneme segmentation tests. Twenty-seven first graders took the *Phoneme Segmentation Fluency* and *Nonsense Word Fluency* subtests of the *Dynamic Indicators of Basic Early Literacy Skills* (Smith & Good, 1988), and the tonal and rhythmic subtests of the *Primary Measures of Music Audiation* (Gordon, 1986). The researchers found significant correlations between the mean composite scores for *PMMA* and phoneme segmentation, tonal *PMMA* and phoneme segmentation, and rhythm *PMMA* and phoneme segmentation.

The results of their study indicate that internal understanding of the implicit sounds that make up written words are important. In other words, a strong connection between the sight and sound of words is important to fluency in reading words; one would assume the same for musical processing. The report seems to suggest this can happen as early as first grade. Although young children typically have more experience with language than with music, the necessary cognitive tools appear to be in place at that age.

Specific connections between aural and visual learning approaches have been found in relation to melody perception. Mikumo (1994) explored whether motor skills could help piano majors remember short melodies. Although she was not specifically looking for relationships between the aural and visual domains, she found them in the course of her study.

Well-trained musicians tapped their fingers as if playing the piano during the memorization phase between pairs of target melodies and matched/unmatched melodies. A computer recorded the tapping movements of each subject. These examples were compared with another set of examples in which the subjects had been told specifically not to do any tapping, but rather to use any method they wanted to remember each target melody. These two methods of pitch encoding were carried out for melodies of six, eight, and ten tones, and were also presented in two other experiments. In Experiment 2, an intervening melody was employed to counterbalance the effect of possible subvocalization by subjects, and in Experiment 3, an intervening series of solmization syllables was read to combat solmization as an aid to pitch memorization.

These distracting interventions should not have disturbed someone using visual imagery to retain the target melodies. Indeed, Mikumo reports

> In the case of Subjects 2, 4, 7, and 10, there was little difference among the numbers of incorrect responses in Session 1 [the session without tapping] of the three experiments; that is, there was little disruptive effect of the interfering melodies or the series of note names. This suggests that these subjects were using an encoding strategy other than pitch rehearsal or verbal encoding. They reported that they used a visualizing strategy in which tones were visualized in their image, as contours (Subject 2), in notation (Subjects 4 and 10), or on a keyboard (Subject 7). (p. 185)

While these may seem like only a few specific cases, they are actually one third of the group. It seems that visual imagery for melody memory is indeed employed by some trained musicians.

Visual imagery related to musical tasks has been discussed mostly in terms of music notation so far. Dictation-takers may employ an array of visual schemes to represent aural stimuli. Karpinski (2000) discusses "protonotation," a system of quick, efficient dashes and marks that later can be translated into formal notation. In my experience, students have used visualized strings of numbers, contour outlines, and the piano keyboard to contextualize percepts. All these approaches merit future study.

Giannakis and Smith (2001) observed relationships between aural and visual domains using sound sequences and color variables. Subjects representing a wide range of musical experience were presented the task of matching six color choices to six tones in each of 11 sequences. The color palette consisted of 216 total choices based on variations in hue, saturation, and value. Based on sequences of colors chosen to match the sequences of pitches, the researchers were able to map general correlation trends between these three color attributes and the sound attributes under scrutiny, loudness and pitch.

The most salient results of the experiment occurred in sequences where only loudness was altered and in sequences in which loudness and pitch were altered inversely, although the overall findings are not completely convincing. In sequences (three per subject) where only loudness was altered, "…hue varied in 1/72, saturation in 38/72, and value in 9/72 sequences" (p. 168). In sequences of descending loudness and ascending pitch, hue remained constant in 17/24 sequences and with ascending loudness and descending pitch, hue remained constant in 16/24 sequences. Saturation and value were the two color attributes that varied the most in these examples.

The researchers concluded that "Quiet tones were associated with low levels of saturation while louder tones with increasing levels of saturation. Furthermore, low-pitched tones evoked dark (low levels of light intensity) colour selections while high-pitched tones were associated with lighter colours" (pp. 172-75). Although the presentation and analysis of numerical results is questionable, the general trends observed in this study are intriguing. The results and conclusions are not thoroughly convincing, but they do support the idea that relationships between the visual and aural domains should be further investigated and applied to music learning.

The studies of relationships between the aural and visual domains described in this section suggest that a cross-modal approach to memory for percepts in either domain may increase perception accuracy for subsequent information in the same domain. Children appear to be able to create visual images representing aural information as early as kindergarten, a skill which may increase focus on the information presented, especially when images are dynamic and manipulable. Some advanced musicians use visual imagery of melodic contours, music notation, or the piano keyboard to encode musical aural percepts and retain them in memory.

Skilled Imagery

Skilled imagery implies two characteristics: (1) images are dynamic and manipulable, and; (2) images are related to a specific skilled task. For example, the studies that follow investigate the dynamic imagery schema used by math students, expert abacus operators, blindfold chess players, and musicians, to carry out specific tasks in their fields.

Pinto and Tall (2002) conducted a case study of Chris, a first-year university mathematics student. Through a series of interviews, they tracked his thought processes and visualization tactics for understanding the mathematical concept of the limit of a sequence. Following are several quotes of the authors of the article. They are presented in full because they are rather illustrative of the process of skilled imagery development, and can be applied directly to dictation strategies.

> It [the visual image of the limit of a sequence] moves up and down in a more general manner. He uses it not as a specific picture, but as a generic picture, one that represents the sequence in a manner that is as general as he possibly can make it…. His combination of mental imagery, its verbal equivalent and its ensuing properties all fit together as a powerful cognitive unit….To be able to realize his error and to self-correct requires him to build a stable schema, one in which dissonant elements can be perceived and corrected whilst other parts remain coherently in place. It is precisely such a mental schema that Chris builds by manipulating his imagery and adapting it as his ideas mature. He has a broad enough grasp of the principles to be able to deal with errors and misconceptions, and to reconstruct them in his quest for a coherent whole. (p. 3)

> For Chris, it is not a matter of *first* the process, *then* the object; rather, he has a dynamic gestalt which comprises them both. (p. 4)

> The crucial idea is to understand how his image characterizes the mental concept of convergence that he is attempting to construct. This mental construction involves playing with the image in various

> ways….To summarise, Chris interprets the definition in terms of his old knowledge, explores the concept through thought experiment and reconstructs his understanding of the concept definition. He compresses information in a picture, which he evokes when writing down the definition. (p. 5)

These thoughts could be easily transferred to visually-imaged music theory concepts and used in melodic dictation. With well-defined, clearly-connected music theory concepts at the ready, students may be better equipped to contextualize dictation percepts efficiently and accurately. As new information is learned and connected in new ways, the overall schema improves, fostering more refined aural-visual transference skills.

Attention to encoding of information is vital in memory research. Jakobson, Lewycky, Kilgour, and Stoesz (2008) highlight the importance of sequencing, both when memorizing music and when memorizing poetry or prose. In both cases, individual elements are important, but their temporal relationships to one another are equally important. Preceding events can serve as triggers to memory of subsequent events.

The authors used verbal and visual learning tests to measure differences between musicians and non-musicians on memory retrieval tasks. They concluded that "Music training is associated with superior delayed free recall of both verbal and visual information" (p. 50). They note two possible explanations of test scores: subjects may differ in their ability to encode information for later retrieval, and subjects may differ in their ability to access the encoded information. Because the musician subjects scored better than non-musicians, but not comparably to each other, the authors suggest that the advantage likely lies in encoding of the information. Musicians may have used better encoding strategies, involving higher-order semantic information, when memorizing the verbal information.

The authors suggest the same for visual task performance: musicians demonstrated better long-term recall of simple visual designs based on their ability to effectively encode the information in systematic ways. The authors warn that their results cannot be assumed to show causality, but that their findings do indeed demonstrate the importance of encoding strategies in the processing of specific information, and in the long-term shaping of cognitive processes.

Encoding strategies may begin with physical manipulation of information and then transfer to imagined manipulation of information. Hishitani (1990) investigated skilled imagery of expert and non-expert abacus operators, ages 13 to 18. The researcher was looking for relationships between expert/non-expert status (based on standardized ranks), relative levels of command of skilled abacus imagery, and subsequent dependence on verbal short-term memorization of numbers during numerical imagery tasks. Variables that influenced subjects' response times to probed digits were number presentation (whole number versus sequence of digits), probe position (within the multi-digit number), and imagery ability of subjects.

Hishitani predicted that expert abacus operators would use an image of the abacus to encode the numbers as they were read, and would therefore demonstrate shorter response times than non-experts, whom he assumed would rely on verbal encoding of the numbers. This would be especially apparent in numbers presented in whole form, for example, "one thousand six hundred eighty-four" instead of "one-six-eight-four." Hishitani also entertained the possibility that the experts might use some combination of mental imagery and verbal encoding if the given number of digits was beyond the image capacities of the subjects.

Hishitani had to rely on subjects' introspection to determine whether they were using imagery, to what extent, and what the capacity for imagery of digits by abacus was for each subject. This introspective method of inquiry was defended thus: "Imagery is a private mental practice….if there is no essential principle for judging plausibility [of an image theory], to the extent that a plausible theory can be constituted, there is no reason why introspection cannot be a source of data" (p. 33).

Hishitani found that non-experts employed verbal encoding, while experts used verbal encoding combined with dynamic imagery. Non-expert response times were longer than expert response times, and longer series produced longer response times overall.

Hishitani also found that experts did not differ significantly in response times based on presentation of numbers (whole versus digit-by-digit), because they were able to transfer the information immediately to an abacus image for later recall, whereas non-experts had to review the number subvocally to arrive at the probe digit. However, the experts did show a significant difference in this aspect for 6-digit numbers, the longest in this study. These longer numbers probably exceeded the image capacity of the experts, forcing them to rely on subvocal extensions to complete the memory store (Miller, 1956). These findings have important implications for memory of melodic pitch information: music students may be able to use visual imagery to store and recall a limited number of notes on the musical equivalent of a mental abacus, for faster retrieval. The image is manipulated to show newly-encoded information in both cases.

Manipulable imagery constructs have also been studied in blindfold chess players. Blindfold chess refers to the standard game, played without the benefit of seeing the board. Players' moves are spoken, and each has to remember the entire board layout as it changes—a manipulable imagery construct. Saariluoma and Kalakoski (1997) compared master players with intermediate players, based on official chess ranking, in a variety of measures of skilled imagery ability. The participants were asked to keep track of three computer-simulation chess games at once, and to demonstrate recall of the chessboard layout at certain points in each game.

In Experiment 1, the researchers found no statistically significant difference in recall based on presentation of dots on the screen versus presentation of chess piece images. In Experiment 2, overall recall was significantly lower when the board was "transposed." A transposed chessboard refers to an imaginary chessboard that has been cut in half and reversed (like cutting a deck of cards).   In Experiment 3, no significant difference was found based on visual directions versus aural directions, and in Experiment 4 results were non-significant for recall of masters compared to recall of intermediates, under imposed time limits.

The results of this study indicate that skilled imagery can be visualized equally well in multiple guises, but should be kept within an established order or system. The study also suggests that aural and written prompts work equally well in stimulation and recall of skilled imagery contents, and that time limits imposed on the process are not necessarily more taxing on less advanced practitioners. These guidelines may be helpful to students and teachers tackling the skill of music theory imagery, for use in dictation and other musical tasks.

In another study, Saariluoma and Kalakoski (1998) investigated the possibility that skilled imagery was affected by the relative functional importance of a given chessboard image. They asked blindfold chess players whether given positions of chess pieces were the same or different as the positions that the participants had been instructed to memorize. They found that “There is a large and statistically significant difference between significant and insignificant minimal transformations [of the images]” (p. 72). In Experiment 2, participants who had been told to find the best next move in a board arrangement displayed a “clear difference in correct answers between insignificant change and significant change positions” (p. 75), while those who had been told to simply count the pieces of each color on the board did not show a significant difference.

The results of Experiment 3 showed that the background arrangement (functionally important versus randomized) of chess pieces also affects accuracy of recall. “Subjects could recall more pieces from the control positions than from background randomized positions” (p. 78). In Experiment 4, the researchers again found no statistically significant difference between visual and auditory explanations of the chessboard layout. The results most relevant to the possibility of imagery use during melodic dictation are the importance of functionality of the target images, the background against which they are formed, and the viability of aural triggers for visual imagery manipulation.

The studies of skilled imagery presented in this section suggest that development of a visual mental schema representing the content of a skilled discipline contributes to greater understanding of it. Dynamic visual images are constantly reconstructed to accommodate new information, and to organize it for effective application. Skilled imagery success appears to depend on level of expertise of those who employ it, and the relative functionality of new information to be assimilated. It appears that skilled imagery can be triggered by either visual or aural directions.

Imagery in Musical Applications

Researchers have explored the imagery paradigms discussed thus far in relation to specific musical tasks. They reveal insights on the cognitive processing necessary to accomplish tasks involving transfer between aural and visual channels. These studies shed light on this discussion, and provide a conceptual framework for the current study.

Aural skills

Melodic memory is directly related to aural skills in music. In fact, it may be considered a part of aural skills. Aural skills are defined here as the broader ability to perceive and make musical decisions about aural stimuli.

Parallel to the chessboard examples described in the previous section, the function of target images in music helps the listener organize percepts according to musical expectations. Subjects may perceive melodies using multiple strategies, including recognizing small clusters of notes as an outlined chord, and anchoring a tone to a later tone of resolution (Povel & Jansen, 2001). Using sets of tones incorporating members of the C7 chord, and the tones F-sharp and G, Povel and Jansen used subjects' "goodness ratings" to determine if those ratings were related to the complexity of the perceptual task (i.e., if the series of tones required chord recognition, anchoring, both, or neither).

Fragments containing only chord tones were rated highest, those with an F-sharp-to-G resolution ranked second, and those with an unresolved F-sharp were rated lowest. The importance and validity of their results are compromised by the tenuous connection they draw between "difficulty" of the perceptual task (necessity of chord recognition and anchoring) and "goodness" ratings of the subjects. The researchers assume that perceptual tasks that require more strategies will be counted as less "good." It seems, however, that the exact opposite could be true: satisfaction of chord recognition and anchoring expectancies could improve subjects' "goodness" ratings.

Povel and Jansen's subsequent experiment involves the same concepts of chord recognition and anchoring, but in a melody completion task. The questionable measures of melody "goodness" no longer distract the reader, and results are based on the notes that subjects supplied to complete each tone series. The researchers found that chord recognition and anchoring were again important factors in subjects' responses. These strategies may be useful when taking dictation, as they represent two methods of melodic "chunking" (Madsen & Staum, 1983; Miller, 1956).

Povel and Jansen suggest a certain degree of flexibility in the mind while a melodic sequence unfolds. Listeners may alternate between uncertain and certain percepts as their expectations are unfulfilled or fulfilled, respectively. Dictation-takers therefore must alter their percepts in real time, or shortly thereafter. The difficulty lies in focusing the mind on the percept being updated. If subsequent portions of the melody are playing, listeners need a way to represent perceived information in the mind, without the distraction of additional information. Visual imagery of the percepts may allow the auditory channel to be kept clear, so that functional decisions can be made.

A manipulable schema serves the mind to field incoming information, compare and contrast it to existing knowledge, and adjust mental constructs if necessary. In music, working schema and memory systems may be approached through a cue abstraction mechanism (Koniari, Predazzer, & Mélen, 2001). This technique involves listening for aural "landmarks" in music, bits that stand out from the surrounding music and serve as important cues. The sections of a piece are then able to be organized in the listener's mind for memory, analysis, manipulation, retrieval, and synthesis.

Koniari, Predazzer, and Mélen investigated the cue abstraction mechanism in relation to segmentation, categorization, and reconstruction abilities of 10- to 11-year-old musicians and non-musicians. The reconstruction task is most relevant to this discussion because the culminating step of melodic dictation is reconstruction of aurally-perceived events in correct order. The researchers found that musicians performed significantly better on the reconstruction task, with primacy and recency effects (correct first and last sections selected correctly most frequently). They suggested that mental processes used in the task depend both upon general, long-term understanding of good order and organization in music, and specific memory for music under study. As new music is perceived, it must be evaluated in relation to existing musical schema to be used effectively in musical tasks.

Koniari, Predazzer, & Mélen's experiment involved putting together large excerpts of full pieces (the *Finale* of Diabelli's *C major Sonatina* and Schubert's *Dances for Piano*), not short monodic melodies. However, these approaches to schematization and reconstruction appear quite applicable to short melodies used for dictation as well. Perceiving new aural information correctly, evaluating it according to existing musical schema, and reconstructing it in a musical task are essentially the core requirements of dictation. The way listeners carry out these tasks has not been studied thoroughly. Do they process each step through inner hearing? Through visual representation? Through kinesthetic connections? Through verbal thoughts? Through some combination of these processes?

Palmer and Krumhansl (1987) presented subjects with three versions of a Bach fugue subject: the original melody, the pitches intact with altered rhythm, and the rhythm intact with altered pitches. In Experiment 1 the alterations were accomplished by simply standardizing rhythm (equal note lengths) or pitch (all rhythmic values presented on the same note). In Experiment 2 the variations of the melody were created by shifting either the pitch or the rhythm, thereby distorting the original melody, but maintaining "active" aspects of both pitch and rhythm. The subjects rated each melody presentation on a scale of 1 to 7, based on how complete the phrase sounded. Both experiments were designed to compare the three conditions, and the researchers found significant correlations among the results of both experiments for melody, temporal, and pitch conditions.

Experiment 1 revealed a significant correlation between ratings under melody and temporal conditions, and between ratings under melody and pitch conditions, but not between ratings under pitch and temporal conditions. Experiment 2 showed similar results: significance between ratings under melody and temporal shift, and between ratings under melody and pitch shift, but not between ratings under temporal shift and pitch shift. Palmer and Krumhansl concluded that both pitch and temporal considerations seem to affect perceptions of musical phrasing independently, but that the two are not interrelated in the process of perceiving music. This may have implications for the development of strategies for hearing, encoding, remembering, and notating melodic information. Students may be focusing on pitch only, or rhythm only, at any given point in the dictation process. The current study explores this focus, based on participant scores in each of two answer categories, pitch and rhythm.

Beckett (1997) compared "pitch-first," "rhythm-first," and "non-directed" strategies in students' accuracy with two-part dictation. Sixty second-year ear training students took two-part dictation examples in three different sessions (representing the three strategies), each 10 to 14 days apart. Students were given four listenings in each session. They were given specific instructions that on the first two listenings of each example they were to notate only rhythm in the rhythm session and only pitch in the pitch session, filling in the rest on the remaining two listenings. In the non-directed session they were allowed to take the dictation however they wished.

Results showed that rhythm scores from the "rhythm first" condition were significantly higher than rhythm scores from the "non-directed" condition and "pitch-first" condition. There were no significant differences between rhythm scores of the "pitch-first" and "non-directed" conditions, nor were there any significant differences between any of the pitch means. Beckett reports: "Subjects said that without rhythmic organization they could not hold pitches in memory" and concludes that "Having the rhythmic structure in place might aid pitch memory" (p. 621). This conclusion strengthens the case for the utility of a manipulable framework on which to peg incoming melodic information.

Pembrook (1986) tested 136 undergraduate music students to compare the effectiveness of three strategies for taking melodic dictation: simultaneously listening and writing, listening to the complete melody before writing, and listening, singing, then writing. He found no significant differences between the strategies.

Pembrook did, however, find statistical evidence that subjects who listened to melodies twice performed better than those who heard melodies only once, and that overall accuracy decreased as melody length increased from six to ten to sixteen notes. Regarding repetition, he concludes: "It is possible that a first hearing provides a schemata or expectancy that allows more meaningful interpretations of the events during a second playing" (p. 254). As for melody length, he suggests

> Errors in dictation for longer melodies more likely reflect an inability to remember all of the tones rather than deficiencies in the process of notating them. In this instance, emphasis should thus be placed on careful listening or coding such that the problem of limited memory can be overcome. (p. 260)

Pembrook is suggesting a weak or missing memory link between perceiving notes correctly and transforming them into written notation.

A benefit of using visual imagery to represent systems of information is that the visual image can be scanned quickly and manipulated at will. Scanning of visual images may indeed operate in much the same way as scanning of physically-present visual percepts (Weber & Harnish, 1974). If visual imagery is to be used effectively in musical tasks, the skill of image scanning must be explored and improved by those struggling with the tasks.

Perceptual span of musical notes may be increased with rehearsal (Burman & Booth, 2009). As musicians repeatedly rehearsed a passage at the piano, their range of visual perception (measured in number of notes) increased as they were tested for identification of a target error note. Burman and Booth noted that, when allowed to visually review the notation of a passage, subjects who could not read music at all did not increase perceptual span range. Their inexperience with music notation essentially blocked their ability to increase perceptual span on music notation tasks.

Perceptual span capabilities for musical notes, much like perceptual span capabilities for words, appear to depend on familiarity with the note patterns in a given example. Skilled musicians in Burman and Booths' study displayed greater perceptual span than less-skilled musicians when tested on new musical excerpts with familiar note patterns. Their strong level of familiarity with common patterns enabled them to scan new material quickly and efficiently.

Burman and Booth suggested that once note patterns have been truly memorized and connected to motor movements (as in the later stages of their experiment), perceptual span becomes less important. Therefore, they concluded that increasing perceptual span is most important during the initial rehearsal of material.

Since melodic dictation generally involves only a few listenings, attention to perceptual span may be an important link in pedagogical approaches. As students attempt to process unfamiliar melodic information as accurately and quickly as possible, they may wish to employ any of various visual representations of the aural information. If so, they may capitalize on chunking familiar patterns, thus increasing perceptual span within visual imagery.

The concept of chunking melodic information is described in a study by Madsen and Staum (1983). They tested undergraduate non-music majors on their ability to identify same and varied (by mode or meter) renditions of a stimulus melody. Following each stimulus, subjects heard eight additional melodies. One was the same as the stimulus, one was altered only in mode or meter, and the others were different from the stimulus. The subjects decided whether each of the eight probe melodies was the same or different as the stimulus melody.

The researchers calculated the percentage of subjects who marked "same" for each probe melody and ranked them to see what the subjects perceived. They found that, overall, "Subjects do evidence ability to discriminate among melodies that are identical or very similar to an initial melody over an extended period of time" and that "It may be that melodies function similarly to the verbal phenomenon of 'chunking,' in which a comprehensible unit that is formed is less vulnerable to interference than unrelated single items" (p. 24). Skill in chunking could prove a valuable asset to melodic dictation students, in terms of both comprehension and memorization of melodies.

Researchers have studied various specific approaches to melodic memory. Pembrook (1987) conducted an experiment with undergraduate music theory students, looking for effects of singing back a melody on retention of the melody. He assigned them to three groups to measure their ability to identify pairs of melodies as the same or different. Students in the first group decided after a two-second pause between melodies, while those in the second group endured a nineteen-second pause between melodies. The third group sang back the first melody before the second melody was played.

Scores from the first and second groups were significantly higher than those of the third group. Apparently students' singing actually hindered their mental storage of the melodies. Pembrook cites poor singing of tritone intervals as an example of one of the difficulties that may have negatively affected students' performances on the discrimination tasks. When students sang the melodies completely correctly, they performed very well on the written responses; however only 11% of 597 total responses fell in this category. Interestingly, Pembrook found no significant difference between the scores of vocal music majors and others. Apparently the singing itself was not the problem; it was rather a lack of mental understanding and memory of the musical material.

Pembrook concludes from the study that

> A majority of students do not sing accurately enough after one hearing of a melody to benefit from this additional audio information. In fact, singing seems to interfere with students' ability to remember the original stimuli. (p. 167)

While singing back can be helpful when done absolutely correctly, students need other approaches to melody memorization as well.

Even within a single college program, dictation and memorization strategies may be quite diverse. Thompson (2004) surveyed her undergraduate aural skills students and their approaches to dictation in a qualitative study. She placed her students into certain categories, based on her results:

1) the “follower,” who depends on other singers or instrumentalists for musical help;
2) the “button-pusher,” who connects notes to the sense of touch while playing an instrument, without actually hearing the sound internally;
3) the “contour-singer,” who can follow and imitate the general shape of a melody, but misses intervals and/or tonality considerations;
4) the “tonal-thinker,” who bases everything on the tonic triad and scale;
5) the “builder,” who constructs a melody interval by interval, and therefore often goes off track for an extended passage;
6) and the “pitcher,” who has absolute pitch recall.

Thompson found that some combination of approaches was the general trend for older, more experienced students, and that the “tone-builder” approach, a combination of interval and tonality approaches, represented the most successful students. She makes an interesting conclusion at the end of the study:

> Successful students like Judy and Etta in this study had already internalized sound patterns which they were easily able to recognize and label. Other students had not defined such a storehouse of tonal patterns in their memory, though they had an intuitive sense of what sounded right or wrong. In the past I have been too quick to apply

> syllables to notation before students had connected them with sound apart from notation. The perceptual way to make this connection would be to hear sound patterns with tonal relationships, label them with solfege or numbers by ear, and then discover how the patterns are notated (p. 98).

Perhaps it would be useful to extend the end of the last sentence to say "…discover how the patterns are notated by eye." In other words, the researcher is advocating storage of aurally-imaged patterns and schema. For students taking melodic dictation, that is an important asset. However, to be able to write what they have aurally understood, they need to have those imagined sounds connected to imagined notes. When students attain fluency across the two domains, they may be able to write from dictation and sing from notation with greater success.

Error detection requires the same skills as melodic dictation, assuming that those performing error detection tasks have little or no time to study the music beforehand. In both tasks students translate aural information into visual information and evaluate the accuracy of the translation. Students must determine how a sounded excerpt should look, and either discriminate between matched and unmatched notation (error detection), or produce the notation themselves (melodic dictation).

Brand and Burnsed (1981) cautioned that sight-singing and ear training skill were not found to be statistically significant predictors of error detection ability in their study of undergraduate instrumental music education majors. Correlational analyses of error detection scores and five predictor variables (number of instruments played, ensemble experience, music theory grades, sight-singing and ear training grades, and years of private instrumental instruction) revealed no significant differences. Correlations among the five predictor variables showed no significant differences either.

The results contradict the previous assertion that melodic dictation (ear training) and error detection share the same requisite skills. However, the study used semester course grades as the measure of sight-singing and ear training ability. Course grades may reflect multiple confounding variables unrelated to the musical tasks themselves. Therefore, the results of Brand and Burnsed's study taken alone may not be very useful.

Other researchers have investigated the transfer of information from the aural to the visual domain during error detection. Byo (1997) found that graduate and undergraduate music majors scored significantly higher on error detection for one-part excerpts than for two-part excerpts, and for two-part excerpts than for three-part excerpts. These results seem logical but the question remains: What causes the differences? Byo discusses the results as they relate to other questions in the study, but he does not speculate as to the cause of the results themselves. It may involve a combination of insufficient task time and students' inadequate audio-visual transference abilities. As the texture of the excerpt becomes thicker, the mind has more information to transfer before the eyes can scan for discrepancies.

Sight Singing

Sight singing is essentially the reverse of dictation. When sight singing, students view written music, imagine the sounds it represents, and reproduce those sounds vocally. In dictation, students listen to music, imagine the notation that represents that music, and reproduce that notation visually. Insights into the imagery transfer phase in either task may inform approaches to the other.

Although MacKnight's (1975) study involved instrumental music reading, her experimental group learned and used the same skills needed for sight singing. Ninety fourth-grade students were divided into an experimental group which learned new pitches through tonal patterns in context, and two control groups which learned new pitches through traditional note-by-note visual presentation. The sequence of presentation in each group merits mention here: The experimental group (1) heard a new pattern, (2) heard and saw a new pattern, and (3) heard and saw a new pattern in context. The control groups (1) saw a new note, (2) learned the letter name of the new note, and (3) played the new note on their instruments. Experimental group participants heard first and saw later; control group participants saw first and heard later.

The experimental group scored significantly higher overall on both the *Watkins-Farnum Performance Scale* (Watkins & Farnum, 1954) and the *Music Achievement Test* (Colwell, 1969). Results also showed a significant main effect of musical aptitude and a significant interaction of treatment and musical aptitude for both tests. MacKnight points out that the interaction was different for each of the tests, because of the scores of the low aptitude students in the three groups. The low aptitude students from the experimental group scored higher on the *MAP* (a written evaluation), while the low aptitude students from the control groups scored higher on the *WFPS* (a performance evaluation).

MacKnight lists several possible explanations for the discrepancy, including brief experimental period, time of day of testing, and age (too old to develop musical aptitude) of the students. The explanation most pertinent to the current discussion is "It is conceivable then that low musical aptitude students took longer to integrate their bimodal responses, and that additional training in left to right scanning of notation might improve their performance" (p. 32).

MacKnight has given educators a true gem of wisdom: students need to be shown how to process aural and visual stimuli together—especially in music, a language in which they do not necessarily converse every day. Understanding and becoming conversant in music aurally, visually, and aural-visually are three distinct processes. They involve skills that must be practiced in different ways, with different exercises. Although the three processes are intimately related, they cannot be assumed to be identical.

McClung (2001) surveyed 2,115 senior All-State Chorus members from Arkansas, Alabama, Georgia, Louisiana, Tennessee, and Mississippi with one question: "In which sight-singing system have you received the most instruction?" Results of the survey showed the following participant responses: melody pitch numbers (58%), movable-do (19%), neutral syllables (13%), other (6%), and fixed-do (4%).

The top two response categories (77% combined) both incorporate tonal context into sight-singing, lending support to the idea that students benefit from strategies that help them organize theoretical concepts cognitively. That organization may help them directly in musical applications like sight-singing and melodic dictation. When students have developed a fluent connection between a dynamic music theory imagery construct and accurate musical perception, both sight-singing and dictation become quite manageable.

Manipulable mental constructs and their applications to musical tasks may be tied to greater overall cognitive development and skills as well. Harrison (1990) explored connections between a host of variables (SAT math and verbal components, GPA, gender, scores on Schleuter's *Musical Aptitude Profile* (out of print), piano study, experience on major instrument, experience on multiple instruments, and total experience on all instruments) and the grades of entering college music majors in four parts of music theory coursework (ear training, sight singing, keyboard harmony, and written work). For both the first and second semesters of study, SAT math scores were the best predictors of written theory work grades, while musical experience and aptitude were the best predictors of sight singing and keyboard harmony grades. Ear-training grades drew a split decision: SAT math scores were the best predictors for the first semester, musical experience and aptitude for the second.

As previously discussed, course grades can create spurious results. However, grades in this study were fairly specific to given skills, rather than a general "music theory" grade, for instance. The link between ear training grades, written theory work, and SAT math scores creates several questions: What skills do high SAT math scores indicate? What skills do ear training and written theory work require? What skills common to all three might successful students use? SAT math tasks and written theory work both require a solid knowledge base and the ability to visually organize bits of knowledge into meaningful wholes. Ear training (especially dictation) requires the same skills, except that the information to be organized is presented aurally, not visually.

In math operations, a situation parallel to dictation would be a verbal account of a set of data and the operations to be performed on it. Test-takers would have to process the information aurally and transfer it to written form, while or after solving it. Perhaps some students in Harrison's study used these skills in both areas. It must be noted, however, that SAT math scores were not the greatest predictor of ear training grades in the second semester.

Because the mind has to make aural-visual transfers in dictation and error detection, and visual-aural transfers in sight singing, one would think that additional interfering information in either form would hinder the communication. However, Halfpenny, Ho, Kurosawa, and Wollner (2003) found evidence to the contrary: subjects in their study sang equally accurately with and without interfering music played in their headphones. All subjects sang in random order the same two melodies, one with interference and one without. In addition to analysis of pitch and rhythm errors, subjects were rated by a panel of experts on continuity, fluency, and overall quality. Although performances in the non-interference condition received statistically higher "overall quality" ratings, none of the other dependent variables revealed statistically significant results.

Halfpenny, et al.'s study suggests that sight singers can ignore interfering aural information. Although 65% of the participants said that the interference condition made inner hearing more difficult, they obviously were able to overcome that challenge. Dictation-takers, of course, are dealing with a slightly different challenge. They are not physically producing music, but rather trying to remember aural stimuli, in the midst of interference of the subsequent measures of melody. Focused exercises isolating the transfer from visual domain to aural, and vice-versa, might help students improve their performance and efficiency in both tasks.

Killian and Henry (2005) explored the influence of study time on sight singing accuracy. Subjects were divided into three groups: low-, medium-, and high-level sight singers. Each subject sang two melodies, one with 30 seconds of preparation time and one with no preparation time. Medium- and high-level singers performed significantly more accurately under the preparation condition, but low-level singers did not. The researchers suggested that the low-level singers may not have had the skills needed to take advantage of preparation time. In other words, they had not developed enough facility in transferring notation into aural imagination. This suggests that students may need instructional guidance in developing this facility for accuracy and efficiency.

Killian and Henry also reported that those who sang aloud performed better than those that did not. They suggest that singing out loud during the preparation time helped singers to reinforce sections of the melody they had mastered and to isolate and correct more challenging sections. Unfortunately, most sight singing situations do not allow for individual audible preparation, and most dictation situations do not allow audible singing at all. Students have to develop those same skills internally. They should be prepared to check and re-check their mastery of the material without making a sound.

Aural skills and sight singing

Researchers investigating various musical tasks in combination illustrate their similarities and differences, thus clarifying thought about overarching skills like aural-visual information transfer. Grutzmacher (1987) compared pretest and posttest scores of forty 5th- and 6th-grade beginning instrumentalists on the *Iowa Test of Music Literacy* (Gordon, 1991) and a researcher-designed sight reading test. The experimental group was taught using tonal patterns reinforced with harmonization and vocalization. The control group was taught using traditional note-by-note music learning. The experimental group scored significantly higher than the control group on the tests pertaining to melodic sight reading and aural identification of major and minor tonalities, but not on comparisons of aural and visual perception of tonal patterns. The tonal pattern approach with reinforcing exercises improved the connections between aural and visual tasks, but perhaps lacked specific training in the mental coordination of aural and visual images of the same patterns.

Specific training is indeed an important component of successful synthesis of musical skills. Bluestine (2007) compared four groups of elementary school students on sight reading and sight singing skills, after a period of differentiated instruction in beginning-level tonal music reading. The unique instruction for each of the four groups was (1) singing whole patterns, (2) singing individual pitches, (3) singing whole patterns, then individual pitches, (4) singing individual pitches, then whole patterns. Bluestine found no significant differences among any groups. A root cause may have been overall lack of familiarity and training in these two skills.

Norris (2003) examined relationships between college freshmen scores on melodic dictation and sight singing exams, spanning the course of a semester. He questioned whether melodic dictation or sight singing pretest scores could predict sight singing posttest scores. The same melodic dictation and sight singing examples were used for both pretest and posttest and the scores were calculated by two experienced sight singing teachers, yielding reliability correlations of .99 (dictation) and .97 (sight singing).

Norris found results that could be expected: statistically significant improvements in scores in both tasks from pretest to posttest over the course of the semester and a decrease in variance within each task. He also found that the correlation between dictation and sight singing scores did not change much from pretest to posttest ($r = .62$ and $r = .57$ respectively) and that pretest sight singing was the only strong predictor of posttest sight singing ($r = .68$).

He concluded that “Relationships between sight singing and melodic dictation (.62 and .57, pre- and posttest, respectively) are not as large as might be expected, but their moderate strengths and significance support the likelihood that sight singing and melodic dictation skills are related” (p. 48). He noted that additional research needs to be undertaken “to continue the search for more effective instructional strategies and contexts” (p. 51).

Killian (1991) compared junior high students’ scores on sight singing (notation), sight singing (solfeggio), and error detection. Based on their sight singing (notation) scores, she divided the students into three categories: high, medium, and low scorers. For high and medium scorers, there were no significant differences in scores for the three tasks. Low scorers scored significantly higher on error detection than on either of the sight singing tasks. Killian concluded that chorus teachers should consider using error detection training as a way to build inexperienced sight singers' confidence and to help them master skills that can be transferred to successful sight singing later. She also cited Brand and Burnsed's aforementioned warning that error detection skills and sight singing skills may not be the same. Skills for each task may require distinct guidance, followed by demonstration of connections across task skill sets.

Larson (1977) ran correlations among error detection, melodic dictation, and sight singing scores of junior and senior undergraduate music majors. Each of the three task types included diatonic, chromatic, and atonal examples. All correlations were significant except those between atonal sight singing and diatonic error dictation, and between atonal sight singing and chromatic error detection. Larson suggested that the difficulty of the atonal sight singing tasks may have skewed the results for the two non-significant findings. He noted that "relationships generally were higher between error detection and dictation scores than between error detection and sight singing scores" (p. 268).

Larson suggested that ear training instructors should "regard dictation as an important contributory means of aiding in the prognosis and development of more purely functional aural-visual discriminatory abilities rather than as an important terminal competency" (p. 270). Perhaps dictation teachers succumb too often to the "teach to the test" method, focusing more on students' ability to get the right answers than on the broad skills students can learn and apply to other tasks, musical and nonmusical. Although dictation tasks effectively measure aural-visual skills, they are not the only measure of those skills nor do they measure the skills completely (Gordon, 2007). Larson's conclusion suggests that if students struggle with melodic dictation, they may struggle with aural-visual skills application in general.

The applicability of aural-visual transference skill development extends across multiple disciplines of musical study. Sheldon (1998) gave pretests and posttests to two groups of undergraduate music education majors taking instrumental methods courses, to observe effects of contextual sight singing and aural skills training on error detection skill for conductors. Using solfeggio, the experimental group sang tonic and dominant patterns, single-line excerpts in unison, and multiple-line excerpts as a group during the course of the semester. The control group received normal methods-course instruction. The pretest and posttest consisted of excerpts of band literature with added errors in pitch and rhythm. Scoring of the tests incorporated both correct error detection responses and incorrect responses ("detecting" an error when none was present).

Analyses of pretests showed no significant main effect of group assignment on either correct responses or incorrect responses, but analyses of posttests revealed that the experimental group responded with significantly more accuracy in both. Sheldon concluded that the sight singing and aural skills training may have improved the experimental group's error detection ability. She also found a significant main effect of error type across both categories of responses on both tests: subjects detected rhythm errors more accurately than pitch errors.

Sheldon concluded that the semester-long training may not have been long enough to improve pitch error detection abilities as much as rhythm error detection abilities. She warned against breaking error detection tasks into isolated components in future studies, because it could ruin the practicality of results for conductors. The ultimate importance of aural-visual processing research depends on its applicability to musical tasks.

The studies of imagery in musical applications outlined in this section suggest that skilled imagery constructs can be useful in completing musical tasks. Musicians can develop manipulable imagery schema based on accumulated training and experience. Application of musical skilled imagery may follow hierarchical guidelines, imposed by the musician, to maximize the organization of various musical aspects of a given stimulus. Temporal length and textural complexity of a musical stimulus appear to affect processing times for musical imagery, but efficiency can be increased through "chunking." Aural skills students may employ a variety of strategies in tackling musical tasks, some of which may actually hinder their success. Imagery skills used for error detection, sight singing, and aural skills tasks may be interrelated, and may be most effective when they are purposefully taught, cultivated, and applied in the classroom.

Conclusions

Aural perception in music depends on multiple factors, including musical experience, percept complexity, and aural imagery scanning efficiency. These factors in combination play an important role in taking melodic dictation. Dictation is complicated by time constraints, limited repetitions of the melody, and students' short-term memory overload. Creative approaches to improving short-term memory during dictation may increase efficiency and accuracy for the overall task.

Visual imagery has been shown to enable "chunking" for quick identification and retrieval of information, back-and-forth scanning for efficient task completion, and conceptual stabilization for assimilation of new information. Researchers have also illustrated the following information about aural and visual channels: combining them during processing tasks may help keep one or the other channel clear, dynamic visual imagery may aid listeners as they track sequential aural stimuli, and some advanced musicians use visual imagery to encode melodic information.

Dynamic visual imagery is the basis of skilled imagery, employed in specific tasks within many disciplines. Skilled imagery generally involves a working, manipulable conceptual framework, to which new information can be attached and processed. The basic framework is developed in the mind through sensory experience and illustrative examples. Similar to a computer-based document template, it stays the same in its original form, while allowing changes to create variations of itself for various user needs.

Studies of imagery applications to musical tasks have shown similarities and differences among various aural skills and sight singing skills. The most important similarity for this discussion is the connection between visual and aural representation of music in, for example, both melodic dictation and sight singing. Melodic dictation requires accurate hearing and transfer to visual output, while sight singing requires accurate seeing and transfer to aural output. Research subjects have reported multiple strategies to accomplish these transfer tasks, and experimental studies have demonstrated that some strategies work better than others, for certain subjects. Students likely will benefit from the widest array possible of viable strategies.

These studies set the stage for exploration of the use of visual imagery in melodic dictation. It seems that some musicians already use this skill successfully (Mikumo, 1994). Others may choose to use it if it can be shown to be beneficial to melodic dictation scores, and if it can be taught effectively and efficiently.

Implications for Music Education

Visual imagery is employed in numerous skilled tasks in other disciplines, and has been shown to be compatible with aural stimuli in those tasks. However, it may need to be specifically taught to students as a possible strategy. For prose listening comprehension, Gambrell and Bales (1987) advocate the "think-aloud procedure," in which a teacher describes in detail the mental imaging that could accompany a spoken passage. This serves as a model for visual imaging that students can learn to use on their own (p. 150). This may be necessary in the aural skills classroom for those who have never used visual imagery as a memory strategy in melodic dictation.

Discoveries about communication between the aural and visual domains of the brain during melodic dictation tasks will guide music educators at all levels, in teaching their students how to develop that communication. Olson (1978) studied first graders' abilities to match aural stimuli (short melodies) to visual stimuli (contour maps), visual stimuli to aural stimuli, and aural stimuli to aural stimuli. Students scored significantly higher on aural-aural matching tasks than on aural-visual and visual-aural. First graders apparently cannot skillfully translate musical information between aural and visual domains. By the time they are college music majors they likely will be able, but only if someone has trained them to do so.

Livingston and Ackman (2003) surveyed fifty college music students to find out which factors were most important in their theory training during high school, and to compare the results to a parallel survey conducted in 1981. They reported raw percentages from both surveys, noting a sharp increase in two categories. High school band (from 10.3% to 40%) and high school theory class (from 13.7% to 32%) were the most important influences in their theory development. The authors cite implementation of the National Standards, incorporation of theory concepts into band method books, and increased focus on academics in high schools as possible reasons for these increases. High school music teachers and directors are therefore key influences in the development of the skills discussed in the current study.

While the survey indicated the importance of high school band and theory classes for college theory readiness, the meaning of results of studies in this field for teachers of younger students is equally important. Melodic dictation represents correct perception, comprehension, and notation of sound. Efficient and accurate melodic memory is essential to success in accomplishing these tasks. The basis for these skills can begin with a child's very first musical experience and continue throughout schooling and beyond. Educators should continue to explore ways to reinforce connections between sound and symbol in the music classroom.

Additional questions for further study have been raised throughout this chapter. As new studies reveal specific information about cognitive processing during melodic dictation, music educators will understand how best to approach teaching the requisite skills for successful dictation. If these skills can be fostered at an early age, and throughout a child's music education, melodic dictation in college music classes should be a manageable task.

Purpose of the Current Study

The purpose of this study was to determine how melodic memory is affected by aural distractions when melodic stimuli are presented both visually and aurally, as compared to aurally only. The null hypothesis was that melodic memory scores are the same, whether melodies are heard and seen, or heard only. The specific research questions were:

1. Does visual reinforcement of aural stimuli increase subjects' aural memory for those stimuli, specifically regarding pitch?
2. Does visual reinforcement of aural stimuli increase subjects' aural memory for those stimuli, specifically regarding rhythm?

Limitations

To maximize usefulness of results, this study was limited in several ways. Participants were college undergraduate and graduate music majors in a single geographical area, all of whom had passed the first two semesters of undergraduate music theory. Subjects who had not passed the first two semesters of college theory were excluded, because they may not have possessed the basic musical skills required to complete the experimental tasks, thus potentially skewing results. Students with absolute pitch were also excluded, to avoid the potential of their reliance on this relatively rare skill to complete the experimental tasks.

Using such a specific and limited sample should not hinder generalizability of results, as subjects were tested against themselves. In fact, the main questions of this study centered on the way a single brain functions under two conditions. The fact that the sample was drawn from a single institution helps to ensure that all participants possessed sufficient baseline musical skills to complete experimental tasks.

Definitions

For purposes of the current study, the following definitions are used throughout:

aural imagery: production, retrieval, or manipulation of imaginary sound

aural perception: mental registry of aural stimuli that are physically present

aural skills: the ability to perceive and make musical decisions about aural stimuli

melodic dictation: writing a melody that has been heard several times

visual imagery: production, retrieval, or manipulation of imaginary sight

visual perception: mental registry of visual stimuli that are physically present

CHAPTER 3

METHOD

Design

Participants completed two sections of a melodic memory test in a 2x2 factorial experimental research design. The two independent variables were presentation format of experimental task (aural-only or aural-visual), and answer discrepancy type of the incorrect answer choice in each test item (pitch difference or rhythm difference). The dependent variable was experimental test scores.

Participants

The participants (*n*=41) in this study were undergraduate and graduate music majors. The only prerequisite to participation was successful completion of the first two semesters of college music theory, to insure that all participants possessed the basic musical skills required to complete the experimental tasks. Students who possess absolute pitch were excluded, to avoid the potential of their reliance on this relatively rare skill to complete the experimental tasks.

Potential participants were contacted through brief visits by the researcher to sophomore music theory classes and graduate music theory seminars, to insure that all contacted students had already passed the first two semesters of theory. The researcher gave a brief outline of the project and distributed voluntary consent forms.

Students were assured that their decision to participate held no implications for their evaluation in class, and that all results would be completely anonymous.

Experimental Instrument

The test instrument for this experiment consisted of two audio-video-interleave (AVI) files, created by the researcher using Windows Movie Maker (2009). The files presented six test items each, corresponding to the two test formats: (1) aural-only, and (2) aural-visual. Each test item consisted of a sequence of recorded verbal instructions, aural establishment of key, target melody, distraction melody, and two answer choice melodies. The aural-visual test included notation of the target melody, shown on the screen just after the final note of the target was sounded. The screen remained black throughout the rest of the item. In the aural-only test, the screen remained black throughout. Both tests contained an equal mix of answer discrepancies: three items of each test contained a pitch change in the incorrect answer choice, and three items contained a rhythm change.

Target and distraction melodies were diatonic, in the major mode, representing key signatures with a maximum of two sharps or flats (see Appendixes A and B). A consistent level of smooth melodic contour was maintained across items through the avoidance of any leaps larger than a perfect fifth. The inherent harmonic underpinning of all target and distraction melodies centered on the tonic and dominant chords of the established key for the item. All of these melodic characteristics were intended to minimize the need for complex theoretical knowledge and advanced aural skills experiences, thus focusing attention on memory for musical materials familiar to all participants.

Target and distraction melodies were created by the researcher, in consultation with two professors of music theory at the university where the experimental work was completed. As a result of these consultations, draft melodies were edited to exhibit greater clarity in rhythmic organization and beaming. Optimal legibility was necessary because participants would have only a brief view of the melodies.

All melodies were generated by an *Encore* (2009) piano patch with *Encore* notation. Piano timbre was chosen because it is the most universal of instruments in university music curricula. Computer-generated audio was chosen so that target melodies and answer choices would sound exactly the same, except for the deliberate pitch or rhythm discrepancy in the incorrect answer choice of each example.

Norris (2000) noted important factors to be considered in tests of tonal memory. He specifically pointed out the difference between recognition versions and reproduction versions of these tests. Recognition tonal memory tests require subjects to decide whether playback answer choices match a stimulus tonal sequence; reproduction tests require students to recreate, usually by singing, the stimulus tonal sequence.

Several potentially confounding factors must be considered when giving reproduction tonal memory tests (e.g., singing ability and confidence, changing voices, and subjects' comfort level with performance the task.) On the other hand, recognition tests may lend themselves to guesswork when matching. However, if answer choices such as “neither” or “none of the above” are included, guesswork can be effectively limited. Based on Norris's conclusions and careful consideration of the purpose, questions, and design of this study, it was determined that a recognition test, with two answer choices and a third choice of "neither," was the most reasonable way to proceed.

Target melodies were identical across both tests, with the exception of a practice melody at the beginning of each, used to familiarize the participants with the presentation format (see Appendixes A and B). As shown in Appendixes A and B, the six melodies were presented in different order on each of the two tests, to avoid pattern recognition or transference of any memorized answers from one test to the other. To avoid presentation order bias, half the participants received the aural-only examples first; the other half received the aural-visual examples first. Pilot test data had revealed that use of identical melodies yielded no practice effect. Pilot test participants had actually performed better overall on the first test, than on the second (m = 3.75, m = 3.33, respectively). It was therefore determined that identical melodies, in different presentation orders, would be used in formal data collection as well.

Pilot Test

The pilot test was administered to twelve musicians of comparable dictation ability. Validity and reliability of test questions, clarity of directions and procedures, and overall test difficulty were found to be quite consistent and appropriate, based on examination of the resultant data and interviews with the pilot participants. Table 1 displays the calculated difficulty index for each test item. Because there were three answer choices, chance difficulty index for each question was 0.33. All difficulty indices were above chance, and well below a potential ceiling effect. Therefore, no major adjustments to the testing materials were made before data collection began, although several stages of revisions had already been made before the pilot.

Table 1

Test Item Difficulty Indices

Test item	Item Difficulty Index
Melody 1	0.71
Melody 2	0.58
Melody 3	0.54
Melody 4	0.42
Melody 5	0.71
Melody 6	0.58

Procedures

Each participant completed the experimental tasks individually, in a small, quiet office. Slight extraneous noise in the office was eliminated with Sony MDR-NC7 Noise Canceling Headphones during the trials. The instrument for this study was completely self-contained. Participants used a Dell Inspiron E1405 laptop computer to listen to and view the two AVI programs.

Before each test began, student status (undergraduate or graduate) and major instrument data were collected, for additional exploratory analysis. Data were also recorded of the order in which each participant took the two test formats. Participants then read the following directions respectively.

Aural presentation:

> Each item in this program begins with a I-V7-I chord progression to establish the key. After the chord progression, listen to the first melodic phrase and try to remember it. After a short pause, you will hear a distraction melodic phrase. Focus on remembering the original phrase. You will then hear two answer choices. Circle the answer choice that matches the original exactly. If neither choice matches the original exactly, circle "neither."

Aural and visual presentation:

> Each item in this program begins with a I-V7-I chord progression to establish the key. After the chord progression, listen to the first melodic phrase, view its notation on the screen immediately following, and try to remember it. After the notation disappears, you will hear a distraction melodic phrase. Focus on remembering the original phrase. You will then hear two answer choices. Circle the answer choice that matches the original exactly. If neither choice matches the original exactly, circle "neither."

After reading the directions for themselves, participants were asked to summarize the directions verbally to the researcher, to ensure understanding of the procedures.

On both tests, before each test item, the key of the stimulus was established aurally, using the chords I-V7-I, by the same *Encore* piano patch as described previously for the target melodies. After each target melody, the notation of the melody appeared on the screen in the aural-visual test. In the aural-only test, the screen remained black during this same time interval following the target melody. The ensuing distraction fragment was essentially the melodic consequent of the stimulus. In other words, the target stimulus was the musical "question" and the distraction was the musical "answer." This format was meant to represent the common dictation approach of focusing on processing only the first part of a melody first, while the rest of the melody continues to sound. After the distraction fragment, participants heard each answer choice, decided if either choice matched the target melody, and marked their answer sheets. Five seconds after the second answer choice was heard, the announcement of the next example began.

Anonymity Measures

To ensure anonymity during procedures and analysis, each test was marked on the back with a four-digit number. A small card with matching number was affixed on top of this participant number. Participants detached the card to take with them, leaving the matching participant number on the test. As they finished the test, participants placed their answer sheets into a large envelope themselves. Once all data had been calculated, a mass e-mail was sent to all participants containing a spreadsheet file of all data, by participant number. Participants' e-mail addresses were listed only in the blind carbon copy "recipients" box, to maintain privacy. Participants were able to locate their specific data by the number on their card.

Analysis

The independent variables in this study were presentation format (aural or aural-visual), and answer discrepancy type (pitch or rhythm) of the incorrect answer choice in each test item. The dependent variable was experimental test scores. A 2x2 factorial ANOVA, with presentation format and answer discrepancy type as the factors, was used to analyze, at the .05 alpha level, the collected data from each participant. The 2x2 ANOVA is a statistical comparison of means that detects a significant main effect of either factor, and any interaction between factors. In this study, it would reveal whether participants performed better under either of the two presentation formats, either of the two discrepancy types, or some specific combination of these two factors.

Each participant completed the experimental tasks under all four combinations of factors (see Figure 1). Therefore, data represent 164 total experimental observations. Data exhibit a high level of balance due to the experimental design: participants are compared to themselves under the given conditions, and test items are identical for both tests.

Figure 1. 2x2 ANOVA cells illustrating experimental design.

Aural-Only Presentation Rhythm Discrepancies	Aural-Visual Presentation Rhythm Discrepancies
Aural-Only Presentation Pitch Discrepancies	Aural-Visual Presentation Pitch Discrepancies

Additional Analyses

Test Order

Half the participants took the aural-only test first; the other half took the aural-visual test first. The purpose of administering the tests in this manner was to mitigate a potential practice effect, due to the identical items on both test formats. For example, if all participants took the aural-visual format first, and scored higher on the aural-only format, it could be argued that their familiarity with the test items helped them to score higher on the second format, regardless of the format itself.

In addition to mitigating the potentially confounding effect described just described, equal administration of the two test orders enabled specific statistical analysis of any relationship between test order and participant scores. A *t*-test, by test order, was conducted to detect any significant differences in test scores, based on the order in which participants completed the two test formats. Because the scores of interest for this test were aural-only and aural-visual scores intact, parsed data were collapsed into 82 total observations (two observations per participant).

Major Instrument

Data were collected on the major instrument of each participant. For the purposes of exploratory study, participants were grouped into one of four instrumental categories: vocal (n=11), keyboard (n=7), wind and percussion (n=18), or stringed instrument (n=5). Because of the small, unequal sample sizes created by this division, a non-parametric Kruskal-Wallis H test was used to detect any differences in overall scores, pitch scores, rhythm scores, aural-only scores, and aural-visual scores, by major instrument. Circumstances were far from ideal for this particular test, and results, presented in the next chapter, should be approached with caution. The purpose of this analysis was exploratory only. A dramatic difference between any two or more groups on any of the score categories might suggest specific differences in instrumental training, thus indicating possibilities for future research.

Student Status

The participants in this study were undergraduate and graduate students. Pilot data had demonstrated that neither age nor years of experience created a ceiling effect in data distribution (see Table 1). Therefore, it was determined that both groups of students could be pooled together without skewing data. To test this assumption, student status data (undergraduate or graduate) from formal data collection were incorporated in a 2x2x2 analysis, with test format and answer discrepancy as the other two independent variables. This analysis would reveal any statistically significant differences between graduate and undergraduate students on each of the four data cells shown in Figure 1.

If a significant difference were to be found in any of the test score categories, results might suggest particular differences in length of musical training, level of aural skills development, level of fluency with notation, efficiency and accuracy of melodic memory, or some combination of these factors. While a comparison of these two groups of students was not a specific focus of this study, it was considered useful for informing potential future research.

CHAPTER 4

RESULTS

Descriptive Statistics

Table 2 displays descriptive statistics of the experimental data collected.

Table 2

Descriptive Statistics

Test Format	Discrepancy	Statistics
Aural-Visual	Pitch	*m* = 1.39, *sd* = 0.97
Aural-Visual	Rhythm	*m* = 1.71, *sd* = 0.75
Aural-Only	Pitch	*m* = 1.51, *sd* = 1.03
Aural-Only	Rhythm	*m* = 1.49, *sd* = 0.95

Analysis of Variance Assumptions

Use of the analysis of variance (ANOVA) statistical test requires that four assumptions about the data be met: randomness, independence, normality, and homogeneity of variance. The first two assumptions were met automatically, through the experimental design. Because all participants completed the experimental tasks in all four data cells, their scores in each cell were essentially compared to their own scores in the other cells. Therefore, randomness was satisfied by the fact that there was no group assignment necessary. Independence was satisfied in that each participant had a separate score in each cell. In other words, it was as if each participant took four tests, representing each combination of conditions established by the independent variables.

Normality was checked by reviewing skewness and kurtosis measures for the data. Both fell within acceptable ranges. Homogeneity of variance was tested with Levene's Test of Equality of Error Variances, and was also found to be acceptable. It was determined that the ANOVA could be effectively employed for data analysis.

ANOVA Results of Presentation Format by Answer Discrepancy

As Table 3 shows, there was no significant difference in test scores based on presentation format. There was also no significant difference in test scores based on answer discrepancy, and no interaction between factors.

Table 3

2x2 ANOVA results by test format and discrepancy type

Source	Type III Sum of Squares	*Df*	Mean Square	*F*	Sig.
Corrected Model	2.171	3	.724	.835	.477
Intercept	381.098	1	381.098	439.522	.000
Test Format	.098	1	.098	.113	.738
Discrepancy Type	.878	1	.878	1.013	.316
Format * Discrepancy	1.195	1	1.195	1.378	.242
Error	138.732	160	.867		
Total	522.000	164			
Corrected Total	140.902	163			

Additional Analyses Results

Test Order

Half the participants took the aural-only test first; the other half took the aural-visual test first. A t-test was employed to determine any significant difference between test scores, based on test order. Participants' first test scores were compared to their second test scores. As evidenced in Table 4, there was no significant difference in test scores, based on test order.

Table 4

T-test results by test order

	Statistic	*df1*	*df2*	Sig.
t-test	.492	1	79.742	.485

Major instrument

Because of the small, uneven groups created by categorizing participants according to major instrument, nonparametric tests were needed to analyze these data. The Kruskal-Wallis *H* test was conducted to detect any differences in overall scores, pitch scores, rhythm scores, aural-only scores, and aural-visual scores, by major instrument. As mentioned previously, these results should be approached with caution. They were included in this study for exploratory purposes. As shown in Table 5, no significant differences were found among any groups on any of the test score categories.

Table 5

Kruskal-Wallis H test results by major instrument on overall scores, pitch scores, rhythm scores, aural-only scores, and aural-visual scores

	Overall Scores	Pitch Scores	Rhythm Scores	Aural-Only Scores	Aural-Visual Scores
Chi-Square	2.730	6.586	.594	1.105	4.102
Df	3	3	3	3	3
Asymp. Sig.	.435	.086	.898	.776	.251

Student status

The 2x2x2 ANOVA compared undergraduate participants to graduate participants, based on test format and answer discrepancy type. Because of unequal sample sizes (*n*=22, *n*=19, respectively), three undergraduate participants' data were randomly removed from the data set. As Table 6 shows, there were no significant differences in test scores, based on student status, test format, and answer discrepancy. There also were no interactions among any of the factors.

Table 6

2x2x2 ANOVA results by student status, test format, and discrepancy type

Source	Type III Sum of Squares	*df*	Mean Square	*F*	Sig.
Corrected Model	2.520	7	.360	.394	.905
Intercept	339.007	1	339.007	371.306	.000
Student Status	.007	1	.007	.007	.932
Test Format	.059	1	.059	.065	.799
Discrepancy Type	1.480	1	1.480	1.621	.205
Student Status * Test Format	.007	1	.007	.007	.932
Student Status * Answer Discrepancy	.007	1	.007	.007	.932
Test Format * Answer Discrepancy	.796	1	.796	.872	.352
Student Status * Test Format * Answer Discrepancy	.164	1	.164	.180	.672
Error	131.474	144	.913		
Total	473.000	152			
Corrected Total	133.993	151			

CHAPTER 5

DISCUSSION

Summary

Introduction and Purpose

Melodic dictation tests the synthesis of aural and visual cognitive skills in music. Students taking dictation typically are challenged by lack of time, lack of repetitions of the target melody, and lack of a silent environment in which to process percepts. These challenges are underscored by the human inability to remember a whole melody (Miller, 1956), or to remember even part of a melody while the rest of the melody continues to sound. When focusing on a portion of a target melody, quick transfer of aural percepts to visual imagery may increase melodic memory in the midst of this subsequent aural interference (Cassidy, 2001; Center, Freeman, Robertson, & Outhred, 1999; Mikumo, 1994; Segal & Fusella, 1970). The purpose of this study was to determine whether melodic memory is less affected by aural distractions when melodic stimuli are presented both visually and aurally, as compared to aurally only.

Method

The participants in this study were undergraduate and graduate music majors (*n*=41) at a large northeastern university. All participants had passed the first two semesters of college-level music theory, and none had perfect pitch. Participants were contacted through upper-level theory courses and graduate courses at the university. Data collection and analysis were kept strictly confidential through participant number identification.

This study was based on a 2x2 factorial experimental design. The two independent variables were test format (aural-only or aural-visual) and answer discrepancy type (pitch or rhythm) of the incorrect answer choice for each item. The dependent variable was participant test scores. The experimental instrument consisted of two AVI files, created by the researcher using Windows Movie Maker (2009). One file was an aural-only melodic memory recognition test; the other was an aural-visual melodic memory test. Both tests contained the same six melodies, presented in different order to avoid a practice effect in resultant scores. To further mitigate any test order practice effect, half the participants took the aural-only test first, and the other half took the aural-visual test first.

Tests were completely self-contained. Participants read instructions for each test and proceeded through a sequence of key establishment, target melody, distraction melody, and answer choices for each test item. The only difference between the two test formats was the appearance of the notation representing each target melody—just after the conclusion of the target melody—on the aural-visual test. During the aural-only test, the screen remained black throughout. Both tests contained an equal balance of answer discrepancies: three items of each test contained a pitch change in the incorrect answer choice, and three items contained a rhythm change.

A 2x2 ANOVA, with test format and answer discrepancy as the factors, was conducted to determine any statistically significant differences in test scores based on either factor, and any interaction among factors. Additional analyses were conducted to determine any effects of test administration order, major instrument of participants, or student status (undergraduate or graduate), on test scores.

Results

The 2x2 ANOVA, with test format and answer discrepancy type as main factors, revealed no significant differences in test scores by either factor. There was also no interaction between factors. In the additional analyses, a *t*-test revealed no differences in test scores based on test administration order. Kruskal-Wallis *H*-test results revealed no significant differences in overall scores, pitch scores, rhythm scores, aural-only scores, or aural-visual scores, by major instrument. A 2x2x2 ANOVA, with student status, test format, and discrepancy type as main factors, revealed no significant differences in test scores on any factor. There was also no interaction among any of the three factors.

Conclusions

The central purpose of this study was to determine how pitch and rhythm aspects of melodic memory are affected by aural distractions when melodic stimuli are presented both visually and aurally, as compared to aurally only. Based on the results of this study, visual reinforcement—as presented in this experiment—does not significantly affect aural memory for melodic stimuli in the midst of aural distractions, in terms of either rhythm or pitch.

Melodic stimuli were identical across both test formats, and tests were taken consecutively in one sitting. To avoid a practice effect in resultant data, the presentation order of test items was different for each of the two test formats. Half the participants took the aural-only test first; the other half took the aural-visual test first. To evaluate the efficacy of this administration procedure, a *t*-test was conducted to determine any significant difference in test scores, based on test order. The null hypothesis was that there is no significant difference in test scores based on test order. Results of the *t*-test, at alpha level of .05, suggest that the null hypothesis cannot be rejected. Based on the results of this study, the six identical melodies played in different order on both testing formats yielded no practice effect, even when experimental tasks were completed consecutively in one sitting. The identical melodies yielded no confounding practice effect among participants.

Data were collected of all participants' major instruments. For exploratory purposes, participants were grouped into four categories: vocal, keyboard, wind and percussion, and strings. Resultant groups were compared, using the Kruskal-Wallis *H* test, on overall scores, pitch scores, rhythm scores, aural-only scores, and aural-visual scores. The null hypothesis was that there is no significant difference in any of these test scores, based on major instrument. Results of the Kruskal-Wallis *H* Test, at alpha level of .05, suggest that the null hypothesis cannot be rejected. The reader is reminded that these particular results should be approached with caution, because of small, unequal sample sizes created by instrumental categorization. Based on these results, type of instrument studied has no significant effect on melodic memory scores.

Undergraduate and graduate students were pooled together in this study, because the pilot study had revealed no ceiling effect based on age or musical experience. For exploratory purposes, undergraduates and graduates were compared on the test format and answer discrepancy factors described previously, using a 2x2x2 ANOVA. The null hypothesis was that there are no significant differences in test scores by student status. Results of the 2x2x2 ANOVA, at alpha level of .05, suggest that the null hypothesis cannot be rejected. Undergraduate and graduate students did not show significant score differences on either factor, nor was there any interaction between factors. Graduate students' additional formal musical study did not contribute to higher melodic memory test scores in this study.

Discussion and Recommendations

Background of Study

The foundation for this study derived from a practical problem in high school and collegiate music education: poor student performance in melodic dictation. Supported by extant literature across multiple disciplines, the problem was further focused. Dictation students are likely not able to remember a whole melody at once, or even a portion of the melody, given the aurally distracting nature of the sounding of the rest of the melody (Miller, 1956). Synthesis of previous studies suggested that a visual snapshot of the perceived melodic information might help mitigate the ensuing aural distraction (Cassidy, 2001; Center, Freeman, Robertson, & Outhred, 1999; Mikumo, 1994; Segal & Fusella, 1970). If dictation students could transfer aural percepts into visual imagination, they might be able to remember them better, in the midst of subsequent aural percepts.

Experimental Design

Accurate measurement of imagery is still relatively undeveloped. While fMRI technology sheds light on which parts of the brain are active during particular musical tasks, it cannot track which strategies subjects employ to complete those tasks, or to what extent (Morrison, Demorest, Aylward, Cramer, & Maravilla, 2003). Self-reporting is a common approach to studies of imagery, but is prone to inaccuracy, due to confounding variables such as faulty memory, exaggeration, and desire to report what researchers are seeking (Burton, 2003).

Because of these difficulties, it was determined that the most effective approach to testing the perception assertions just presented would be overt visual presentation of target melodies serving as a proxy for visual imagination of those melodies. Participants actually saw the notation representing the percepts, rather than being asked to imagine it. This aspect of the design was intended to avoid the potential confounding variables presented by self-reporting, and to provide all participants with consistent, correct visual information for testing purposes.

This study was designed to eliminate as many confounding variables as possible, while maintaining applicability to real teaching and learning situations. Overt presentation of target melodies increased the effectiveness of experimental procedures, but may not have served successfully as a proxy for participants' own imaginative faculties. As experimental technologies are improved, future research may involve more specific and precise analyses of these imaginative faculties in relation to music cognition and memory tasks.

Results

Informal discussion with participants, colleagues, and students yielded the general opinion that visual reinforcement of the melodies would no doubt produce higher scores. This was not the case, according to the results of the study. Some participants may have found the visual input distracting. Some, in fact, reported being discouraged by the fact that they could not “see their way through” the entire target before it disappeared from the screen. Some participants may be “aural learners,” better able to focus during the memorization stage without any visual input at all (Beheshti, 2009). Future research might involve pairing melodic memory tests with a learning modalities survey, enabling score analysis based on modality groups.

If overt, correct visual reinforcement of target melodies does not improve aural memory for those melodies, it seems unlikely that self-generated visual imagery would either. During dictation, if students create their own visual images of aural percepts, they run the risk of embedding pitch or rhythm errors in the image. Even if they transfer all information correctly to the visual image, they must be able to maintain the image strongly enough to remember it later. A correct, clear, sizable image on the computer screen did not accomplish this effect for participants overall in this study. It seems unlikely that students’ own mental images of the same information would be any more helpful.

It is important to note that the visual component of the computer program used in these experimental trials was predetermined. Images were consistent and notation standard, but participants had to work with exactly what was shown them.

They had to perceive the notation as displayed, and make quick decisions about how to use it in memory, in relation to the matching aural percept. When students create their own visual images, they are able to customize them for optimum utility and lasting memory. They can choose to use, for example, vivid colors, intricate designs, and visual backgrounds and highlights, to effectively store information, prioritize it, and make it available for the task at hand. Overt visual presentation of melodies may have eliminated these options for participants in this study. Future related research might attempt to capture the nature of aural and visual imagery generation during musical tasks, through detailed interviews with successful and unsuccessful dictation students, and through longitudinal studies of visual imagery applied to aural skills study.

Additional Analyses

Test order. A common concern of those engaged in informal discussion of this study was that participants would be able to take advantage of the identical melodies presented in both test formats. A practice effect would confound results, analysis, and subsequent validity of conclusions. Several steps were taken to mitigate this potential problem: practice items were different on each test format, test items were presented in different order on each test format, and the order of test formats was switched for every other participant. The first two steps were intended to avoid a practice effect for any given individual participant. The third was intended to equalize the effect across participants, should one have arisen.

In both the pilot test and in formal data collection, there was no significant difference in test scores based on test order. Participants did not score better on the second test, even when they had already heard the melodies on the first test. While this result contributes to the validity of the process and conclusions of this study, it raises a further question about melodic memory. Why did scores not improve on the second testing of the same material?

As evidenced in Appendix A, the discrepancies in items' incorrect answer choices were slight. This was intended to accomplish two objectives: avoiding obvious wrong answers, and keeping discrepancies within only the pitch or the rhythm category, not both. An obvious wrong answer would likely not affect presentation format data, as participants would benefit from the easy item on both aural-only and aural-visual forms. However, it would skew answer discrepancy data toward whichever category, pitch or rhythm, it represented. Keeping discrepancies to one or the other of the two categories was necessary for effective data analysis. Such slight discrepancies may have kept participants from generating a practice effect, even if they "knew" the target melodies better the second time. Many participants reported that the test was difficult because the answer choices sounded so much like the target melody. In the second test for each participant, not only were target melodies lingering in memory from the previous testing, but answer choices as well. The tests indeed appear to have forced each participant to focus only on the melody at hand, with no preconceived assumptions possible. Future research might involve memory stamina studies, in which truly identical tests are taken one after another, to measure intersections of participants' memory capacity, mental focus, and discrimination abilities.

Major instrument. Various major instruments present inherent idiomatic differences in musical study. Approaches to instruction and study on a given instrument are determined by the physical requirements for playing or singing, the ensembles in which the instrument is typically employed, and the specific legacy of master teachers, repertoire, and pedagogical methods associated with it. Because comparison by major instrument group was not a central focus of this study, participants were not specifically approached based on controls of any of these factors. The exploratory analysis of data by major instrument in this study revealed no significant differences on any test score measures. Future studies might compare musical perception and cognition of participants by instrumental grouping, taking into account specific musical training baseline data. For example, students of various instruments who had studied for a specific time period, from a given method book, or within a given program or school, could be tested on various musical tasks, to observe how inherent differences in instrumental training may affect musical perception and cognition.

Student status. Comparison of undergraduate and graduate students' scores was not a central focus of this study either. Based on pilot study results, it was assumed that all students could be pooled together for formal data collection without any resultant floor or ceiling effect. This assumption appears to have been correct. The follow-up analysis of these data indicates that graduate students did not score significantly higher than undergraduate students. Based on the results of this study, graduate students' additional years of formal instruction does not improve their melodic memory. This finding seems counterintuitive. We assume that musical skills of graduate students are typically more developed than those of undergraduates. In other words, graduate clarinet majors are, overall, better players than undergraduate clarinet majors, and graduate music history majors are better historians than undergraduates, for example. The key to this improvement is likely accumulated practice in the specific discipline. Graduate curricula tend to be more focused on the particular major of interest, therefore necessarily leaving out other areas of musical skill-building. Clarinet and history majors may have very little formal contact with melodic dictation, for instance, after their sophomore or junior year of undergraduate work. Even graduate theory majors may not have as much exposure to dictation in graduate school as they did during undergraduate work.

Many formally-trained musicians may be more successful in dictation tasks in their undergraduate years than they ever will be afterward, regardless of their subsequent career paths. During enrollment in aural skills classes, undergraduates benefit from regular practice with dictation skills. When these courses have been completed, skills may be weakened by less frequent exposure to the activities used to build them. Future research on this topic might analyze gradual disintegration over time of any of various musical skills universal to undergraduate curricula. Studies of this type could shed light on the nature of musical skill acquisition and maintenance in general, and provide impetus for potential revisions of undergraduate and graduate curricula.

Implications for Music Education

One of the unique challenges in the experimental trials of this study, and in melodic dictation, is that the melodies were unfamiliar. If the participants of this study had been asked to identify slight rhythm and pitch discrepancies in famous or familiar melodies, they would likely have performed much better because they would be making comparisons to a known aural referent. One might argue that this type of error detection would be more appropriate for, and useful to, music educators. In the profession, teachers and directors listen to rehearsals and classes and detect slight deviations from the score or from their intended aural image of a piece. The music is generally familiar to them, and they have had time to study it and to prepare themselves for the task. Is the ability to perform this task with unfamiliar music and little time a worthwhile skill?

Sight reading is perhaps the most obvious reason for developing cognitive fluency in relation to unfamiliar music. Music educators in virtually every teaching situation encounter the opportunity to sight read—conducting a new score, playing piano accompaniments, or modeling voice parts, for example. Even if teachers play it safe, never allowing themselves to sight read in the classroom, they must sight read for themselves the first time they look at new music during their preparation. In this case, what is at stake is teachers' efficiency, a high priority in a demanding profession.

Sight reading, like dictation, is a good measure of development of connections between the visual and aural aspects of music. Ostensibly, sight reading measures visual-aural transfer, and dictation measures aural-visual transfer. In reality, both tasks measure both types of transfer, because a constant cycle of checking and rechecking occurs in both. Sight readers see music, create a quick aural preview of how it should sound, and sing or play it. The cycle only becomes complete when they hear what they performed, compare it to what they intended, and “see” the discrepancies as compared to the written music, thus informing subsequent reading. Likewise, dictation-takers must look at what they have sketched, “hear” what it sounds like, and compare it to what they physically heard, in a continuous cycle of observation and revision.

Mastery of these cycles of aural-visual connections is vitally important to music educators’ skill repertoire for rehearsals, private and group lessons, improvisation and composition activities, and score study, for example. Music teachers are constantly comparing and evaluating visual and aural renditions of musical materials, and using the results to inform short- and long-term decisions affecting their curricula. Not only are aural-visual connection skills a necessary part of teachers’ own skill repertoire, they are worthy of student learning at all levels. Through performance, improvisation, composition, listening, analysis, and critique activities, music educators have many opportunities to help students develop strong cognitive connections between the aural and visual aspects of music, a skill universally useful in experiences with Western music.

REFERENCES

Beckett, C. A. (1997). Directing student attention during two-part dictation. *Journal of Research in Music Education, 45*(4), 613-625.

Beheshti, S. (2009). Improving studio music teaching through understanding learning styles. *International Journal of Music Education, 27*(2), 107–115.

Burton, L. J. (2003). Examining the relation between visual imagery and spatial ability tests. *International Journal of Testing, 3*(3), 277-291.

Bluestine, E. (2007). *A comparative study of four approaches to teaching tonal music reading to a select group of students in third, fourth, and fifth grade* (Doctoral dissertation). Available from ProQuest Dissertations and Theses Database. (UMI No. 3268133)

Boisen, R. (1981). The effect of melodic context on students' aural perception of rhythm. *Journal of Research in Music Education, 29*(3), 165-172.

Brand, M., & Burnsed, V. (1981). Music abilities and experiences as predictors of error-detection skill. *Journal of Research in Music Education, 29*(2), 91-96.

Burman, D. D., & Booth, J. R. (2009). Music rehearsal increases the perceptual span for notation. *Music Perception, 26*(4), 303-320.

Byo, J.L. (1997). The effects of texture and number of parts on the ability of music majors to detect performance errors. *Journal of Research in Music Education, 45*(1), 51-66.

Byrne, B., & Fielding-Bardsley, R. (1995). Evaluation of a programme to teach phonemic awareness to young children: A 2- and 3-year follow-up. *Journal of Educational Psychology, 87*(1), 488-503.

Cassidy, J. W. (2001). Listening maps: Undergraduate students' ability to interpret various iconic representations. *Update – Applications of Research in Music Education, 19*(2), 15-19.

Center, Y., Freeman, L., Robertson, G., & Outhred, L. (1999). The effect of visual imagery training on the reading and listening comprehension of low listening comprehenders in Year 2. *Journal of Research in Reading, 22*(3), 241-256.

Colwell, R. J. (1969). *Music Achievement Test*. Chicago: Follett Educational Corporation.

Covington, K. (1992). An alternate approach to aural training. *Journal of Music Theory Pedagogy, 6*, 5-18.

Cumming, N. (2000). *The Sonic Self*. Bloomington: Indiana University Press.

Elliott, D. J. (1995). Music matters: a new philosophy of music education. Oxford: Oxford University Press.

Forgeard, M., Schlaug, G., Norton, A., Rosam, C., Iyengar, U., & Winner, E. (2008). The relation between music and phonological processing in normal-reading children and children with dyslexia. *Music Perception 25*(4), 383-390.

Gambrell, L., & Bales, R. (1987). Visual imagery: A strategy for enhancing listening, reading, and writing. *Australian Journal of Reading, 10*(3), 147-53.

Giannakis, K., & Smith, M. (2001). Imaging soundscapes: Identifying cognitive associations between auditory and visual dimensions. In R. Godoy & H. Jorgensen (Eds.), *Musical imagery* (pp. 161-179). Lisse: Swets & Zeitlinger.

Good, R., & Kaminski, R. (1988). *Dynamic indicators of basic early literacy skills*. Frederick, CO: Sopris West.

Goodman, N. (1978). *Ways of Worldmaking*. Indianapolis: Hackett.

Gordon, E. E. (1986). *Intermediate measures of music audiation*. Chicago: GIA Publications.

Gordon, E. E. (1986). *Primary measures of music audiation*. Chicago: GIA Publications.

Gordon, E. E. (1991). *Iowa test of music literacy*. Chicago: GIA Publications.

Gordon, E. E. (2007). *Learning sequences in music: a contemporary music learning theory*. Chicago: GIA Publications.

Gordon, E. E. (2007a). *Awakening newborns, children, and adults to the world of audiation*. Chicago: GIA Publications.

Grutzmacher, P. A. (1987). The effect of tonal pattern training on the aural perception, reading recognition, and melodic sight-reading achievement of first-year instrumental music students. *Journal of Research in Music Education, 35*(3), 171-181.

Guttman, J., Levin, J. R., & Pressley, M. (1977). Pictures, partial pictures, and young children's oral prose learning. *Journal of Educational Psychology, 69*(5), 473-480.

GVOX, Inc. (2009). *Encore Notational Software*. East Brunswick, NJ: GVOX, Inc.

Halfpenny, E., Ho, S., Kurosawa, K., & Wollner, C. (2003). The effects of distracted inner hearing on sight-reading. *Psychology of Music, 31*(4), 377-389.

Halpern, A. R. (1988). Mental scanning in auditory imagery for songs. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*(3), 434-443.

Harrison, C. S. (1990). Relationships between grades in the components of freshman music theory and selected background variables. *Journal of Research in Music Education, 38*(3), 175-186.

Hishitani, S. (1990). Imagery experts: How do expert abacus operators process imagery? *Applied Cognitive Psychology, 4*, 33-46.

Hubbard, T. L., & Stoeckig, K. (1988). Musical imagery: Generation of tones and chords. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*(4), 656-667.

Iwamiya, S. (1994). Interactions between auditory and visual processing when listening to music in an audiovisual context: 1. Matching 2. Audio quality. *Psychomusicology, 13*, 133-154.

Jakobson, L. S., Lewycky, S. T., Kilgour, A. R., & Stoesz, B. M. (2008). Memory for verbal and visual material in highly trained musicians. *Music Perception* *26*(1), 41-55.

Karpinski, G. (2000). *Aural skills acquisition*. Oxford: Oxford University Press.

Killian, J. N. (1991). The relationship between sightsinging accuracy and error detection in junior high singers. *Journal of Research in Music Education,* *39*(3), 216-224.

Killian, J. N., & Henry, M. L. (2005). A comparison of successful and unsuccessful strategies in individual sight-singing preparation and performance. *Journal of Research in Music Education, 53*(1), 51-66.

Kizkin, S., Karlidag, R., Ozcan, C., & Ozisik, H. I. (2006). Reduced P50 auditory sensory gating response in professional musicians. *Brain and Cognition, 61*(3), 249-254.

Koniari, D., Predazzer, S., & Mélen, M. (2001). Categorization and schematization processes used in music perception by 10- to 11-year-old children. *Music Perception, 18*(3), 297- 324.

Larson, R. C. (1977). Relationships between melodic error detection, melodic dictation, and melodic sightsinging. *Journal of Research in Music Education, 25*(4), 264-271.

Livingston, C., & Ackman, J. (2003). Changing trends in preparing students for college level theory. *American Music Teacher, 53*(1), 26-29.

Long, P. A. (1977). Relationships between pitch memory in short melodies and selected factors. *Journal of Research in Music Education, 25*(4), 272-282.

Lucas, J. R., & Gromko, J. E. (2007). The relationship of musical pattern discrimination skill and phonemic awareness in beginning readers. *Contributions to Music Education, 34,* 9-17.

MacKnight, C. B. (1975). Music reading ability of beginning wind instrumentalists after melodic instruction. *Journal of Research in Music Education, 23*(1), 23-34.

Madsen, C. K., & Staum, M. J. (1983). Discrimination and interference in the recall of melodic stimuli. *Journal of Research in Music Education, 31*(1), 15-31.

Margulis, E. H. (2005). A model of melodic expectation. *Music Perception 22*(4), 663-714.

McClung, A. C. (2001). Sight-singing systems: current practice and survey of all-state choristers. *Update-Applications of Research in Music Education, 20*(1), 3-8.

Meyer, L. B. (1968). *Emotion and meaning in music*. Chicago: University of Chicago Press.

Microsoft Corporation (2009). *Windows Movie Maker*. Redmond, WA: Microsoft.

Mikumo, M. (1994). Motor encoding strategy for pitches of melodies. *Music Perception, 12*(2), 175-197.

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review, 63,* 81-97.

Morrison, S. J., Demorest, S. M., Aylward, E. H., Cramer, S. C., & Maravilla, K. R. (2003). fMRI investigation of cross-cultural music comprehension. *Neuroimage, 20,* 378–384.

Neale, M. D. (1999). *Neale analysis of reading ability*. Melbourne: Acer Press.

Norris, C. E. (2000). Factors related to the validity of reproduction tonal memory tests. *Journal of Research in Music Education*, *48*(1), 52-64.

Norris, C. E. (2003). The relationship between sight singing achievement and melodic dictation achievement. *Contributions to Music Education, 30*(1), 39-53.

Olson, G. (1978). Intersensory and intrasensory transfer of melodic contour perception by children. *Journal of Research in Music Education, 26*(1), 41-47.

Palmer, C., and Krumhansl, C. L. (1987). Independent temporal and pitch structures in determination of musical phrases. *Journal of Experimental Psychology: Human Perception and Performance, 13*(1), 116-126.

Pembrook, R. G. (1986). Interference of the transcription process and other selected variables on perception and memory during melodic dictation. *Journal of Research in Music Education, 34*(4), 238-261.

Pembrook, R. G. (1987). The effect of vocalization on melodic memory conservation. *Journal of Research in Music Education*, *35*(3), 155-169.

Pinto, M., & Tall, D. (2002). Building formal mathematics on visual imagery; A case study and a theory. *For The Learning of Mathematics, 22*(1), 2-10.

Povel, D., & Jansen, E. (2001). Perceptual mechanisms in music processing. *Music Perception, 19*(2), 169-198.

Povel, D., & Jansen, E. (2002). Harmonic factors in the perception of tonal melodies. *Music Perception, 20*(1), 51-85.

Rinck, M., & Denis, M. (2004). The metrics of spatial distance traversed during mental imagery. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*(6), 1211-1218.

Saariluoma, P., & Kalakoski, V. (1997). Skilled imagery and long-term working memory. *The American Journal of Psychology, 110*(2), 177-201.

Saariluoma, P., & Kalakoski, V. (1998). Apperception and imagery in blindfold chess. *Memory, 6*(1), 67-90.

Segal, S. J., & Fusella, V. (1970). Influence of imaged pictures and sounds on detection of visual and auditory signals. *Journal of Experimental Psychology, 83*(3), 458-464.

Sheldon, D. A. (1998). Effects of contextual sight-singing and aural skills training on error detection abilities. *Journal of Research in Music Education, 46*(3), 384-395.

Telesco, P. (1991). Contextual ear training. *Journal of Music Theory Pedagogy, 5*, 179-190.

Thompson, K. A. (2004). Thinking in sound: A qualitative study of metaphors for pitch perception. *Journal of Music Theory Pedagogy, 18*, 81-107.

Watkins, J. G., & Farnum, S. E. (1954). *The Watkins-Farnum Performance Scale*. Winona, Minn: Hal Leonard Music.

Weber, R., & Harnish R. (1974). Visual imagery for words – the Hebb test. *Journal of Experimental Psychology, 102*(3), 409-414.

Weber, R. J., Kelley, J., & Little, S. (1972). Is visual image sequencing under verbal control? *Journal of Experimental Psychology, 96*(2), 354-362.

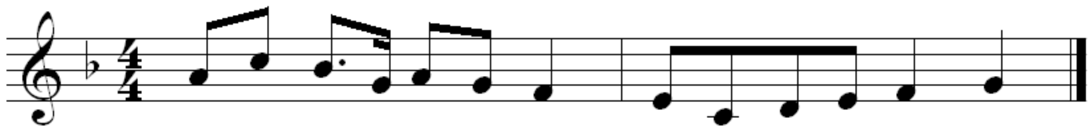
Wilson S. J., & Saling, M. M. (2008). Contributions of the right and left mesial temporal lobes to music memory: Evidence from melodic learning difficulties. *Music Perception 25*, 303-314.

APPENDIX A

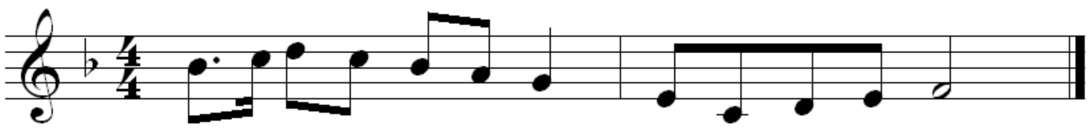
AURAL-ONLY TEST MELODIES

Note: MM=120 for all melodies on this test.

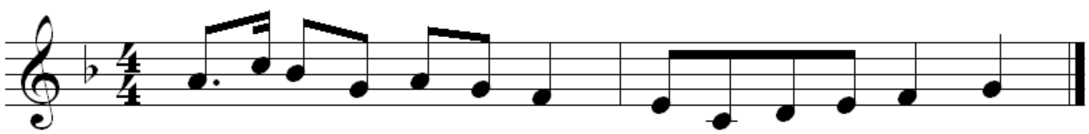
Target melody (practice)



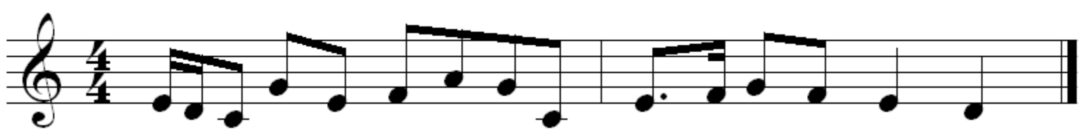
Distraction melody (practice)



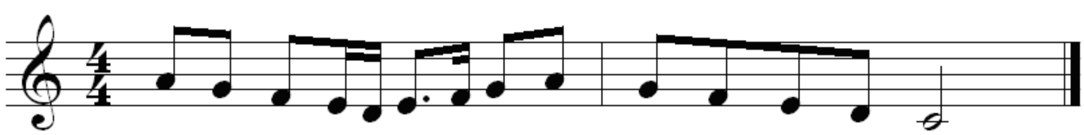
Incorrect answer choice (practice)



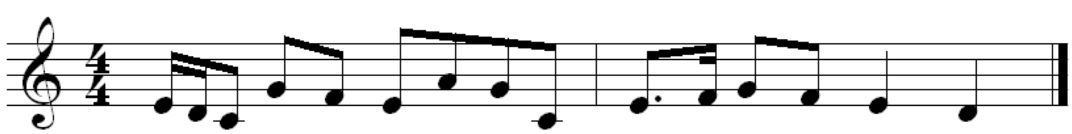
Target melody 1



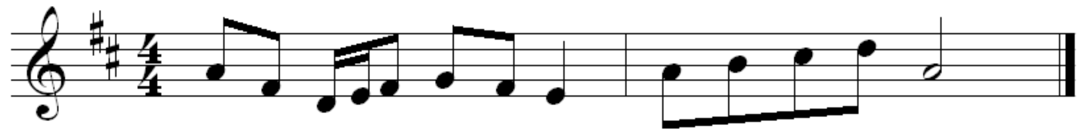
Distraction melody 1



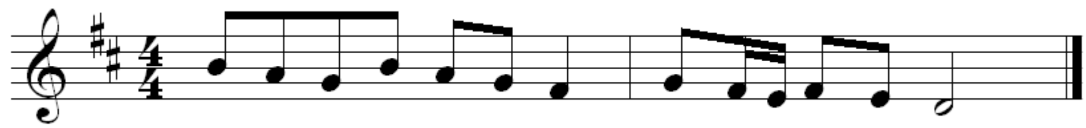
Incorrect answer choice 1



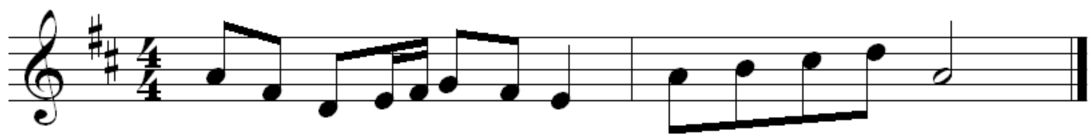
Target melody 2



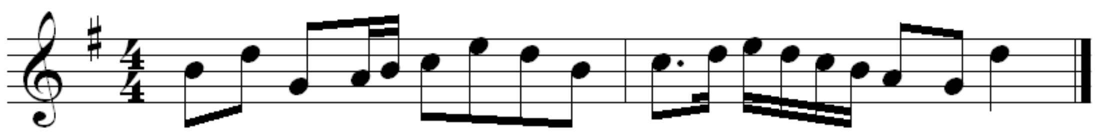
Distraction melody 2



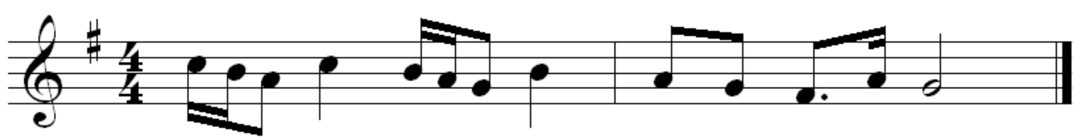
Incorrect answer choice 2



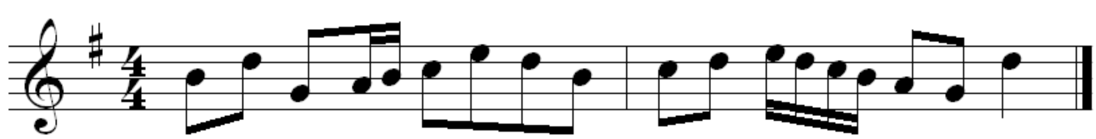
Target melody 3



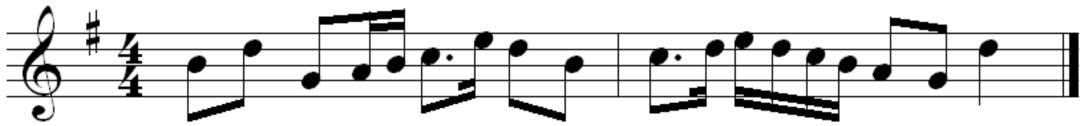
Distraction melody 3



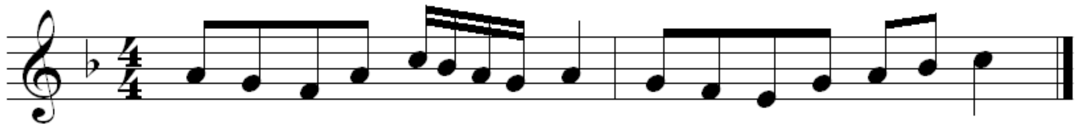
Incorrect answer choice 3a



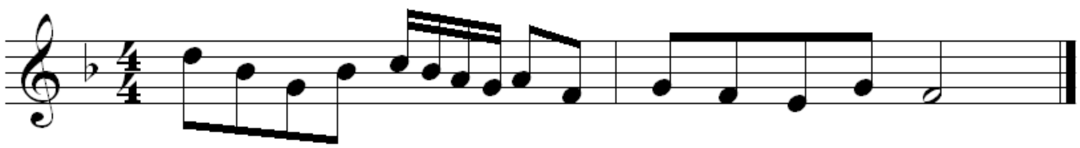
Incorrect answer choice 3b



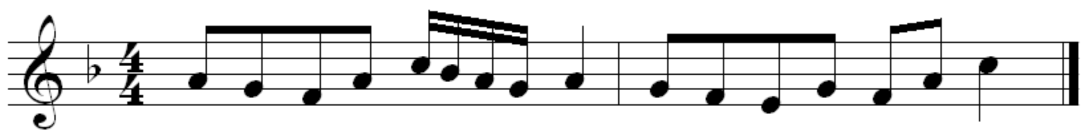
Target melody 4



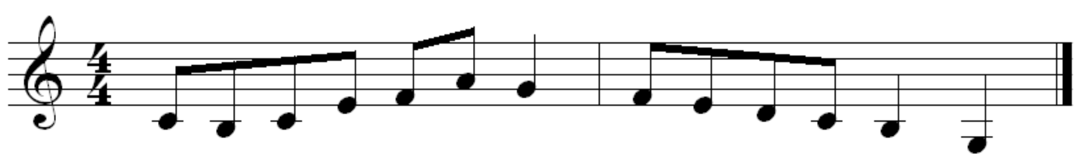
Distraction melody 4



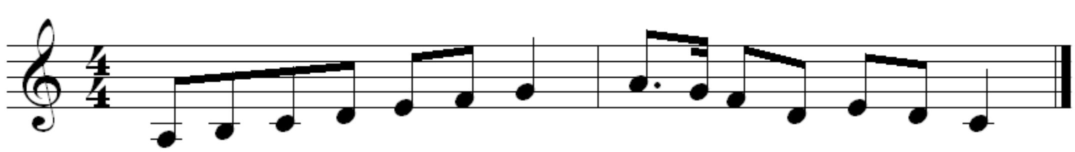
Incorrect answer choice 4



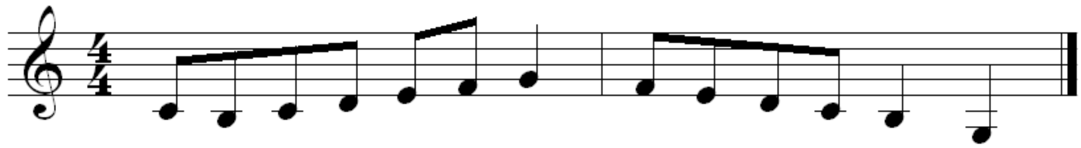
Target melody 5



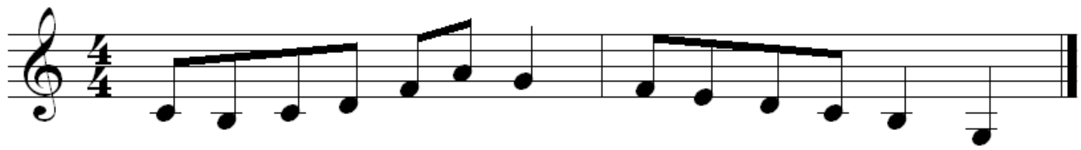
Distraction melody 5



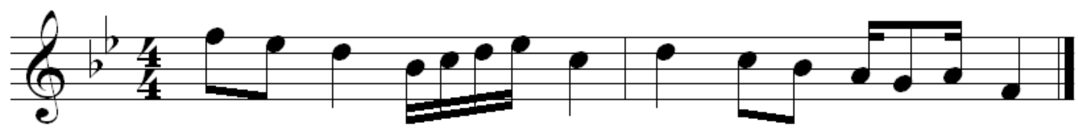
Incorrect answer choice 5a



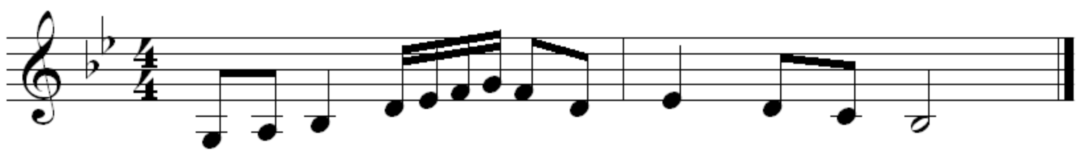
Incorrect answer choice 5b



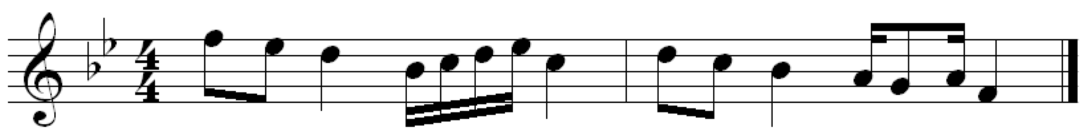
Target melody 6



Distraction melody 6



Incorrect answer choice 6

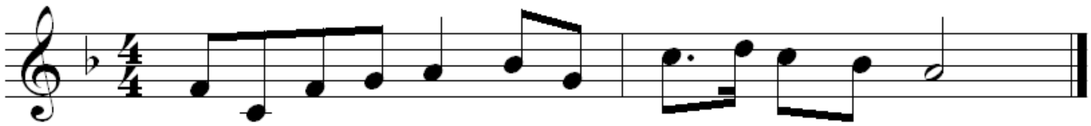


APPENDIX B

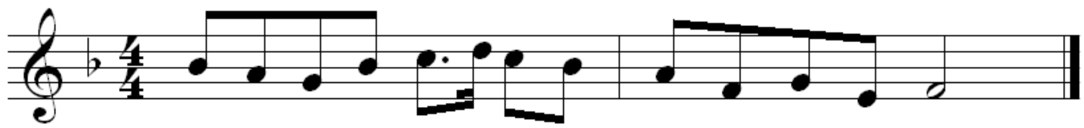
AURAL-VISUAL TEST MELODIES

Note: MM=120 for all melodies on this test.

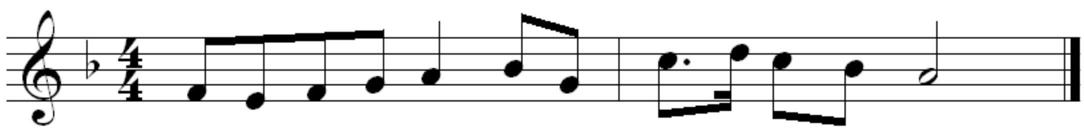
Target melody (practice)



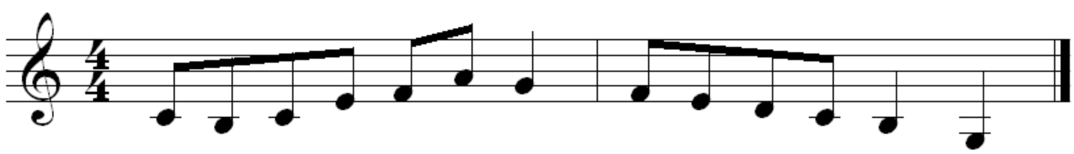
Distraction melody (practice)



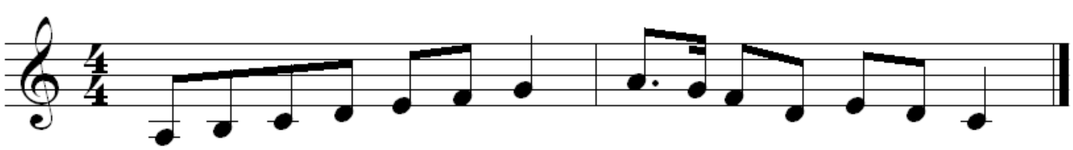
Incorrect answer choice (practice)



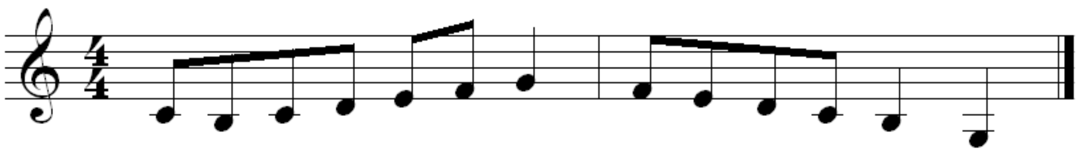
Target melody 1



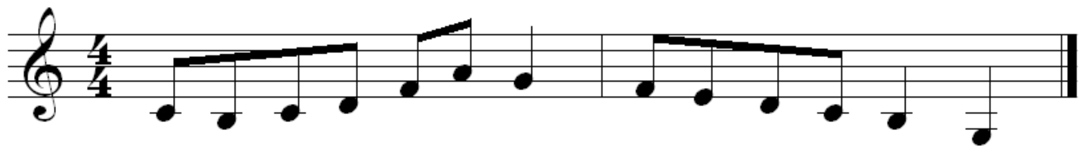
Distraction melody 1



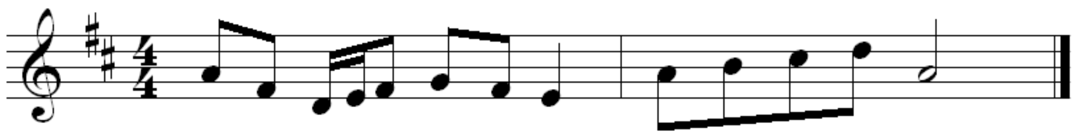
Incorrect answer choice 1a



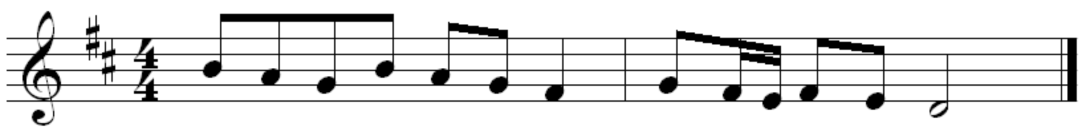
Incorrect answer choice 1b



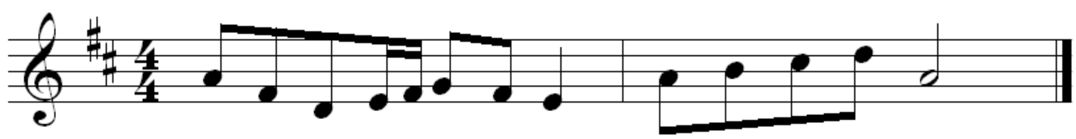
Target melody 2



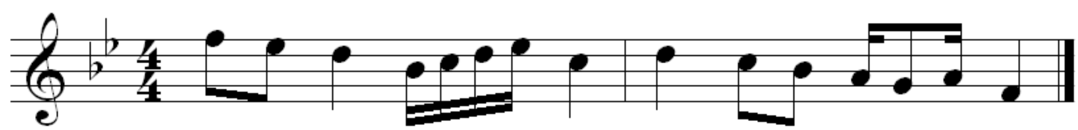
Distraction melody 2



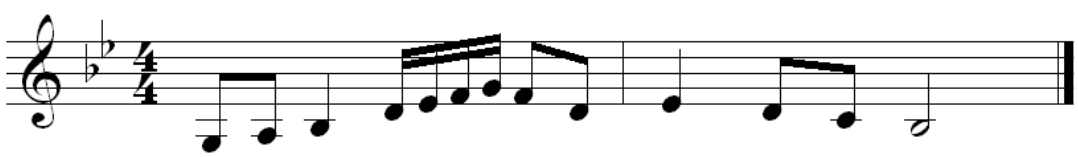
Incorrect answer choice 2



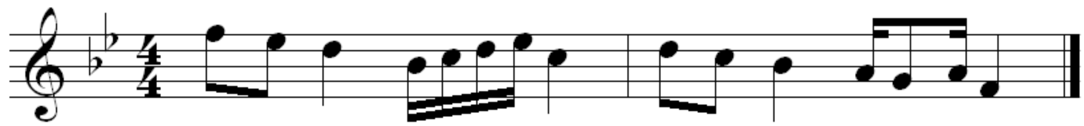
Target melody 3



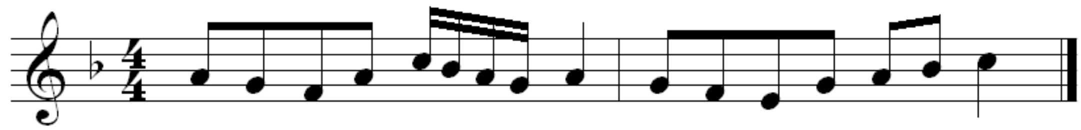
Distraction melody 3



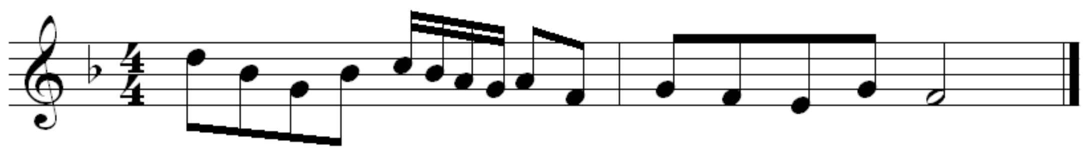
Incorrect answer choice 3



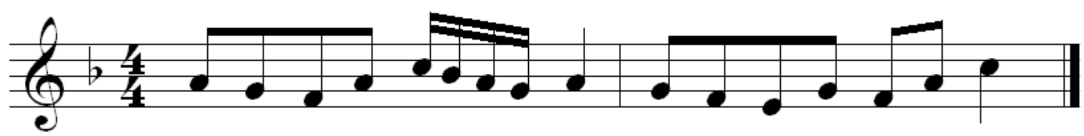
Target melody 4



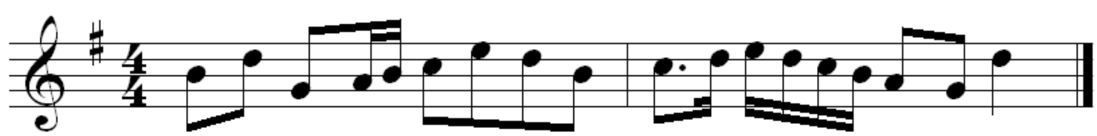
Distraction melody 4



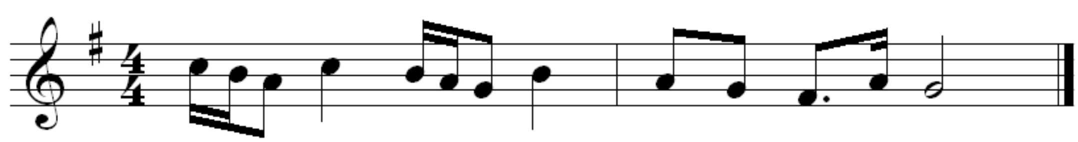
Incorrect answer choice 4



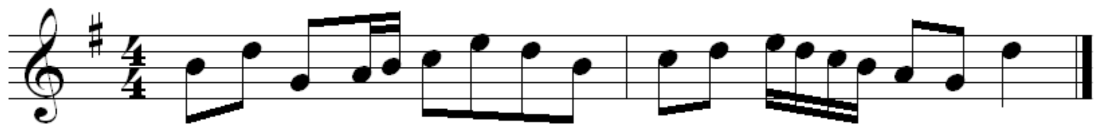
Target melody 5



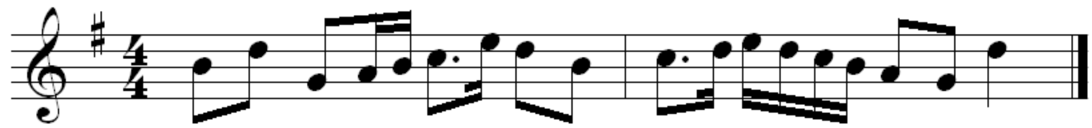
Distraction melody 5



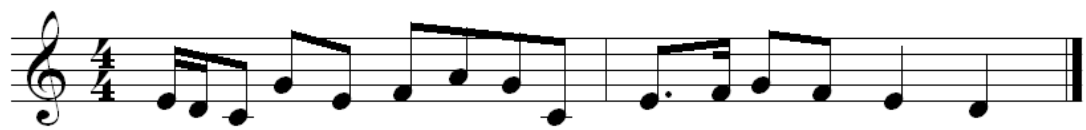
Incorrect answer choice 5a



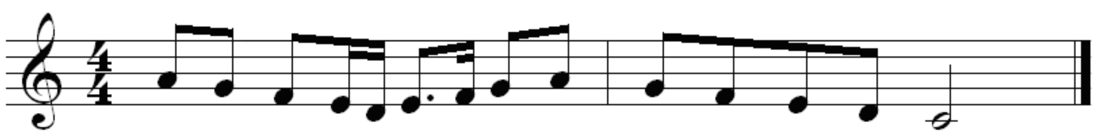
Incorrect answer choice 5b



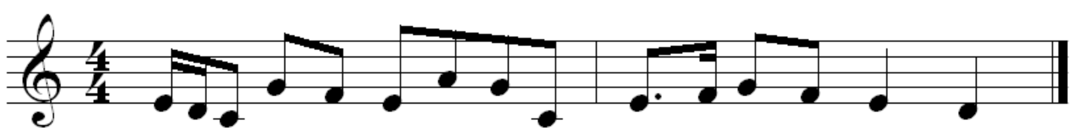
Target melody 6



Distraction melody 6



Incorrect answer choice 6



APPENDIX C

AURAL-ONLY ANSWER SHEET

Each item in this program begins with a I-V7-I chord progression to establish the key. After the chord progression, listen to the first melodic phrase and try to remember it. After a short pause, you will hear a distraction melodic phrase. Focus on remembering the original phrase. You will then hear two answer choices. Circle the answer choice that matches the original exactly. If neither choice matches the original exactly, circle "neither."

(Practice)	Melody A	Melody B	Neither
1.	Melody A	Melody B	Neither
2.	Melody A	Melody B	Neither
3.	Melody A	Melody B	Neither
4.	Melody A	Melody B	Neither
5.	Melody A	Melody B	Neither
6.	Melody A	Melody B	Neither

APPENDIX D

AURAL-VISUAL ANSWER SHEET

Each item in this program begins with a I-V7-I chord progression to establish the key. After the chord progression, listen to the first melodic phrase, view its notation on the screen immediately following, and try to remember it. After the notation disappears, you will hear a distraction melodic phrase. Focus on remembering the original phrase. You will then hear two answer choices. Circle the answer choice that matches the original exactly. If neither choice matches the original exactly, circle "neither."

(Practice)	Melody A	Melody B	Neither
1.	Melody A	Melody B	Neither
2.	Melody A	Melody B	Neither
3.	Melody A	Melody B	Neither
4.	Melody A	Melody B	Neither
5.	Melody A	Melody B	Neither
6.	Melody A	Melody B	Neither

APPENDIX E

IRB PROTOCOL

May 1, 2009

Richard Throm, Program Manager & Coordinator
Institutional Review Board
3400 N. Broad Street (509-00)
Philadelphia, PA 19140

Dear Richard,

Enclosed please find three (3) copies each of the following items as they pertain to research entitled *Effects of Visual Presentation on Aural Memory for Melodies:*

- IRB Request for Protocol Review Cover Page (Behavioral and Social Sciences)
- IRB Protocol
- Participant Invitation Letter
- Participant Consent Form
- Experimental Instrument

The pages immediately following this cover letter confirm that our IRB research training is current. Thank you for your consideration of this protocol. We appreciate your time and look forward to hearing from you.

Sincerely,

Deborah Sheldon, PhD
Professor and Chair, Department of Music Education and Therapy
Boyer College of Music and Dance, Temple University

Nathan Buonviri
Dean's Appointment, Department of Music Education and Therapy
Boyer College of Music and Dance, Temple University

	TEMPLE UNIVERSITY Office of the Vice President for Research **Institutional Review Board** **Committee B** (215) 707-8757 Fax: (215) 707-8387 www.research.temple.edu/irb	COMMITTEE USE ONLY PROTOCOL NUMBER

REQUEST FOR PROTOCOL REVIEW (BEHAVIORAL & SOCIAL SCIENCES)

Please follow all instructions. Use additional paper when necessary. This form can be downloaded at the website referenced above.

I. PRINCIPAL INVESTIGATOR – IF STUDENT RESEARCH, ADVISOR IS PRINCIPAL INVESTIGATOR

NAME, DEGREE	AFFILIATION WITH TEMPLE	PHONE	FAX
Deborah A. Sheldon, PhD	Professor	204-8649	204-1982

SCHOOL/COLLEGE, CENTER/DEPARTMENT, AND SECTION	TEMPLE EMAIL (REQUIRED)
Boyer College of Music and Dance – Music Education	dsheldon@temple.edu

PREFERRED MAILING ADDRESS
****

ACCESSNET ID (REQUIRED)	9 DIGIT TUID (REQUIRED)
dsheldon	*********

SIGNATURE OF PRINCIPAL INVESTIGATOR - IF STUDENT RESEARCH, ADVISOR IS PRINCIPAL INVESTIGATOR	DATE
Signature: ______________________________ Printed Name: Deborah A. Sheldon	May 1, 2009

II. STUDENT INVESTIGATOR – TEMPLE STUDENT

NAME, DEGREE	AFFILIATION WITH TEMPLE	PHONE	FAX
Nathan O. Buonviri, PhD Student	Student	267-265-0236	204-1982

SCHOOL/COLLEGE, CENTER/DEPARTMENT, AND SECTION	TEMPLE EMAIL (REQUIRED)
Boyer College of Music and Dance – Music Education	nathan.buonviri@temple.edu

PREFERRED MAILING ADDRESS
****

ACCESSNET ID (REQUIRED)	9 DIGIT TUID (REQUIRED)
tua63242	*********

SIGNATURE OF STUDENT RESEARCHER	DATE
Signature: ______________________________ Printed Name: Nathan O. Buonviri	May 1, 2009

III. PROJECT CATEGORY

☐	Faculty Research	☒	Dissertation Research
☐	Master's Research	☐	Other Graduate Research
☐	Undergraduate Research	☐	Undergraduate Independent Study
☐	Undergraduate Course Requirement	☐	Administrative Research
☐	Other (please specify):		

IV. PROJECT DATA

TITLE OF PROJECT

Effects of Visual Presentation on Aural Memory for Melodies

FUNDING AGENCY	PROPOSED STARTING DATE August 20, 2009	ESTIMATED DURATION 4 months

STUDY LOCATION
Presser Hall, Boyer College of Music and Dance, Temple University

IS DATA FOR THIS STUDY BEING OBTAINED FROM ANOTHER SOURCE?
☒ **No** ☐ **Yes**

IF YES, IDENTIFY THE SOURCE AND PROVIDE DOCUMENTED PERMISSION TO USE THE DATA.

PLEASE NOTE

IF YOUR PROTOCOL IS DETERMINED TO REQUIRE FULL COMMITTEE REVIEW, YOU WILL BE REQUESTED TO PROVIDE ADDITIONAL COPIES (20 TOTAL)
FORWARD THREE (3) COPIES OF THIS FORM WITH PROTOCOL AND CONSENT FORM(S) TO:
RICHARD THROM, DIRECTOR, OFFICE FOR HUMAN SUBJECTS PROTECTION
PROGRAM MANAGER & COORDINATOR, IRB
3rd FLOOR HUDSON BUILDING (555-00)
3425 NORTH CARLISLE STREET
PHILADELPHIA, PA 19140

Part I. Characteristics of Potential Subjects

A. About how many subjects will you need?

This study will require forty participants.

B. Describe the potential subjects in terms of gender, age range, ethnic group, economic status, and any other significant descriptors.

Subjects are college music majors, ages 18-21, who have passed the first two semesters of Music Theory at Temple University. No potential subject will be excluded based on gender, ethnic group, or economic status. It is assumed, though not guaranteed, that subjects will represent substantial diversity regarding these descriptors.

C. Indicate any special subject characteristics, such as persons with mental handicaps, prisoners, pregnant women, etc.

None

D. What is the general state of health of the subjects (physical and mental)?

All subjects are presumed to be healthy.

E. Describe how you will gain access to these potential subjects?

Potential subjects will be identified through music theory courses at Temple. A letter of introduction and invitation to potential subjects is attached to this protocol (Appendix A).

F. How will subjects be selected or excluded from the study?

Students who have not passed the first two semesters of Music Theory at Temple University will not be invited to participate in the study.

G. If subjects are from an institution other than Temple University, please indicate the name of the office responsible for granting access to the subjects.

N/A

H. If the subjects are children, anyone suffering from a known psychiatric condition, or legally restricted, please explain why it is necessary to use these persons as subjects.

N/A

Part II. Experimental or Research Procedure

A. Please describe the intended experimental or research procedure. This should include a description of what the subject will experience or be required to do. Please attach a copy of all questionnaires or instruments to be used.

The purpose of this study is to determine whether melodic memory is less affected by aural distractions when melodic stimuli are presented both visually and aurally, as compared to aurally only. Pembrook (1986) noted the human inability to remember all the information in a melody long enough to be able to write it down (a process called melodic dictation). The process of remembering even a portion of a melodic dictation is complicated by the aural distractions of ensuing sections of the target melody. Other studies (e.g. Segal & Fusella (1970), Mikumo (1994), and Cassidy (2001)) suggest that students may benefit from visualizing aural stimuli to overcome aural distractions, and to effectively store and retrieve the aurally-perceived information. While it is difficult to test whether subjects are visualizing a melody, it is feasible and worthwhile to explore whether concurrent aural and visual presentation of a melody improves melodic memory, as compared to aural presentation only. This project consists of a self-contained computer program that tests subjects' ability to remember melodic fragments in the midst of distraction melodies provided by the program. Subjects will perform the memory task under two conditions:

1. **Aural-only presentation of the melodic stimuli**
2. **Aural and visual presentation of the melodic stimuli**

Results will be analyzed in terms of effects on memory for two characteristics of melody: rhythm and pitch. Subjects' scores will be compared using a 2x2 (presentation format x melodic characteristic) ANOVA.

Please see Appendix C, "Experimental Instrument".

B. Will the subjects be deceived in any way? If yes, please describe below.

Subjects will not be deceived in any way.

C. To what extent will the routine activities of the subject be interrupted during the course of the study?

Each subject will participate in one 20-minute self-guided trial. Routine activities of the subject will not be interrupted.

D. Indicate any compensation for the subjects.

None

Part III. Data Confidentiality

A. What procedure(s) will you use to insure confidentiality of the data. How will you preserve subject anonymity?

All data will be coded by participant number and handled only by the two investigators.

Part IV. Consent Procedures

A. Attach a copy of consent form to be used. If non-written consent is to be used, attach a statement describing exactly what the subjects will be told.

Please see Appendix B, "Consent Form".

B. Describe how you will handle consent procedure for minors, mentally challenged persons, and persons with significant emotional disturbances.

N/A

Part V. Benefits of the Study

A. How will any one subject benefit from participation in this study?

While no benefits are guaranteed, the results of this study may help subjects to better understand and synthesize visual and aural strategies for melodic memory, a worthwhile skill for music majors. They may discover new ways to connect musical skills and improve overall musicality.

B. How will society, in general benefit from the conduct of this study?

While no benefits to society are guaranteed, music teachers may find the results of this study helpful in guiding and revising music curricula. The results of the study may also be helpful to teachers and students of other subjects in terms of cognitive strategies for processing aural and visual stimuli, separately and in combination.

Part VI. Risks/Discomforts to Subjects

A. Describe any aspect of the research project that might cause discomfort, inconvenience, or physical danger to the subjects.

No aspect of this project will cause discomfort, inconvenience, or physical danger to subjects.

B. Describe any long range risks to the subjects.

There are no risks to the subjects, either in the short or long term.

C. What is the rationale for exposing the subjects to these risks?

N/A

IRB Protocol Appendix A
Participant Invitation Letter

May 1, 2009

(Full Name)
Boyer College of Music and Dance
Temple University
Philadelphia, PA, 19122

Dear (Name),

Greetings from the Department of Music Education and Therapy! We are currently engaged in a study of melodic memory strategies. Our records show that you have successfully passed the first two semesters of Music Theory at Temple. This letter is to invite you to participate in our study. Participants will be asked to complete one 20-minute self-guided program on the computer. The program will consist of short melodic memory self-tests in the midst of distraction melodies.

We think that results of this research may provide valuable insights into visual and auditory processing of music, and teaching and learning of melodic dictation. If you are willing to participate in this study, please read the attached consent form. If you are in agreement with the terms of the consent form, please sign and return to us.

Thank you for your time.

Sincerely,

Deborah Sheldon
Professor and Chair, Department of Music Education and Therapy
Boyer College of Music and Dance
Temple University
deborah.sheldon@temple.edu

Nathan Buonviri
Graduate Fellow, Department of Music Education and Therapy
Boyer College of Music and Dance
Temple University
nathan.buonviri@temple.edu

IRB Protocol Appendix B

Consent Form

Consent Form

Effects of Visual Presentation on Aural Memory for Melodies

Deborah A. Sheldon
Nathan O. Buonviri
Department of Music Education and Therapy
(215) 204-8649

We are currently engaged in a study of melodic memory strategies. You have been invited to participate because of your success in Music Theory at the college level. To help us gain further insights into this topic we ask you to complete one 20-minute self-guided program on the computer. All data will be handled only by the two of us, and will be maintained anonymously through the use of participant numbers.

We welcome questions about the experiment at any time. Your participation in this study is on a voluntary basis, and your refusal to participate at any time will not prejudice future interactions with us or Temple University. There is no compensation for participation in the study, but benefits may include better understanding of musical memory strategies, enhanced connection of musical skills, and improved musicality.

Signing your name below indicates that you have read and understand the contents of this Consent Form, that you agree to take part in this study, and that you have read the following statement:

"I understand that if I wish further information regarding my rights as a research subject, I may contact Richard Throm, Program Manager & Coordinator at Office of the Vice President for Research of Temple University by phoning (215) 707-8757."

__________________________________ __________________

Participant's Signature Date

__________________________________ __________________

Primary Investigator's Signature Date

IRB Protocol Appendix C

Aural-Visual Answer Sheet

Each item in this program begins with a I-V7-I chord progression to establish the key. After the chord progression, listen to the first melodic phrase, view its notation on the screen immediately following, and try to remember it. After the notation disappears, you will hear a distraction melodic phrase. Focus on remembering the original phrase. You will then hear two answer choices. Circle the answer choice that matches the original exactly. If neither choice matches the original exactly, circle "neither."

(Practice)	Melody A	Melody B	Neither
1.	Melody A	Melody B	Neither
2.	Melody A	Melody B	Neither
3.	Melody A	Melody B	Neither
4.	Melody A	Melody B	Neither
5.	Melody A	Melody B	Neither
6.	Melody A	Melody B	Neither

IRB Protocol Appendix D

Aural-Only Answer Sheet

Each item in this program begins with a I-V7-I chord progression to establish the key. After the chord progression, listen to the first melodic phrase and try to remember it. After a short pause, you will hear a distraction melodic phrase. Focus on remembering the original phrase. You will then hear two answer choices. Circle the answer choice that matches the original exactly. If neither choice matches the original exactly, circle "neither."

(Practice)	Melody A	Melody B	Neither
1.	Melody A	Melody B	Neither
2.	Melody A	Melody B	Neither
3.	Melody A	Melody B	Neither
4.	Melody A	Melody B	Neither
5.	Melody A	Melody B	Neither
6.	Melody A	Melody B	Neither

APPENDIX F

EXPERIMENTAL DATA

Aural-Only Test Aural-Visual Test

	Item	1	2	3	4	5	6		1	2	3	4	5	6
	Discrepancy Type	P	R	R	P	P	R		P	R	R	P	R	P
Correct Answer		2	1	3	1	3	2		3	1	2	1	3	2
Participant														
1		2	1	1	2	1	2		1	1	2	3	1	2
2		2	2	1	1	3	2		1	2	2	2	3	3
3		2	2	2	2	2	1		1	1	2	2	1	2
4		2	1	3	1	2	2		3	1	1	1	1	2
5		2	1	2	1	3	2		1	2	1	1	3	2
6		2	1	3	1	2	2		1	1	2	2	3	1
7		2	1	1	3	2	3		2	1	2	1	1	2
8		1	1	3	3	1	1		3	3	2	2	1	1
9		2	1	2	1	3	2		1	2	2	2	3	2
10		2	3	1	2	1	2		3	2	1	1	2	2
11		1	2	2	3	1	1		2	1	2	3	1	2
12		2	2	3	1	1	3		3	3	1	1	3	2
13		2	1	1	2	3	1		1	1	2	2	1	2
14		3	1	3	2	3	1		1	3	2	1	3	2
15		2	2	1	2	1	1		1	1	3	1	2	2
16		2	2	3	1	1	1		1	1	2	2	2	1
17		2	1	3	2	1	1		1	1	2	2	2	1
18		2	1	1	2	3	1		1	2	2	1	3	1
19		1	1	3	3	3	1		1	3	2	1	1	2
20		2	1	3	1	3	2		3	1	2	1	3	3
21		2	2	1	1	1	2		1	3	2	1	3	1
22		2	2	3	1	3	2		1	2	1	2	3	1
23		1	3	2	3	3	1		2	2	2	1	1	2
24		2	1	1	2	3	1		3	1	2	1	2	3
25		2	1	1	2	1	3		1	1	2	1	3	1
26		2	2	3	2	3	2		3	2	2	1	3	2
27		2	3	1	1	3	1		3	1	1	3	2	3
28		1	1	3	2	1	2		1	2	2	1	3	2
29		1	3	3	1	2	1		3	3	2	1	3	1
30		1	2	3	2	1	2		1	1	2	2	3	1
31		2	1	2	1	1	2		1	1	2	2	3	1
32		3	3	1	3	2	2		2	3	2	3	3	2
33		2	1	3	2	2	2		3	1	2	1	1	2
34		2	1	3	2	3	1		2	1	1	2	2	2
35		2	1	2	2	3	1		3	1	1	2	3	2
36		2	1	2	1	3	1		2	1	2	1	1	2
37		2	2	1	1	3	2		1	1	1	3	1	2
38		2	1	3	2	3	2		3	2	2	1	3	2
39		3	1	3	2	2	2		2	1	1	3	3	1
40		3	3	1	2	1	1		1	2	2	2	1	2
41		3	2	2	3	2	2		2	2	1	3	1	2