COTTER, KATHERINE N., Ph.D. Mental Control of Musical Imagery: Combining Behavioral and Experience-Sampling Approaches. (2020)
Directed by Dr. Paul J. Silvia. 115 pp.

Mental control of musical imagery is a complex, but understudied, process that consists of two components: *initiation*—whether the musical imagery experience began voluntarily or involuntarily—and *management*—whether instances of control occur after the experience has begun (e.g., changing the song, stopping the experience). The present research examined these two components of mental control using both behavioral lab-based musical imagery tasks and self-reports of mental control in daily life using experience sampling methods. Both music students and members of the general university community participated. This project had four primary aims: (1) examining the relationship between initiation and management of musical imagery; (2) assessing how mental control abilities differ as a function of stimulus type; (3) describing perceptions of initiation and management in daily life; and (4) evaluating how well performance on lab-based behavioral tasks aligns with self-reported mental control in daily life. The findings suggest that initiation and management abilities are closely related, people perform equivalently when asked to control tonal stimuli and song stimuli, people generally report the ability to control musical imagery in daily life, and self-report and behavioral assessments of mental control of musical imagery show a modest association. These findings have implications for current understandings of control of musical imagery and identify several avenues for future research.

MENTAL CONTROL OF MUSICAL IMAGERY:

COMBINING BEHAVIORAL AND

EXPERIENCE-SAMPLING

APPROACHES

by

Katherine N. Cotter

A Dissertation Submitted to
the Faculty of The Graduate School at
The University of North Carolina at Greensboro
in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Greensboro
2020

Approved by

____
Committee Chair

APPROVAL PAGE

This dissertation written by KATHERINE N. COTTER has been approved by the following committee of the Faculty of The Graduate School at the University of North Carolina at Greensboro.

Committee Chair ______________________________

Committee Members ______________________________

____

____

____
Date of Acceptance by Committee

____
Date of Final Oral Examination

ACKNOWLEDGMENTS

I would like to thank my advisor, Paul Silvia, and doctoral committee members, Michael Kane, Timothy Johnston, and Dayna Touron, for their assistance and comments. I would also like to acknowledge Psi Chi, the International Honors Society in Psychology, for providing funding for this project and Starbucks for keeping me well-caffeinated while completing my dissertation.

TABLE OF CONTENTS

Page

LIST OF TABLES v

LIST OF FIGURES vi

CHAPTER

I. INTRODUCTION 1

II. METHOD 19

III. RESULTS 36

IV. DISCUSSION 49

REFERENCES 65

APPENDIX A. TABLES 76

APPENDIX B. FIGURES 99

APPENDIX C. MUSICAL IMAGERY EXPERIENCE-SAMPLING SURVEY 113

APPENDIX D. INDIVIDUAL DIFFERENCE MEASURES 115

LIST OF TABLES

Page

Table 1. Descriptions of Auditory Imagery Tasks 76

Table 2. Details for Frequently Played Song Stimuli 78

Table 3. Details for Infrequently Played Song Stimuli 82

Table 4. Auditory Imagery Task Exclusions 85

Table 5. Descriptive Statistics for Imagery Tasks and *t*-tests Testing Task Performance against Chance Levels 87

Table 6. Correlations among Imagery Tasks 88

Table 7. Descriptive Statistics and Correlations for Daily Life Musical Imagery Items 89

Table 8. Correlations Between ESM Responses and Lab Imagery Control Tasks 91

Table 9. Correlations Between ESM Responses and Latent Growth Curve Factors 92

Table 10. Performance on Lab Musical Imagery Tasks Predicting Self-Reported Mental Control Abilities in Daily Life 93

Table 11. Communities and Hybrid Centrality for Lab Auditory Imagery Tasks 96

Table 12. Communities and Hybrid Centrality for Between-Person Experience-Sampling Network 97

Table 13. Communities and Hybrid Centrality for Lab Auditory Imagery Tasks and Experience-Sampling Network 98

LIST OF FIGURES

Page

Figure 1. Power Curves based on Monte Carlo Simulations 99

Figure 2. Sample Trials of Tonal Auditory Imagery Tasks 101

Figure 3. Distributions of Auditory Imagery Task Performance (Proportion of Trials Answered Correctly) 102

Figure 4. Confirmatory Factor Analyses of Auditory Imagery Tasks 103

Figure 5. Latent Growth Curve Analyses of Auditory Imagery Tasks 104

Figure 6. Intraclass Correlation Coefficients for Experience-Sampling Items 105

Figure 7. Within-Person Distributions of Experience-Sampling Mental Control Ability Items 106

Figure 8. Between-Person Distributions of Experience-Sampling Mental Control Ability Items 107

Figure 9. Within-Person Distributions of Experience-Sampling Mental Control Difficulty Items 108

Figure 10. Between-Person Distributions of Experience-Sampling Mental Control Difficulty Items 109

Figure 11. Network of Auditory Imagery Tasks (Clusters are indicated by Node Color) 110

Figure 12. Network of Experience-Sampling Items at the Between-Person Level (Clusters are indicated by Node Color) 111

Figure 13. Network of Auditory Imagery Tasks and Experience-Sampling Items at the Between-Person Level (Clusters are indicated by Node Color) 112

CHAPTER I

INTRODUCTION

Music engagement is pervasive in modern society—we subscribe to music-listening services, star in mini-concerts during our showers and commutes, and are bombarded by upbeat, bopping tunes when getting our groceries. But our musical experiences are not limited to those in our external environment—we also hear music in our "mind's ear." Musical imagery in its broadest sense can be described as hearing music in one's head that is not simultaneously present in the environment (Bailes, 2007; Cotter, Christensen, & Silvia, 2019). People report hearing musical imagery often in their everyday lives (approximately 25% of the time; Bailes, 2006; Cotter et al., 2019; Liikkanen, 2011).

Musical imagery is a dynamic, complex phenomenon. In some experiences, people only imagine select components of the music, such as the melody or vocals; in others, people report experiencing more subtle components of the music, such as harmonic lines or the timbres of different instruments (Bailes, 2007). Further, these experiences need not be solely auditory. In many cases, people's musical imagery experiences are multimodal and include visual or kinesthetic imagery (e.g., Bowes, 2009) or involve moving or humming to the imagined music (e.g., Cotter et al., 2019; Floridou, Williamson, Stewart, & Müllensiefen, 2015). Musical imagery can be embedded in rich internal narratives, such as envisioning yourself performing in a desired role (Bowes, 2009), or echo your current state of mind and personal concerns (Floridou et al., 2015).

One multi-dimensional model of musical imagery (Cotter et al., 2019) identifies five dimensions of everyday musical imagery experiences: affective valence, repetitiveness, vividness, length, and mental control. Research has revealed considerable variability in people's musical imagery experiences, but it has also found some common themes. For example, people report having some negative experiences (e.g., Liikkanen, 2011; Williamson & Jilka, 2014), but most musical imagery episodes are positive (e.g., Beaman & Williams, 2010; Beaty et al., 2013; Cotter et al., 2019). Similarly, research has demonstrated substantial within-person and between-person variability in the repetitiveness (e.g., Cotter et al., 2019; Kvavilashvili & Anthony, 2012; Margulis, 2014), vividness (e.g., Bailes, 2007, 2015; Campbell & Margulis, 2015; Cotter et al., 2019), and length (e.g., Beaman & Williams, 2010; Brown, 2006; Cotter et al., 2019; Lipson, 2006) of musical imagery experiences in everyday life.

Most research on musical imagery in daily life has emphasized describing the *what* of the experience—what song, what trigger, what valence, what episode length. Less attention has been given to *how* these experiences unfold, change, and stop. In particular, the *mental control* of musical imagery—an intriguing process involved in starting, stopping, shaping, elaborating, and maintaining musical imagery—is complex, nuanced, and relatively understudied. Mental control of musical imagery, instead of being a unitary construct, can be broken down into two distinct components—initiation and management. *Initiation* refers to how the episode of musical imagery begins—was it started on purpose or did the music appear spontaneously? *Management* refers to attempts to control the musical imagery episode after it has begun and can take different forms (e.g., altering components of the music, sustaining the experience in the face of distraction, stopping the entire experience). When thinking about control in this manner, it becomes evident that the same episode of musical imagery can be controlled in one way but involuntary in another.

By viewing mental control of musical imagery as multi-faceted, we can reflect on what the field has already examined and propose candidates for future research. Additionally, when using this lens of mental control to re-evaluate prior research, it's evident this framework provides new ways to organize and interpret what we know about mental control of musical imagery and demonstrates that seemingly different musical imagery experiences have more in common than they first appear to. Further, by introducing a common language with which to describe these mental control processes, we can better articulate what we already know, develop research questions that arise once operating within this framework, and refine our assessment of mental control.

Approaches in Musical Imagery Research

Four perspectives on musical imagery have emerged in the literature, but they have not directly tackled the issue of its mental control (see Cotter, 2019, for additional review). The first approach—the cognitive psychology of auditory imagery—emphasizes the use of behavioral tasks to assess imagery abilities and treats musical imagery as one instance of auditory imagery. This research has often used simple tonal stimuli to investigate principles of people's auditory imagery experiences. These tonal auditory imagery studies assess a range of people's auditory imagery capabilities, including information about their abilities to initiate and manage musical imagery.

The other three approaches emphasize musical imagery experiences that often occur in everyday life. The involuntary musical imagery approach emphasizes experiences that are "spontaneous" and "uncontrolled" (e.g., Liikkanen, 2008, 2011; Floridou et al., 2015). In general, these experiences are involuntarily initiated (Williams, 2015), though some researchers also state that involuntary musical imagery cannot be managed (e.g., Beaman & Williams, 2010, 2013; Floridou & Müllensiefen, 2015; Floridou et al., 2015; Jakubowski, Bashir, Farrugia, & Stewart, 2018).

Another approach focuses on musicians' uses of imagery as a rehearsal and composition tool. Unsurprisingly, musicians frequently use imagery techniques to enhance their performances and compositions. The rich qualitative tradition of this literature suggests that musicians purposefully use musical imagery to improve the technical or stylistic aspects of a piece (Bowes, 2009; Fine, Wise, Goldemberg, & Bravo, 2015; Gregg, Clark, & Hall, 2008; Holmes, 2005; Saintilan, 2014). Additional quantitative studies show that using mental rehearsal has been associated with the ability to more quickly memorize a new piece of music (Rubin-Rabson, 1941), reduce errors in performance (Bernardi, De Buglio, Trimarchi, Chielli, & Bricolo, 2013; Wöllner & Williamon, 2007), and improve confidence (Johnson, 2011).

The final approach—ecological musical imagery—examines musical imagery in-the-moment as it occurs in people's everyday lives. Researchers in the ecological musical imagery tradition tend to take a descriptive, exploratory approach: they seek to describe what people's everyday musical imagery experiences are like (e.g., Bailes, 2007, 2015; Beaty et al., 2013; Cotter et al., 2019). The studies that take this approach capture involuntary, voluntary, and creative musical imagery experiences and discuss musical imagery as a general phenomenon experienced by musicians and non-musicians alike.

The four approaches employ a range of assessment techniques and emphasize different flavors of musical imagery experiences. Research has used both behavioral and self-report measures of mental control, and past work suggests that people are generally pretty good at controlling their musical imagery (e.g., Bowes, 2009; Cotter & Silvia, in press; Foster & Zatorre, 2010; Holmes, 2005; Janata & Paroo, 2006). But there are several ways in which these four perspectives differ in their approaches to mental control. For the purposes of this project, I focused on the auditory imagery and ecological musical imagery approaches.

Cognitive psychology of auditory imagery. Auditory imagery research, rooted in cognitive psychology, has often used tonal stimuli in its lab-based paradigms to investigate the principles of people's auditory imagery experiences—this section focuses only on studies using musical stimuli. These tonal-based auditory imagery studies assess a range of people's auditory imagery capabilities, from simple imagery-assisted pitch discrimination to complex transformations of melodies. Although this literature does not formally discuss mental control, the natures of the tasks do provide information regarding people's initiation and management abilities. Table 1 provides descriptions of the task paradigms used in auditory imagery research.

Initiating musical imagery. Inherent in any auditory imagery task is the need to construct a mental image. In early work, the imagery tasks were relatively simple—imagining the pitch of a presented tone and completing a signal detection task (Farah & Smith, 1983). These results suggest that people can form images of single tones at will, and these images facilitate auditory perception via a reduced detection threshold for imagined pitches as compared to non-imagined pitches. In Pitch Discrimination tasks (see Table 1), participants imagine specified tones, chords, or short passages of music (e.g., musical scales, simple melodies) and assess whether auditory probes match the pitch of their constructed image (Herholz et al., 2008; Janata & Paroo, 2006). On average, people can form the requested images with reasonable accuracy for single tones and chords (60-95% correct; Hubbard & Stoeckig, 1988), musical scales (78% correct when probe in tune; Janata & Paroo, 2006), and simple melodies (60 and 87% correct for non-musicians and musicians, respectively; Herholz et al., 2008). Collectively, the literature suggests that people can, when instructed, initiate a variety of simple musical images.

Sustaining musical imagery. Several studies also assess people's ability to manage their established images, such as deliberately sustaining the image—Hubbard's (2018) recent review of auditory imagery suggests this may be an overlooked dimension of control. In several Pitch Discrimination and Timing Judgment studies (see Table 1), participants hear the first few notes of a musical passage and imagine the remainder of the passage to determine whether a subsequent probe tone matches the pitch or timing of their imagined music (Bailes & Bigand, 2004; Herholz et al., 2008; Janata & Paroo, 2006; Weir et al., 2015). In one Timing Judgment study, participants were instructed to imagine the continuation of music for as long as possible and, when they were no longer able to continue the imagined music, to "check in" with the actual progression of the song by raising the volume of the stimulus song (Bailes & Bigand, 2004). The results indicated that the check-ins were related to structural properties in the music, suggesting that people can sustain images of sections of music, but when the piece shifts to a new section, people have difficulties imagining these transitions.

Other sustention work uses Temporal Accuracy tasks (see Table 1), which require participants to indicate when their image of a designated musical passage has reached the end (Halpern & Zatorre, 1999) or when a specific point in the passage is reached (Halpern, 1988). In Lyric Comparison studies (see Table 1), people are presented with two lyrics from a well-known tune (e.g., "Happy Birthday") and are asked which of two lyrics has an associated note higher in pitch (Aleman, Nieuwenstein, Bocker, & Haan, 2000; Zatorre & Halpern, 1993). In these basic sustention studies, people can maintain short images of familiar tunes (Aleman et al., 2000; Herholz et al., 2008; Weir et al., 2015; Zatorre & Halpern, 1993) and musical scales (Janata & Paroo, 2006) to perform the necessary Pitch Discrimination and Timing Judgments.

Researchers have also used more complicated sustention tasks that involve continuous monitoring of an image. A more complex Pitch Discrimination task involved listening to a simple melody and judging whether the subsequently presented notation matched the heard melody (Bailes, Bishop, Stevens & Dean, 2012). To evaluate similarity, participants needed to generate an image of the notation and monitor their image for deviations from the target melody previously heard—on average, participants were accurate 70% of the time. Additionally, Contour Tracking work (see Table 1) finds that people can monitor changes in pitch across a musical passage via reporting whether a pitch is higher or lower than the preceding pitch (Weber & Brown, 1986).

In Loudness Profile studies (see Table 1), participants listen to passages of music, paying special attention to changes in loudness throughout the piece. They then imagine the musical passage and indicate the dynamic contour of the piece using a slider when listening to and imagining the music (Bailes et al., 2012; Bishop, Bailes, & Dean., 2013). People were able to produce a dynamics profile of their imagery similar to the profile generated when listening to the same musical passage.

Other studies using Tempo Judgment paradigms (see Table 1) ask participants to listen to or imagine specific pieces of music and indicate what they believe to be the correct tempo (Jakubowski, Farrugia, & Stewart, 2016; Jakubowski, Halpern, Grierson, & Stewart, 2016). Unsurprisingly, people are most accurate when listening to a song (Jakubowski et al., 2016), but in both studies people were able to sustain their image to complete the tasks. Collectively, the auditory imagery literature demonstrates people's ability to sustain a musical image and suggests that, in addition to making single, isolated judgments about their musical imagery (i.e., pitch discrimination, timing accuracy), people can also monitor its temporal qualities.

Manipulating musical imagery. Although sustaining musical imagery is one example of management, the more intuitive sense of management is the ability to manipulate and alter aspects of an image. In one Pitch Manipulation study (see Table 1), participants were presented with a single tone or chord and asked to imagine the tone of chord one step higher—their altered image was then probed for accuracy (on average 60-95% correct; Hubbard & Stoeckig, 1988). In a more complex Pitch Manipulation study, participants were presented with the first few notes of an ascending or descending scale and imagined subsequent notes that were higher or lower in pitch as specified via up or down arrows (Gelding, Thompson, & Johnson, 2015). After imagining multiple notes, a probe tone was presented for a pitch discrimination judgment to assess the accuracy of images. Musicians tended to be more accurate than non-musicians (82 vs. 76% accuracy, respectively).

Researchers have also examined people's ability to perform complex mental manipulations using a Melody Transformation task (see Table 1; Foster & Zatorre, 2010; Foster, Halpern, & Zatorre, 2013). Musicians were presented with a target melody and needed to determine whether the test melody was the same as the target melody or if a pitch had been altered. The test melody, however, was presented in one of three forms: reversed (i.e., the melody was presented from the end to beginning), transposed (i.e., the melody was presented in a different key), or control (i.e., the melody was not altered). To determine whether the test and target melodies were identical, participants needed to mentally transform the test melody to be in the same key or temporal order as the target for comparison. Unsurprisingly, people were most accurate when presented with control melodies (between 76 and near 100% accuracy) and were less accurate when presented with transposed (69-90% accuracy) and reversed melodies (80% accuracy; Foster & Zatorre, 2010; Foster et al., 2013). These findings suggest that manipulations people make to their musical imagery can vary in complexity and difficulty.

Ecological musical imagery. The ecological musical imagery approach, in contrast to the auditory imagery approach, uses ecological momentary assessment techniques to explore and describe people's everyday musical imagery experiences. Experience sampling methods (ESM), the most frequently used technique, collect probe-caught musical imagery experiences as they are happening via completion of multiple surveys per day across several days at random time intervals. This method provides researchers with a measure of control over their data collection in people's everyday lives by determining, for example, when people can complete surveys, and how frequently they are probed. This approach also preserves differences between episodes that can be obscured when using other self-report methods, such as retrospective surveys or interviews, that require respondents to pool their musical imagery experiences (see Cotter & Silvia, 2017, for additional details).

Mental control, however, has not been a prominent focus. Most studies using this approach have not differentiated between involuntary and voluntary instances of musical imagery, although some studies have asked questions alluding to people's ability to exert control over their musical imagery. In daily life, people do not report frequently initiating musical imagery (Beaty et al., 2013; Bailes, 2015; Cotter et al., 2019)—when asked if they started an episode of musical imagery on purpose, people report doing so approximately 25% of the time (Cotter et al., 2019). Interestingly, when people are asked to initiate an episode of musical imagery in everyday life, both musicians and non-musicians report the ability to do so most of the time (61% of probes; Cotter & Silvia, in press), and all participants reported the ability to initiate musical imagery in at least one probe in the study. Even though not reported as the dominant way musical imagery begins, people do report initiating musical imagery occasionally in everyday life and report the ability to do so when asked.

Researchers have also assessed people's perceptions of managing their musical imagery—many of these items involve wanting to get rid of or alter the content of an episode. For instance, some work has asked if people wish the imagery contained different music (Bailes, 2007, 2015) or if they wanted the episode to end (Bailes, 2007, 2015; Beaty et al., 2013). These items do not directly assess management, but endorsement of these statements indirectly implies management failure. Although reporting of responses to these items was limited, people did not strongly endorse these statements (Bailes, 2007; Beaty et al., 2013), implying that management failure is not the norm. Indeed, when asked, people reported moderate levels of control over their imagery (Cotter et al., 2019).

One study has also directly investigated perceptions of management ability (Cotter & Silvia, in press). In this study, participants were asked to perform five manipulations to their musical imagery—changing the tempo, key, vocalist's gender, primary instrument, and entire song. Participants reported the ability to perform the manipulations between 47 and 72% of the time. People reported the least success in changing the key of the music and the most success in changing the song. Unsurprisingly, people with greater musical expertise reported a more frequent ability to perform all manipulations. Consistent with the findings from the auditory imagery literature, people reported the ability to manage their musical imagery, but there were instances when they failed.

Bridging the Gap: Lab-based Auditory Imagery and Ecological Musical Imagery

Given the considerable differences between these two approaches, a natural step for future research is to apply the lessons learned from one approach to the other. The lab-based auditory imagery approach has many strengths. First, this approach emphasizes behavioral over self-report measurement of mental control. Given limitations in people's understanding and reporting on their cognitive processing (Nisbett & Wilson, 1977), behavioral measurement of processes like mental control provides greater clarity and validity than other measurement techniques. Additionally, the lab-based paradigms of the auditory imagery approach provide researchers with a fine level of control over the stimuli, tasks, and research environment. Collectively, this allows researchers to isolate specific mechanisms within people's mental control processes (e.g., controlling or monitoring pitch-related versus timing-related aspects of their imagery) and examine people's control abilities under optimal conditions (e.g., reduced distractions, minimal environmental noises).

But there are drawbacks to this approach. In classic auditory imagery studies, the stimuli are single tones, chords, or simple tonal sequences (e.g., Farah & Smith, 1983; Hubbard & Stoeckig, 1988; Janata & Paroo, 2006); more recent studies have used both simple (e.g., Foster et al., 2013; Gelding et al., 2015) and more complex stimuli (e.g., Bailes et al., 2012; Weir et al., 2015). But the considerable heterogeneity and idiosyncratic nature of everyday musical imagery has not been captured in the lab-based auditory imagery literature. Often everyday musical imagery contains familiar songs (Liikkanen, 2008, 2011; Williamson & Jilka, 2014), and people tend to report imagining music that they have heard recently (Bailes, 2015; Williamson et al., 2011). Given the ability to selectively shape the music we listen to (e.g., using personalized music playlists, music streaming services, etc.), the diversity of recently heard music is immense. Further, musicians will also use musical imagery as a rehearsal tool (Bowes, 2009) or to compose original music (Bailes, 2009; Bailes & Bishop, 2012).

The ecological musical imagery approach has the opposite character. Its primary strength is in assessing musical imagery in people's messy and chaotic everyday environments and capturing people's typical musical imagery contents. In some regards, by capturing the breadth of people's environments and imagery contents, the ecological musical imagery approach is better equipped to assess people's ability to control their musical imagery. There are limitations that reduce the ability to draw conclusions, however. Its descriptive and self-report nature does not provide the same clarity and validity as the lab-based auditory imagery studies (see Hubbard, 2013, 2018). Because the everyday musical imagery literatures largely rely on self-report measures (as opposed to the behavioral methods favored in auditory imagery research), it is important to evaluate the validity of such reports.

Nisbett and Wilson (1977) are skeptical of people's ability to report on their higher-level cognitive processes. They demonstrate through a series of their own studies and previously published work that people are often unable to accurately recognize that their decisions were influenced by external experimenter manipulations and instead attribute their decisions to their own internal processes. People confidently and readily supply these mistaken attributions, leading Nisbett and Wilson to claim that people cannot accurately introspect on these processes and thus that self-reports assessing such processes are inherently invalid. In some cases, people might rely on idiosyncratic theories of whether they should be able to control their imagery. For example, people with greater musical expertise may inflate their reported ability to control their musical imagery due to beliefs that they should be able to do so. Indeed, the exclusive use of self-reports of mental control are best understood as representing people's subjective perceptions of control rather than their objective ability to control these experiences.

Nisbett and Wilson (1977) do note, however, that their findings do not preclude the possibility that people *could* provide accurate reports of their internal processes:

> We also wish to acknowledge that the studies do not suffice to show that people *could never* be accurate about the processes involved. To do so would require ecologically meaningless but theoretically interesting procedures such as interrupting a process at the very moment it was occurring, alerting subjects to pay careful attention to their cognitive processes, coaching them in introspective procedures, and so on. (p. 246, italics in original)

This passage suggests that methods such as ESM and other ecological momentary assessment techniques may address some limitations of the studies Nisbett and Wilson review. Furthermore, Hurlburt and Heavey (2001) state that people may not be fully cognizant of their fleeting thoughts and internal experiences when asked about them after they have happened but are capable of introspecting and reporting on these experiences as they are happening (see Cotter & Silvia, 2017, for a comparison of retrospective and ecological measurement techniques in musical imagery). This suggests that in-the-moment assessment may be better equipped to capture such internal experiences. Additionally, the time estimation literature (see Block & Gruber, 2014; Block & Zakay, 1997 for review) suggests that being aware that you will be asked to make a judgment of your perceptual experience prior to making it yields more accurate judgments. In this literature, prospective time estimation tasks (knowing you will be estimating the length of a subsequent interval) yield more accurate estimates than retrospective tasks (being asked to estimate the length of an interval after it has passed without prior knowledge you would make this judgment). This suggests that telling people they will be asked to make control judgments later in the study and using ecological momentary assessment techniques to capture these cognitive processes as close to the experience as possible may result in more valid measurement of mental control than other self-report assessments.

Drawing on the strengths of each approach would benefit both literatures without compromising their respective focuses. Incorporating more complex stimuli in auditory imagery studies, such as music similar to what is heard in everyday life (e.g., pop songs; Bailes, 2015), could be used to objectively evaluate people’s control abilities with greater ecological validity. In fact, some researchers have begun using more ecologically-valid stimuli in lab-based work (e.g., Bishop et al., 2013; Godøy, Haga, & Jensenius, 2006; Jakubowski et al., 2015; Weir et al., 2015) and found that people are able to generate relatively accurate musical images of more complex stimuli. Conversely, ecological studies would benefit from adapting behavioral assessments used in lab-based research to increase the validity of reports. Indeed, a few studies have begun to integrate behavioral and ecological assessment (e.g., singing involuntary musical imagery episodes into a recorder, McNally-Gagnon, 2016; recording tempo of voluntary and involuntary musical imagery via tapping the beat, Jakubowski et al., 2018).

The Present Research

In the present research, I combined both lab-based behavioral and ESM assessments to obtain a multi-faceted understanding of mental control of musical imagery. Participants first completed 11 behavioral musical imagery tasks used in past auditory imagery research or adapted for the present study. After completing the lab session, participants then completed one week of ESM in which they reported their ability to initiate and manage their musical imagery in their everyday lives.

This project has a number of strengths that address limitations inherent in past lab-based auditory imagery and ecological musical imagery research. To address a large limitation of the lab-based auditory imagery approach, the present research used both traditional auditory imagery tasks with relatively simple tonal stimuli and tasks in which stimuli consisted of excerpts from frequently played and popular songs similar to the musical imagery contents reported in ecological musical imagery research. To address the primary limitation of the ecological musical imagery approach, the present research also examined the validity of people’s self-reported mental control abilities. Participants completed a week of ESM in which they were asked to initiate and manage their musical imagery in several ways (e.g., alter the musical key and tempo). Performance on the behavioral imagery tasks completed in the lab were used to predict self-reported control abilities in daily life to examine concordance between the objective behavioral measurements and the subjective self-reports. Additionally, this design provided the ability to evaluate whether performance on tasks with stimuli that more closely resemble musical imagery contents in daily life, as compared with tasks with simple tonal stimuli, better predicted self-reported mental control in daily life.

A sample with diverse musical backgrounds was recruited, including music novices, people with general musical training, and music students who may be more likely to use musical imagery due to their specialized musical goals. At the within-person level (i.e., Level 1), people were asked to initiate and perform manipulations on their musical imagery during one week of ESM data collection. At Level 2, the between-person level, people completed imagery tasks and a battery of individual difference measures—the individual difference measures were not analyzed for the purposes of the present study. The auditory imagery tasks included two initiation tasks and seven management tasks used in prior lab-based auditory imagery research and four auditory imagery management tasks adapted from tasks used in prior research for use with ecologically-valid song stimuli.

The present research expands the current literature in four ways. First, it is the first study to behaviorally assess both initiation and management ability in the same study. Even though the distinction between these two components in the cognitive psychology literature is evident, no prior research has examined how performance on initiation and management tasks are associated. Second, the present research examined whether and how the complexity of musical stimuli influences control abilities. Only one previous study (Weir et al., 2015) has used song stimuli to assess control of musical imagery, and no studies have directly compared control abilities with tonal and song musical stimuli. It may be the case that increased complexity impairs control performance, that the similarity between the song stimuli and people's typical musical imagery contents bolsters control abilities, or that the complexity of stimuli is irrelevant to control abilities—the present research addressed this question. Third, the present research explored people's perceptions of their ability to initiate and manage musical imagery in their everyday lives. Finally, the present research assessed how well self-report ESM measures of control relate to behavioral measures of control. Because ESM is often used to assess people's musical imagery experiences and has a number of advantages over other measurement techniques, it is important to understand the validity of these reports.

CHAPTER II

METHOD

Participants

Sample size determination. The number of participants was determined through Monte Carlo simulations conducted in Mplus 8. Given the complexity of calculating power for multilevel analyses, I varied three important components across simulations: the magnitude of the standardized regression weight (i.e., β_{01j} and β_{02j}), the number of Level-1 units (i.e., number of ESM surveys completed), and the number of Level-2 units (i.e., number of participants). Three values of standardized regression weights (.15, .20, .30), four values of Level-1 units (10, 15, 20, or 25 surveys completed), and four values of Level-2 units (50, 100, 150, or 200 participants) were simulated, resulting in a total of forty-eight simulations (see Figure 1 for power curves). A standardized regression weight equal to .15 is the smallest theoretically interesting effect—simulations using this value were of primary interest; however, larger regression weights were also considered should simulations involving a regression weight of .15 indicate a sample size that could not reasonably be obtained for this project. In all simulations, the intraclass correlation coefficient (ICC) of the outcome variable was set to a value typical of past ESM work (i.e., between .30 and .50; Cotter et al., 2019; Cotter & Silvia, in press), and the correlation between the two predictors was fixed to .50 as a conservative estimate because power decreases as predictors are more closely associated. All simulations were run with 1,000 replications and with the following model:

$$\text{Level 1: } Control_{ij} = \pi_{0ij} + e_{ij}$$

$$\text{Level 2: } \pi_{0ij} = \beta_{00j} + \beta_{01j}IV1_j + \beta_{02}IV2_j + u_{0j}$$

Based on the simulations with a standardized regression weight equal to .15, 200 participants who complete an average of 25 ESM surveys would be needed to achieve power equal to .79. This was a feasible number of participants to recruit in the designated data collection period (January 2019 through September 2019). Additionally, assuming an average of 25 completed ESM surveys per person is in line with response rates of prior research (e.g., 24.29 – 33.48; Cotter et al., 2019; Cotter & Silvia, in press). In past research, approximately 10% of participants were excluded due to poor ESM response rates or inattentiveness during the lab portion of the study. Therefore, I planned to recruit approximately 220 people to account for excluded participants.

Sample characteristics and exclusions. Two hundred and twenty-five students were recruited from the psychology research participation pool and the School of Music. To recruit music students, flyers were posted in the School of Music building and a recruitment table was set up outside the School of Music building. Music student participants were paid $50 in cash, and participants from the psychology research pool received eight research participation credits. The recruited sample included 43 music students and 182 students from other majors. Music students varied in their concentrations, but most students were studying music performance (41.86%) or music education (27.91%). Although music students had different majors, all programs of study included a music performance requirement. Participants who responded to at least 45 ESM surveys during the data collection period were entered into a raffle for one of three $40 cash prizes—12.89% of the sample qualified for raffle entry.

Three factors determined participant exclusion: scores on inattentiveness measures during the lab portion of the study, abnormal behavior during the lab portion of the study, and total number of ESM surveys completed during the data collection period. Participants who failed to pass attention-check survey items and demonstrated inconsistent responding on items designed to capture inattentiveness were excluded from all analyses (see Maniaci & Rogge, 2014; McKibben & Silvia, 2016, 2017). Attention-check items (2 items) required participants to indicate a specific scale response (e.g., "Strongly Agree")—people who failed both items were excluded. Inconsistent responding items (12 items; 5-point Likert scale) were pairs of statements (e.g. "I1", "I2") that attentive participants should respond similarly to. Six items were presented at the start of the lab portion and the remaining items were presented at the end of the lab session. Absolute differences in responding for each item pair was calculated (i.e., responding a "4" on one item and "2" on its pair resulted in a difference of 2), and these difference scores were summed for all item pairs. Participants with total scores greater than 8 were excluded. Participants who failed to follow instructions during the lab session or who demonstrated a lack of care during the lab session (e.g., falling asleep, texting) were also excluded. Additionally, people who completed fewer than 5 ESM surveys, a recommended minimum for daily life research (Bolger & Laurenceau, 2013), were excluded from analyses. Using these criteria, 18 participants were excluded due to their behavior in the lab session, 20 participants were excluded for completing fewer than 5 ESM surveys, and 2 participants were excluded due to lab session behavior and completing fewer than 5 ESM surveys. After exclusions, the final sample consisted of 185 participants—the final sample was young (*M* age = 19.52, *SD* = 2.45, *range* = 18 to 36), predominately female (119 female, 63 male, 3 unreported), and racially diverse (48.57% White, 45.41% Black, 4.32% American Indian or Alaska Native, 3.24% Asian, 2.16% Native Hawaiian or Other Pacific Islander, 8.65% Declined to State; note that participants were permitted to select multiple options). The final sample consisted of 36 music students and 149 students from other fields.

Apparatus

All imagery tasks were completed on Lenovo computers and with Sony MDR-ZX110NC headphones. Prior to completing any imagery tasks, participants adjusted the volume on the headphones to a comfortable level.

Auditory Imagery Tasks

People completed 11 auditory imagery tasks that varied in several ways. First, tasks varied in the stimulus type—in the tonal auditory imagery tasks, stimuli consisted of single tones, chords, and major scales; in the song auditory imagery tasks, stimuli consisted of excerpts from songs by well-known artists. Second, tasks required people to make different kinds of judgments about their imagery—people made pitch (seven tasks), timing (two tasks), tempo (one task), and rhythm (one task) judgments based upon their musical imagery. Finally, people completed tasks assessing both initiation (two tasks) and management (nine tasks) of musical imagery.

Tonal auditory imagery tasks. All stimuli were created using the MuseScore 2 composition software, notes were synthesized as piano tones, and .WAV sound files were created for subsequent editing. All pitch and timing editing was completed manually using Audacity. Tasks were programmed using Direct RT, and both choice performance (correct/incorrect) and reaction times (in ms) were recorded for all responses. Participants were instructed to focus on the accuracy of their responses, not speed of responses, and there was no time limit for responses.

Tone and chord initiation. Both the tone and chord initiation tasks were adapted from those used in Hubbard and Stoeckig (1988, Experiment 3). Stimuli consisted of the 12 pitches of the chromatic scale and the 12 major scale chords. In each trial, participants first heard a cue tone or chord for 2.3 s and were instructed to imagine the pitch(es) of the cue. After forming an image of the cue, participants pressed the keyboard space bar and heard a tone mask (2.7 s) consisting of 16 randomly-selected eighth notes followed by a probe tone or chord for 2.3 s (see Figure 2, panels A and B). The probes took one of three forms: identical pitches to the original cue, pitches sharpened (i.e., higher in pitch) by 30 cents, or pitches flattened (i.e., lower in pitch) by 30 cents (approximately 1/3 of a semitone). The degree of mistuning for the final pitches was determined by Janata and Paroo's (2006) findings that suggest this degree of mistuning is correctly identified approximately 50% of the time. Participants indicated whether the probe was higher than, the same as, or lower than the cue—probes for each trial were randomly selected from the three variants. People completed all tone trials in one block and all chord trials in a separate block. Each block contained 12 trials. Each scale pitch and major chord served as a cue once, and trials were presented in a random order. Prior to completing each test block, participants completed three practice trials to familiarize themselves with the tasks. These practice trials were identical to those in the test block except that the cue and probe stimuli were presented in a different octave than stimuli in the test blocks to control for any effects of multiple presentation of stimuli.

Pitch discrimination. This task was adapted from Janata and Paroo (2006). Stimuli consisted of diatonic scales in all 12 major keys—scales were both ascending (i.e., increasing in pitch across notes) and descending (i.e., decreasing in pitch across notes). In each trial, participants heard the first four notes of a scale and were instructed to imagine the next four notes of the scale. The final note of the scale (i.e., the eighth note one octave above the starting note) was played and took one of three values: an in-tune pitch, a pitch sharpened by 30 cents, or a pitch flattened by 30 cents (see Figure 2, panel C). The degree of mistuning for the final pitches was determined by Janata and Paroo's (2006) findings that suggest this degree of mistuning is correctly identified approximately 50% of the time. Participants indicated whether the final note of the scale was the same as, higher than, or lower than their image of the final note—the tuning of the final note was randomly selected from the three variants. People completed 12 trials (one for each major scale)—whether this was an ascending or descending scale was randomly selected for each trial (e.g., each person completed a trial with either an ascending A major scale or a descending A major scale but did not complete both). Trials were presented in a random order and each lasted for 4.2 s. Prior to completing the test trials, participants heard complete ascending and descending scales and completed four practice trials—two for ascending scales and two for descending scales. These practice trials were identical to the test trials except the scales were presented in a different octave than the scales in the test trials to control for any effects of multiple presentations of stimuli.

Timing judgment. This task was adapted from Janata and Paroo (2006). Stimuli consisted of diatonic scales in all 12 major keys—scales were both ascending and descending. In each trial, participants heard the first four notes of a scale and were instructed to imagine the next four notes of the scale. The final note of the scale (i.e., the eighth note one octave about the starting pitch) was played and took one of three forms: an in-time note, a note played 60 ms early, or a note played 60 ms late (see Figure 2, panel D). The degree of timing difference for final notes was determined by Janata and Paroo's (2006) findings that suggest this degree of timing deviance is correctly identified approximately 50% of the time. Participants indicated whether the final note of the scale was played at the same time as, earlier than, or later than their image of the final note—the timing of the final note was randomly selected from the three variants. People completed 12 trials (one for each major scale)—whether this is an ascending or descending scale was randomly selected for each trial (e.g., each person completed a trial with either an ascending A major scale or a descending A major scale but did not complete both). Trials were presented in a random order and each lasted for 4.2 s. Prior to completing the test trials, participants heard complete ascending and descending scales and completed four practice trials—two for ascending scales and two for descending scales. These practice trials were identical to the test trials except the scales was presented in a different octave than the scales in the test trials to control for any effects of multiple presentations of stimuli.

Rhythm judgment. This task was adapted from Foster and Zatorre (2010). Stimuli consisted of five-beat rhythmic patterns played on a single pitch (middle C/C4). Rhythmic patterns consisted of both played notes and rests (i.e., pauses between notes) to ensure variations between patterns. In each trial, participants heard a cue rhythmic pattern (4.3 s) followed by a 3 s silence and then heard a probe rhythmic pattern (4.3 s). Participants indicated whether the probe rhythmic pattern was the same as or different from the cue rhythmic pattern—whether the probe rhythmic pattern was the same as or different from the cue was randomly selected (see Figure 2, panel E). Probe rhythmic patterns different from cue patterns only contained one difference from the cue patterns (e.g., the length of a note or rest is altered, a rest is added). Participants completed 15 trials presented in a random order. Prior to completing the test trials, participants completed three practice trials identical to test trials—rhythmic patterns used in practice trials did not also appear in test trials.

Simple melody comparison. This task was adapted from Foster and Zatorre (2010) and Foster et al. (2013). Stimuli consisted of novel, unfamiliar five-note melodies played in the C major scale. In each trial, participants heard a cue melody (4.1 s) followed by a 3 s silence and then heard a probe melody (4.1 s). Participants indicated whether the probe melody contained one pitch that was different from in the cue melody—whether the probe melody was the same as or different than the cue melody was randomly selected (see Figure 2, panel F). Probe melodies different from cue melodies had the pitch of one note changed by one or two semitones, but remained in the key of C major, and the changed note did not alter the melodic contour (i.e., whether pitches moved up or down across the melody). The pitch of the probe melody's first note was never changed. Participants completed 15 trials presented in a random order. Prior to completing the test trials, participants completed three practice trials identical to test trials—melodies used in practice trials did not also appear in test trials.

Transposed melody comparison. This task was adapted from Foster and Zatorre (2010) and Foster et al. (2013). Stimuli consisted of novel, unfamiliar five-note melodies played in the C major scale. In each trial, participants heard a cue melody (4.1 s) followed by a 3 s silence and then heard a probe melody with pitches transposed up or down by four semitones (4.1 s). Participants then judged whether the probe melody contained exactly the same melodic contour (i.e., the same pattern of upward and downward changes in pitches) as the cue melody—whether the melodic contour of the probe melody was the same as or different from the cue melody was randomly selected (see Figure 2, panel G). Probe melodies different from cue melodies had the pitch of one note changed by one semitone, and the changed note did not change the overall melodic contour (i.e., whether the notes moved up or down in pitch) but did change the magnitude of the changes in pitch. The pitch of the transposed probe melody's first note was never changed. Participants completed 15 trials, and all trials were presented in a random order. Prior to completing the test trials, participants completed three practice trials identical to test trials—melodies used in practice trials did not also appear in test trials.

Song Auditory Imagery Tasks. Stimuli were selected by: consulting multiple sources listing popular, iconic, or well-known songs (e.g., tops songs on Spotify; rankings of iconic songs on prominent music websites, such as Billboard.com); generating songs frequently played in party or group settings (e.g., Y.M.C.A. by The Village People); generating frequently played seasonal songs (i.e., well-known Christmas songs); and consulting past research using similar tasks (e.g., Weir et al., 2015). From these sources, 64 songs were selected for piloting and were converted into .WAV files. An additional 64 songs were selected as being less frequently played songs by the same artists—selection of the infrequently played songs was done through identifying the song from the same album or time period as the well-known song with the fewest plays listed on Spotify. Thirty-two frequently played songs were randomly assigned to be used as stimuli for the pitch discrimination and timing judgment tasks, and thirty-two infrequently played songs were randomly assigned to be used as stimuli for the key change and tempo change tasks. All stimuli were piloted and the 15 most familiar songs from the pitch discrimination and timing judgment tasks (30 songs total) that did not exhibit ceiling or floor effects for accuracy were used as stimuli in the present research. The artist-matched infrequently played songs (30 songs total) were used as stimuli for the key change and tempo change tasks. Information about the songs used as stimuli in the present research is located in Tables 2 and 3. The tempo of the frequently (M = 116.60, SD = 25.28, *range* = 67.00 – 193.00) and infrequently played (M = 130.50, SD = 29.73, *range* = 82.00 – 194.00) stimuli did not significantly differ, $t(58) = -1.95$, $p = .0559$, 95% CI: [-28.16, .36].

Excerpts from all songs were manually created using Audacity. Excerpts from frequently played songs were selected to include a highly memorable or familiar part of the song (e.g., chorus, beginning of the song) and were cut in musically meaningful places (i.e., not in the middle of a word or phrase). Excerpts from select songs were originally created by Weir et al. (2015)—these are indicated in Table 2. For consistency, any excerpts originally created by Weir et al. (2015) were edited in the same manner as newly created stimuli.

Excerpts from infrequently played songs were selected to feature a less memorable part of the song (e.g., the middle lines of a verse) and were cut in musically meaningful places. The end of all song excerpts faded to silence. All pitch and tempo changes were completed using Ableton Live 10 using the "Complex Pro" algorithm to maintain the sound quality for each excerpt. Tasks were programmed using Direct RT, and both choice performance (correct/incorrect) and reaction times (ms) were recorded for all responses. Participants were instructed to focus on the accuracy of their responses, not speed of responses, and there was no time limit on responses.

Pitch discrimination. This task was adapted from Weir et al. (2015). Stimuli consisted of frequently played songs likely to be familiar to participants. In each trial, participants heard an excerpt from a song that featured what is likely to be a highly memorable or familiar part of the song (e.g., chorus, beginning of the song). Excerpts were cut in musically meaningful places (i.e., not in the middle of a word or phrase), and the average excerpt length was 28.48 s (*SD* = 5.05 s; *Range* = 20.02 – 37.30 s). People heard a portion of the excerpt (*M* = 12.88 s, *SD* = 4.28 s, *Range* = 6.79 – 19.39 s) and the music disappeared for several seconds (*M* = 9.96 s, *SD* = 1.56 s, *Range* = 8.00 – 12.70 s). During this silence, participants were instructed to imagine the song continuing. After the silence, the music returned (*M* = 5.64 s, *SD* = 1.92 s, *Range* = 2.83 – 9.65 s) and took one of three forms: the pitches of the music were unaltered, the pitches of the music were sharpened by one semitone, or the pitches of the music were flattened by one semitone. These pitch shifts were identical to those used by Weir et al. (2015). Participants then indicated whether the pitches after the silence were unaltered or if they had been shifted up or down in pitch. Participants also indicated their familiarity with the song excerpt on a four-point scale (*I have never heard this song before* to *I have heard this song many times and know it very well*). Participants completed 15 trials presented in a random order. Prior to completing the test trials, participants completed three practice trials identical to test trials—songs used in practice trials did not also appear in test trials.

Timing judgment. This task was adapted from Weir et al. (2015). Stimuli consisted of frequently played songs likely to be familiar to participants. In each trial, participants heard an excerpt from a song that featured what is likely to be a highly memorable or familiar part of the song (e.g., chorus, beginning of the song). Excerpts were cut in musically meaningful places (i.e., not in the middle of a word or phrase), and the average excerpt length was 27.67 s (*SD* = 5.11 s, *Range* = 20.77 – 34.36 s). People heard a portion of the excerpt (*M* = 12.47 s, *SD* = 4.79 s, *Range* = 5.91 – 21.57 s) and the music disappeared for several seconds (*M* = 9.21 s, *SD* = 1.10 s, *Range* = 7.60 – 11.14 s). During this silence, participants were instructed to imagine the song continuing. After the silence, the music returned (*M* = 6.00 s, *SD* = 2.23 s, *Range* = 2.97 – 11.59 s) and took one of three forms: the music re-entered on time, the music re-entered two beats early, or the music re-entered two beats late. These timing shifts were identical to those used by Weir et al. (2015). Participants then indicated whether the music entered on time, too early, or too late. Participants also indicated their familiarity with the song excerpt on a four-point scale (*I have never heard this song before* to *I have heard this song many times and know it very well*). Participants completed 15 trials presented in a random order. Prior to completing the test trials, participants completed three practice trials identical to test trials—songs used in practice trials did not also appear in test trials.

Key change. This task was designed to assess people's pitch discrimination ability, similar to the tonal auditory imagery melody comparison tasks. Stimuli consisted of infrequently played songs matched by artist to stimuli used in the song pitch discrimination and timing judgment tasks. In each trial, participants heard an excerpt from a song that featured a less memorable part of the song (e.g., the middle lines of a verse), and the average excerpt length was 5.80 s (*SD* = 0.81 s, *Range* = 4.86 – 7.21 s). People heard the cue excerpt, followed by a tone mask (3 s) consisting of 16 randomly-selected eighth notes and a probe excerpt. The probe excerpt took one of three forms: the pitches of the music were unaltered, the pitches of the music were sharpened by one semitone, or the pitches of the music were flattened by one semitone. Participants then indicated whether the pitches in the probe excerpt were the same as the cue excerpt or if they had been shifted up or down in pitch. Participants also indicated their familiarity with the song excerpt on a four-point scale (*I have never heard this song before* to *I have heard this song many times and know it very well*). Participants completed 15 trials presented in a random order. Prior to completing the test trials, participants completed three practice trials identical to test trials—songs used in practice trials did not also appear in test trials.

Tempo change. This task was designed to assess people's ability to imagine temporal aspects of music, similar to the tonal auditory imagery rhythm judgment task. Stimuli consisted of infrequently played songs matched by artist to stimuli used in the song pitch discrimination and timing judgment tasks. In each trial, participants heard an excerpt from a song that featured a less memorable part of the song (e.g., the middle lines of a verse), and the average excerpt length was 5.94 s (*SD* = 1.02 s, *Range* = 4.60 – 7.99 s). People heard the cue excerpt, followed by a 3 s silence and the probe excerpt. The probe excerpt took one of three forms: the tempo of the music was unaltered, the tempo of the music was increased by 5% of the listed beats per minute (BPM), or the tempo of the music was decreased by 5% of the listed BPM. BPM was determined through the "Sort Your Music" Spotify add-on. Participants then indicated whether the tempo of the probe excerpt was the same as the cue excerpt or if the tempo was faster or slower than the cue excerpt. Participants also indicated their familiarity with the song excerpt on a four-point scale (*I have never heard this song before* to *I have heard this song many times and know it very well*). Participants completed 15 trials presented in a random order. Prior to completing the test trials, participants completed three practice trials identical to test trials—songs used in practice trials did not also appear in test trials.

Experience Sampling

Experience-sampling apparatus. MetricWire is a smartphone application designed for mobile data collection. The ESM surveys were programmed into MetricWire, and participants received a notification when a new survey was available. Each notification consisted of a visual notice of a new survey and the phone's default sound notification for applications—participants were instructed to have sound notifications enabled for the duration of the study. After a notification for a musical imagery survey appeared, people had five minutes to begin the survey—after five minutes the survey was no longer available, and participants waited until they received another notification.

Musical imagery survey (Appendix A)**.** When signaled, people completed a survey about their musical imagery experiences. People first reported if they were hearing music in their head (*Yes* or *No*). People who were not experiencing musical imagery were asked to initiate musical imagery and report whether they were able to initiate imagery (*Yes* or *No*) and the difficulty of initiation (using a seven-point *Not at all difficult* to *Very difficult* scale). People who were already experiencing musical imagery reported whether they purposefully initiated the imagery (using a seven-point *Strongly Disagree* to *Strongly Agree* scale).

People who were able to initiate musical imagery or were hearing musical imagery when signaled were asked to complete a series of manipulations to their imagery by: changing the tempo; the key; the primary instrument (if the musical was instrumental) or gender of the vocalist (if the music had a vocal track); and switching to a different song. For each manipulation, people reported whether they could complete the change (*Yes* or *No*), and the difficulty of completing the change on a seven-point scale (*Not at all difficult* to *Very difficult*). People who were unable to initiate musical imagery when signaled completed filler questions about the quality of their thoughts to ensure the ESM survey was a similar length. Consistent with ESM practices, all participants completed a series of items about their current feelings, mood, and environment; however, these items were not of primary interest and will not be discussed further.

These manipulations were chosen for their conceptual similarity to prior musical imagery research. Past research has examined imagery for tempo (e.g., Halpern, 1992; Jakubowski et al., 2015, 2016), music in different keys (e.g., Foster et al., 2013; Hubbard & Stoeckig, 1988; Vuvan & Schmuckler, 2011), and timbres of different instruments (e.g., Crowder & Pitt, 1992), and the ability to change imagined songs to dislodge earworms (e.g., Beaman & Williams, 2010; Williamson & Jilka, 2014; Williamson, Liikkanen, Jakubowski, & Stewart, 2014). Additionally, similar manipulations have been used in past research (Cotter & Silvia, in press) and have shown variability in difficulty.

Procedure

Participants first attended a 120-minute lab session to complete the auditory imagery tasks administered through DirectRT and individual difference survey measures administered through MediaLab on desktop computers—individual differences measures (i.e., measures of personality, musical expertise, fluid intelligence, auditory imagery vividness and control) were collected for exploratory analyses and were not analyzed for the purposes of this project (see Appendix B for list of measures). The auditory imagery tasks were completed in two blocks. Participants completed all tonal stimuli tasks followed by all song stimuli tasks. People then downloaded MetricWire on their smartphones (183 participants) or borrowed a lab-owned 7” Android tablet (2 participants) with MetricWire downloaded to complete the study. Research assistants explained how to use MetricWire and guided participants through a sample survey containing all possible survey items. Participants were told to read and respond to all questions and to familiarize themselves with the app and response system; participants were not instructed to form a musical imagery to complete this practice survey.

The ESM data collection period occurred over 7 days. People were signaled 10 times per day at quasi-random times between 8 a.m. and midnight. Signals were constrained to be at least 45 minutes apart. People with low response rates after two days (i.e., fewer than 10 surveys completed) were contacted via email to ensure there were no technical difficulties. All participants were contacted after four days to ensure they were not experiencing any technical difficulties and to update them on their progress toward raffle entry. Upon completion of the study, participants who borrowed a tablet returned the tablet and were thanked for their participation. The rest of the participants were instructed to remove MetricWire from their personal smartphones and were thanked for their participation. Participants who completed at least 45 ESM surveys were entered into a raffle for one of three $40 cash prizes.

CHAPTER III

RESULTS

Control in the Lab

Trial and task exclusions. On all tasks, trials with reaction times faster than 200 ms were excluded. Because participants were told to prioritize accuracy, trials with longer reaction times were not excluded from analyses. Additionally, Song Pitch Discrimination and Song Timing Judgment trials in which participants were unfamiliar with the song (familiarity ratings of 1 or 2) were excluded, and Key Change and Tempo Change trials in which participants were familiar with the song (familiarity ratings of 3 or 4) were excluded. Participants who had fewer than 10 valid trials on a task were excluded from analyses using that task. See Table 4 for exclusion details for all tasks.

Task performance. Descriptive statistics for all tasks are in Table 5, and distributions of accuracy are presented in Figure 3; correlations between mean performance on tasks are in Table 6. Mean proportion of correct trials ranged from .37 (Tonal Timing Judgment) to .68 (Simple Melody Comparison). One-sample *t*-tests were used to determine whether performance differed from chance—performance on all tasks was above chance except the Rhythm Judgment task, which did not differ from chance (see Table 5).

Initiation and management. Our first aim was to examine the relation between initiation and management abilities and whether people show differences in these abilities. Confirmatory factor analysis in Mplus 8 was used to examine the association between performance on initiation and management tasks. Proportion of trials correct on the tone and chord initiation tasks were used as indicators for the initiation latent factor and performance on all other tasks were used as indicators for the management latent factor. Factors variances were fixed to one and all factor loadings were free to vary. In this analysis, the proportion of trials correct was converted to *T* scores ($M = 50$, $SD = 10$) to put task performance on the same scale (i.e., because the tasks vary in the chance rate, a raw score of .45 indicates performance well above chance on some tasks but below chance performance on others). This model fit the data well ($\chi^2(43) = 59.97$, $p = .044$; RMSEA = .05, 90% CI: [.01, .07], $p = .562$; CFI = .95; SRMR = .05). The two factors were highly correlated ($r = .85$, $p < .001$, 95% CI: [.72, .99]) indicating that initiation and management ability are closely related (see Figure 4, top). Because the two factors were so strongly correlated, I also examined a model in which this correlation was fixed to 1 ($\chi^2(44) = 67.39$, $p = .013$; RMSEA = .05, 90% CI: [.03, .08], $p = .381$; CFI = .93; SRMR = .05). This model exhibited significantly worse fit than the model in which the correlation between factors was unconstrained ($\chi^2_{diff}(1) = 7.42$, $p < .01$).

Latent growth curves (Duncan, Duncan, & Strycker, 2006; Preacher, Wichman, MacCallum, & Briggs, 2008) in Mplus 8 were also used to model people's imagery control abilities. This analysis yields intercept and slope factors. To estimate the intercept, regression weights for all tasks were fixed to 1. To estimate the slope, regression weights for the 2 initiation tasks were fixed to -.5, and regression weights for the 9 management tasks were fixed to .5. With this scaling, the intercept represents the sample's estimated average performance (in the *T*-score metric) across all tasks, and the slope represents the difference in performance (in standard deviation units) in initiation versus management tasks. Positive slope values indicate better performance on management relative to initiation tasks. Model fit was poor ($\chi^2(61) = 135.08$, $p < .0001$; RMSEA = .08, 90% CI: [.06, .10], $p = .004$; CFI =.79; SRMR = .18).

The intercept equaled 49.98 and its variance component (33.51) was significantly different from 0 ($p < .001$), indicating that people differ in their average performance on the imagery tasks. The slope equaled .05 and was not significantly different from 0 ($p = .915$), suggesting that, on average, people's performance on initiation and management tasks were equivalent. However, the slope's variance component (14.85) was significantly different from 0 ($p = .009$), suggesting that people vary in their relative performance on initiation versus management tasks. The intercept and slope factors were significantly correlated ($r = -.81$, 95% CI: [-1.00 -.60], $p < .001$), indicating that people who perform better on the tasks overall tend to have a smaller difference in performance between initiation and management tasks (see Figure 5, top).

Tonal and song stimuli. My second aim was to examine the relationship between people's ability to control musical imagery containing tonal versus song stimuli and determine whether people differed in these abilities. A similar CFA in Mplus 8 was used to examine the association between performance on tasks using tonal and song stimuli. Proportion of trials correct on the Song Pitch Discrimination, Song Timing Judgment, Key Change, and Tempo Change tasks were used as indicators for the song factor; all other tasks were used as indicators for the tonal factor. Factor variances were fixed to one and all factor loadings were free to vary. As before, the proportion of trials correct were scaled as *T* scores for this analysis. This model fit the data well ($\chi^2(43) = 59.81$, p = .0456; RMSEA = .05, 90% CI: [.01, .07], p = .567; CFI = .95; SRMR = .05). The two factors were highly correlated (r = .85, p < .001, 95% CI: [.73, .98]), suggesting that performance on tasks with tonal and song stimuli are closely related (see Figure 4, bottom).

Latent growth curves in Mplus 8 were used to model performance on tasks using tonal and song stimuli. The specification mirrored the prior analysis. To estimate the intercept, regression weights for all tasks were fixed to 1. To estimate the slope, regression weights for the 7 tonal stimuli tasks were fixed to -.5, and regression weights for the 4 song stimuli tasks were fixed to .5. Proportion of trials correct were converted to *T* scores. The intercept thus represents average performance (in the T metric) across all tasks, and the slope represents the average difference (in standard deviations) in performance on tasks using tonal and song stimuli. Positive slope values reflect relatively better performance on song stimuli tasks compared to tonal tasks. Model fit was poor ($\chi^2(61) = 162.27$, p < .0001; RMSEA = .10, 90% CI: [.08, .11], p < .001; CFI = .71; SRMR = .22). Because the two factors were so strongly correlated, I also examined a model in which this correlation was fixed to 1 ($\chi^2(44) = 67.39$, p = .013; RMSEA = .05, 90% CI: [.03, .08], p = .381; CFI = .93; SRMR = .05). This model exhibited significantly worse fit than the model in which the correlation between factors was unconstrained ($\chi^2_{diff}(1) = 7.58$, p < .01).

The intercept equaled 50.01 and its variance component (25.97) was significantly different from 0 ($p < .001$), suggesting that people differ in their average performance on the imagery tasks. The slope equaled .12 and was not significantly different from 0 ($p = .789$), suggesting that, on average, people's performance on tasks with tonal and song stimuli are equivalent. In addition, the slope's variance component (5.01) was not significantly different from 0 ($p = .193$), suggesting that people do not vary in their relative performance on tasks with tonal versus song stimuli. The intercept and slope factors were not significantly correlated ($r = .25$, 95% CI: [-.21, .70], $p = .280$; see Figure 5, bottom).

Control in Daily Life

Descriptive statistics. Participants completed a total of 5,104 ESM surveys. The average number of surveys completed per person was 27.59 (*SD* = 14.77, range = 5 – 67). The ESM data have two levels: the within-person level (containing repeated responses to the ESM survey) and the between-person level (containing performance on lab tasks and pooled ESM scores). Descriptive statistics for frequency of musical imagery, and self-reported control ability and difficulty, and within-person and between-person correlations of these variables are shown in Table 7. Distributions of ESM surveys at the within- and between-person levels are presented in Figures 7-10. Intraclass correlation coefficients (ICCs; estimates of the proportion of variance at the between-person level) are depicted in Figure 6. Correlations between ESM control items and lab tasks are reported in Table 8.

Multilevel models were estimated in Mplus 8 using maximum-likelihood with robust standard errors and fixed effects. Scores from both the intercept and slope of the initiation versus management and tonal versus song stimuli latent growth curve models were used to predict frequency of musical imagery and all mental control indices. Correlations between ESM items and the latent growth curve intercepts and slopes are reported in Table 9. All reported regression coefficients are standardized and are reported in full in Table 10. Standardization of regression coefficients in multilevel models is much more complex than for single-level models, especially for models with categorical variables. Therefore, coefficients for continuous outcomes (starting imagery on purpose and all difficulty items) represent standardization of both the outcomes and predictors, and coefficients for binary outcomes (imagery frequency and all control ability items) represent standardization of only the predictors. Cohen's f^2 is used as the measure of effect size for the regression coefficients (Selya, Rose, Dierker, Hedeker, & Mermelstein, 2012), with values of .02, .15, and .35 representing small, medium, and large effects, respectively (Cohen, 1988).

How often did people experience musical imagery? Of the 5,131 ESM surveys, 2,280 (44.67%) captured musical imagery episodes. The ICC for experiencing musical imagery was .19, so reporting musical imagery at any given survey is more strongly influenced by within-person factors that change throughout the day rather than stable between-person differences. Frequency of musical imagery in daily life was predicted by both the intercept ($\beta = 1.04$, 95% CI: [.61, 1.47], $SE = .22$, $p < .001$) and slope ($\beta = .70$, 95% CI: [.25, 1.14], $SE = .23$, $p = .002$) of the initiation versus management model but only the intercept (β = .37, 95% CI: [.21, .53], *SE* = .08, *p* < .001) of the tonal versus song stimuli model (slope: β = .12, 95% CI: [-.04, .27], *SE* = .08, *p* = .143). This suggests that people who perform better on the lab imagery tasks overall and people who performed better on management tasks than on initiation tasks reported more frequent musical imagery in daily life. It should be noted, however, that the relation between frequency and the initiation versus management slope is a small effect (f^2 = .07), and the relations between frequency and the two intercepts (initiation versus management: f^2 = .13; tonal versus song: f^2 = .13) approach a medium effect.

Can people initiate musical imagery? I measured initiation of musical imagery in two ways: whether people already experiencing musical imagery when signaled started it on purpose, and whether people could initiate musical imagery when asked. Consistent with past work, my first method indicated that people infrequently reported initiating their musical imagery episodes (*M* = 2.80, *SD* = 2.14)—approximately 23% of episodes were voluntarily initiated (as indicated by responding above the scale midpoint). Perceptions of initiation of pre-existing musical imagery were unrelated to both the initiation versus management latent variable (intercept: β = -.08, 95% CI: [-.59, .44], *SE* = .26, *p* = .762; slope: β = .03 95% CI: [-.50, .57], *SE* = .27, *p* = .902) and tonal versus song latent variable (intercept: β = -.15, 95% CI: [-.31, .02], *SE* = .09, *p* = .084; slope: β = .09, 95% CI: [-.09, .27], *SE* = .09, *p* = .323).

My second measure of initiation addressed whether people can initiate musical imagery when asked. Overall, people reported the ability to initiate musical imagery on 60.93% of the surveys, and people who were able to initiate did not find it difficult to do so (*M* = 2.07, *SD* = 1.46, *ICC* = .33). Self-reported initiation ability was only related to the intercept of the tonal versus song model (intercept: β = .25, 95% CI: [.06, .43], *SE* = .09, *p* = .009; slope: β = -.07, 95% CI: [-.24, .10], *SE* = .09, *p* = .428) and was unrelated to the intercept and slope of the initiation versus management model (intercept: β = -.10, 95% CI: [-.64, .35], *SE* = .27, *p* = .708; slope: β = -.36, 95% CI: [-.90, .17], *SE* = .27, *p* = .184). Initiation difficulty was unrelated to both the initiation versus management model (intercept: β = -.12, 95% CI: [-.62, .38], *SE* = .26, *p* = .644; slope: β = -.02, 95% CI: [-.50, .46], *SE* = .24, *p* = .931) and tonal versus song model (intercept: β = -.07, 95% CI: [-.24, .11], *SE* = .09, *p* = .475; slope: β = -.07, 95% CI:[-.25, .11], *SE* = .09, *p* = .429).

Can people manipulate musical imagery? I also assessed people's reports about their ability to manipulate the contents of their musical imagery in five ways—altering the tempo, key, vocalist's gender, primary instrument, and entire song—and the difficulty of making these manipulations. Overall, people most frequently reported being able to change the entire song (79.08%) and the tempo (78.52%) of their imagery. People also reported being able to change the gender of the vocalist (67.25%) and primary instrument (66.02%) most of the time and were least likely to indicate an ability to change the key of their musical imagery (57.83%). People who reported completing the manipulations did not find them to be that difficult (*M*s = 2.07 to 2.56 on a 7-point scale). Changing the key of the musical imagery were reported to be the most difficult (*M* = 2.56) whereas altering the imagery's tempo was reported to be the easiest (*M* = 2.07). The ICCs for reported management ability ranged from .40 to .47 and for reported management difficulty from .40 to .51.

Reported ability to manage imagery in daily life was consistently predicted by the intercept of the tonal versus song model (all regression coefficients were significant except for ability to change the key and song and are reported in Table 10) but inconsistently predicted by the intercept of the initiation versus management model—only reported ability to manipulate the primary instrument was predicted by the intercept. For significant coefficients, the relationship indicated that people who were better at the lab tasks overall reported a greater ability to control the qualities of the imagery in daily life. The slope of the initiation versus management model predicted the reported ability to change the key and song; the slope of the tonal versus song model was unrelated to the management items (see Table 10). People who performed better on initiation tasks reported greater ability to change the key and song. Reported difficulty was consistently predicted by the intercept of the tonal versus song model but was not predicted by the intercept of the initiation versus management model or the slopes of either model. For all significant coefficients, the relationship indicated that people who were better at the lab tasks overall reported less difficulty in controlling the qualities of their imagery. Most significant effects were relatively small (f^2s < .08), except for the associations between the intercepts and reported ability to change the primary instrument (initiation versus management: $f^2 = .15$; tonal versus song: $f^2 = .40$) suggesting that most associations between lab task performance and reported management ability and difficulty are not substantial.

Supplementary Exploratory Analyses

The auditory imagery lab tasks were intended to assess initiation and management abilities and the ability to control tonal and song stimuli. In examining the fit indices of the latent growth curve models and variability in strength in the confirmatory factor analyses loadings, this conceptual framing of the tasks does not capture the true nature of the lab tasks, likely underestimating the relations between the lab tasks and self-reports of control in daily life. To better understand the underlying structure of the tasks and their relations with self-reported control in daily life, I constructed network models for three sets of variables: (1) the auditory imagery lab tasks; (2) the ESM control ability and difficulty items; and (3) the auditory imagery lab tasks and ESM control ability and difficulty items. These networks represent the relations between items and identify distinct clusters of similar items.

Network construction. To evaluate the clusters of related items, I applied Exploratory Graph Analysis (EGA; Golino & Demetriou, 2017; Golino & Epskamp, 2017), which constructs a network model and then applies a walktrap community detection algorirthm (Pons & Latapy, 2006) to determine the number of clusters in the network. This algorithm evaluates the boundaries of clusters by identifying groups of densely connected items with few remote connections via random walks (or searches) across connections starting at each item. Thus, these clusters represent groups of items that are more interconnected with one another than they are with other items. To filter the network, I used a *lasso* approach (Epskamp & Fried, 2018; van Borkulo et al., 2018), which generates a network of partial correlations between all variables. These analyses used the EGAnet package (Golino & Christensen, 2019) in *R* (R Core Team, 2018). The thickness of lines in the networks indicate the strength of relationships; the color indicates positive (green) or negative (red) relations. Distance between items—the number of edges or connections along the shortest path to connect items—also indicates the strength of relationship. Items that are farther apart in the network would have a low correlation. I also assessed the hybrid centrality (Pozzi, Di Matteo, & Aste, 2013)—the position of each item in the network based upon its connections with other items and its relative location—of all items to understand the influence of individual items within the network using the NetworkToolbox package (Christensen, 2019) in *R*. This value estimates the centrality of items in relation to all other items, not just those within its own cluster. Items with higher centrality values tend to have many connections within and outside their cluster and can be thought of as being better representations of each latent dimension present in the network than items with low centrality values.

Auditory imagery task network. The network identified two clusters of tasks (see Figure 11 and Table 11). The first cluster contained the Tone Initiation, Chord Initiation, Tonal Pitch Discrimination, Tonal Timing Judgment, Song Pitch Discrimination, and Tempo Change tasks. Of these tasks, the Tone and Chord Initiation tasks were the most central (.89 and .85, respectively). The second cluster contained the Simple Melody Comparison, Song Timing Judgment, and Key Change tasks; the Key Change task was most central (.97). The Transposed Melody Comparison and Rhythm Judgment tasks were not connected to the other tasks and were not designated to any cluster. The two clusters were most strongly connected through the Song Pitch

Discrimination and Key Change tasks. This network suggests that the methods used to organize tasks in the latent growth curve models did not represent the interrelations among the tasks well.

ESM network. The network identified two clusters of items (see Figure 12 and Table 12). The first cluster contained all six difficulty items, and the Song Difficulty and Key Difficulty were most central (.93 and .75, respectively). The second cluster contained all six ability items, and the Speed Ability and Vocal Ability items were most central (.92 and .66, respectively). The items were relatively well-connected within their own cluster but showed weak, negative connections between the clusters. This suggests that people who reported being able to control their imagery found it relatively easy to control.

Auditory imagery task and ESM network. The network identified three clusters (see Figure 13 and Table 13). The first cluster contained all ESM control ability items, and the Key Ability and Vocal Ability items were most central (.92 and .80, respectively). The second cluster contained all ESM control difficulty items, and the Song Difficulty item was most central (.85). The third cluster contained all auditory imagery tasks, except the Transposed Melody Comparison task which was unconnected to the rest of the network and was not designated to any cluster. The most central items in the third cluster were Key Change and Tone Initiation task (.82 and .75, respectively). All items in the two ESM clusters were well-connected within their own cluster and showed negative associations between the two clusters, suggesting that people who more frequently reported controlling their imagery in daily life did not report it to be difficult to control. Similarly, the cluster of ESM ability cluster showed positive associations with the auditory imagery task cluster whereas the ESM difficulty cluster showed negative associations with the task cluster. This suggests that people who exhibited better performance in the lab reported greater success in controlling their imagery in daily life and did not find the process difficult.

CHAPTER IV

DISCUSSION

The present research examined people's ability to control their musical imagery in a series of behavioral lab tasks and self-reported control abilities in daily life. There were four primary aims of this project: (1) to compare the ability to initiate and manage musical imagery using behavioral tasks; (2) to compare the ability to control simple tonal stimuli and more complex song stimuli using behavioral lab tasks; (3) to describe the self-reported ability and difficulty of initiating and managing musical imagery in daily life via ESM reports; and (4) to examine the relationship between lab-based behavioral assessments of control and self-report ESM measures of control.

Initiation versus Management of Musical Imagery in the Lab

To compare the ability to initiate and manage musical imagery, participants completed two tasks assessing initiation ability and nine tasks assessing management ability. Prior lab-based research on control of musical imagery has examined performance on a task-by-task basis—the present research is unique in its examination of mental control of musical imagery using a latent variable approach. The latent initiation and management factors were strongly correlated (r = .85), suggesting the initiation and management abilities are closely intertwined. In the latent growth curve model, there was not an overall main effect for differences in performance on initiation versus management tasks—people performed equivalently on the tasks—but there was significant variability in performance in the tasks overall (i.e., in the intercept) and in the difference in initiation and management task performance (i.e., in the slope). This suggests that people do vary in their abilities to accurately complete the tasks and also vary in the degree of accuracy on initiation versus management tasks. Conceptually, initiation and management are viewed as distinct, but related, components of mental control (Cotter, 2019; Cotter et al., 2019); however, the correlation between these factors perhaps suggests they are inter-dependent and should not be viewed as separable ability factors that underlie this collection of lab tasks.

The association between these factors likely reflects several reasons, including the considerable structural similarity of the initiation and management tasks. Given the shared structure of all the imagery tasks—presentation of a cue to stimulate imagery and evaluating imagery through identifying deviations in a probe—it is possible that the relations between the individual tasks and the latent factors derived from those tasks are influenced by common method variance (Campbell & Fiske, 1959), which could be inflating both the zero-order correlations between these tasks and the relations between the latent factors (Johnson, Rosen, & Djurdjevic, 2010).

Given this possibility, it would be premature to conclude that initiation and management should be treated as a unitary construct, and future research should take different empirical approaches to better understand the relationship between initiation and management abilities. One direction would be to take a formal multi-trait multi-method approach (Campbell & Fiske, 1959) in which the task structure is varied in a way that is orthogonal to the initiation and management distinction. For example, Lyric Comparison tasks (see Table 1) from the auditory imagery literature do not require comparison of an image to a probe but rather comparing one portion of the image to a different portion of the image. The development of other tasks that require comparing different portions of an imagined song on relevant musicological qualities (e.g., pitch, tempo) would address a limitation of the present research. Further, a multi-trait multi-method approach would be a stronger test of distinguishing between initiation and management than the present study.

Alternatively, future research could take a bifactor modeling approach (Little, 2013; Reise, 2012). This approach, like the MTMM approach, would require developing additional initiation tasks so there are more balanced numbers of initiation and management tasks. With a broader task base, it would be possible to distinguish between individual differences in task performance attributable to an overall mental control ability and individual differences due to abilities unique to initiation and management. A bifactor approach may also reveal that initiation and management are essentially the same ability.

Nevertheless, it is important to keep in mind that while performance on initiation and management tasks may or may not reflect distinct cognitive abilities, initiation and management are distinct cognitive acts. People don't often initiate musical imagery in daily life (Bailes, 2015; Cotter et al., 2019, Cotter & Silvia, in press)—one of the major approaches in this field involves the study of involuntarily initiated musical imagery experiences. People do, however, report using a variety of management techniques (e.g., changing the contents, ending the episode) to exercise control over their imagery (e.g., Beaman & Williams, 2010; Williamson et al., 2014). But the intentions and motivations behind initiation and management likely vary. In initiating an episode of musical imagery, people are choosing to engage with this internal experience; in managing ongoing musical imagery, people are choosing to prolong their experience, change the contents or qualities of the experience, or end the experience. The underlying ability to execute these cognitive acts could be the same, but there are likely factors surrounding the decisions to initiate or manage musical imagery that distinguish these components of mental control.

Control of Tonal versus Song Stimuli Imagery in the Lab

To compare the ability to control imagery of tonal and song stimuli, participants completed seven tasks assessing control over tonal imagery and four tasks assessing song imagery. The latent tonal and song factors were strongly associated (r = .85), suggesting control abilities over imagery of tonal and song stimuli are closely related. In the latent growth curve model, there was not an overall main effect or significant variation for differences in performance on tonal versus song stimuli tasks—people performed equivalently on the tasks using the two stimuli types and did not vary in degree of performance difference on the tonal and song stimuli tasks (i.e., the slope). Like the initiation and management latent growth model, however, there was significant variability in the tasks overall (i.e., the intercept).

There are two possible reasons that I did not find a difference in performance on tonal and song stimuli tasks. First, there may be no influence of stimulus type over ability to control musical imagery. Although the stimuli in most auditory imagery studies are a far cry from the contents of typical everyday musical imagery (i.e., familiar, popular music; Bailes, 2015), the underlying processes involved in controlling imagery may be insensitive to these differences. An alternative explanation is that the relative simplicity of the tonal stimuli and the familiarity of the song stimuli made the tasks easier while the novelty of the tonal stimuli and the complexity of the song stimuli simultaneously made the tasks more difficult, resulting in equivalent observed performance.

There are benefits to using ecological stimuli to better understand control processes as they likely occur outside the lab, but one of the challenges in using this type of stimuli is the reduced experimental control over musicological factors. The songs experienced as musical imagery in daily life typically contain lyrics, have both melodic and harmonic elements, and vary in timbre and dynamics (Bailes, 2007). These factors are largely held constant in auditory imagery tasks—stimuli are often sequences of single tones played in a single timbre without any lyrics (e.g., Foster et al., 2013; Foster & Zatorre, 2010; Janata & Paroo, 2006). To more fully understand how, if at all, mental control of musical imagery with song stimuli differs from control of musical imagery with traditional tonal stimuli, we need to understand how specific musicological features aid or hinder the ability to control musical imagery. Future research should isolate and systematically vary these musicological factors to determine their unique influences on ability to control musical imagery. For instance, to examine how the presence of lyrics in song stimuli influences control abilities, researchers can identify well-known instrumental pieces (e.g., theme songs from films) and well-known lyrical pieces—

ensuring the pieces are matched for tempo, key, and other relevant musicological factors—and compare performance on tasks using the paradigms in the present research.

Control of Musical Imagery in Daily Life

To assess perceptions of control over musical imagery, participants completed one week of ESM during which they were asked about their pre-existing musical imagery and told to initiate and manipulate their musical imagery. People reported frequently experiencing musical imagery in their daily lives—approximately 45% of the time. The frequency of musical imagery in daily life is similar to the frequency (52% of signals) reported by Cotter and Silvia (in press), who used a similar participant population and an identical ESM survey; however, this frequency is higher than some other studies (approximately 25% of the time; Cotter et al., 2019; Liikkanen, 2011). In Cotter and Silvia's (in press) study, this high frequency may have been attributable to the high percentage (50%) of music students in the sample, but in the present research music students were only 19% of the final sample. It is possible that participants in the present research were sensitized to their musical imagery experiences through completion of the 11 lab tasks, which may have influenced rates of musical imagery during the subsequent week. Future research should counterbalance the order of the lab session and ESM data collection to account for this possibility.

When people were experiencing musical imagery when signaled, they typically were not purposefully initiating these experiences (only 23% of ongoing experiences were purposefully initiated) in line with prior research (Bailes, 2015; Beaty et al., 2013; Cotter et al., 2019; Cotter & Silvia, in press). However, when asked to attempt initiating an episode of musical imagery, people reported they were able to start an episode 61% of the time, consistent with Cotter and Silvia's (in press; 61% of signals) finding that people generally reported the ability to initiate imagery when asked.

People most frequently reported being able to alter the tempo of the imagery (79% of signals) and the song (79% of signals) and reported the least success in changing the key of the imagery (58% of signals). People also reported that changing the tempo of the imagery and the song were the least difficult manipulations and that changing the key or vocalist's gender in the imagery to be the most difficult. These judgments of management ability are also consistent with past research. Cotter and Silvia (in press) also found changing the tempo (71% of signals) and song (72% of signals) to be the most frequently endorsed changes and changing the key (47% of signals) to be the least frequently endorsed change. The present research found somewhat higher percentages of reported success, but the profile of which changes were more or less frequently reported was replicated in the present study. Overall, these descriptive findings suggest people often report success in controlling their imagery in a variety of ways and found exercising control over their imagery to be rather easy.

Relationship Between Behavioral and Self-Report Assessments of Control

The final aim of the present research was to assess the relationship between behavioral and self-report measures of mental control of musical imagery. To examine this relationship, the latent growth curve models' intercepts and slopes were used to predict responses to the ESM musical imagery reports. Overall, there were few strong associations between the behavioral and self-report measures of control. Higher global performance on the behavioral tasks did, to some extent, predict greater likelihood of reporting being able to initiate and manage imagery and finding controlling imagery to be easier; however, the tonal versus song stimuli intercept more consistently predicted reported control than the initiation versus management task type intercept. The strongest associations were between global behavioral task performance and reported ability to change the primary instrument of musical imagery—people with better task performance reported greater ability to change the primary instrument of their imagery.

Additionally, there were a few modest relationships between the model slopes and reported control in daily life—people who did better on initiation tasks reported greater ability to alter the key and song of their imagery in daily life. Although there were several statistically significant associations, most effect sizes were small, suggesting that people's abilities, assessed behaviorally in the lab, are at most weakly linked to their self-reported success in controlling their musical imagery in daily life.

One potential reason for the small or non-significant associations in these models is the differences in the type of manipulations people were asked to complete in the lab vs. in daily life. In the lab, participants were making only pitch or timing related judgments (i.e., whether pitches or the key were shifted; whether the timing or tempo was altered), but in daily life they were asked to complete a wider range of manipulations (i.e., altering vocalist gender, primary instrument, entire song). It is possible these differences contributed to the small and non-significant relations between the behavioral and self-report measures of control. Indeed, past work has demonstrated that people can control some aspects of their imagery (i.e., pitch) better than others (i.e., timing; Janata &

Paroo, 2006; Weir et al., 2015), and the present findings may also reflect differing abilities to control imagery in different ways. The two most closely related behavioral and self-report management assessments—key and tempo—showed inconsistent relations, however; the behavioral and self-report measures of key control ability were correlated ($r = .34$) but the two measures of tempo control ability were not associated ($r = .12$). Future work should seek to clarify associations between self-report and behavioral measures through more closely matching the contents of these measures.

In addition to the differences in the manipulations people were asked to make in the lab versus in daily life, it is likely there were substantial differences in the contents of their images. In the lab tasks, people's images were constrained by the stimuli that were presented—they had to imagine what they were told to. In daily life, however, people were not told which song to imagine and so were likely imagining a song they were familiar with (Bailes, 2015). Similarly, people had more freedom in how they completed the manipulations in daily life. When people had to imagine a different vocalist singing the song, they could choose a vocalist who was perhaps their favorite or who was particularly salient in the moment rather than imagining a researcher-specified vocalist. In daily life people were afforded greater freedom over how they controlled their imagery, and the genres, songs, and exact nature of the changes people were imagining varied more than in the lab and from person to person. These differences may have contributed to the small relations between lab and daily life measures of control. Future work could constrain the manipulations people make in daily life to be more similar to those performed in the lab (i.e., just pitch or timing related) to better assess how well behavioral and self-report measures align.

A more critical take on these findings would argue that people's judgments of their control abilities are inaccurate. Nisbett and Wilson (1977) claim that people have limited access to the true nature of their higher-order cognitive processes and rely on personal theories of these processes when introspecting. There are important differences, however, between the work Nisbett and Wilson (1977) use to support their stance and the present research. Nisbett and Wilson (1977) examined people's ability to accurately report that a stimulus existed and that the stimulus influenced their response and concluded that people are often unable to identify the stimulus, the response, or that the stimulus influenced the response. Their study assessed constructs that lack concreteness (e.g., cognitive dissonance) and processes that were not salient (e.g., adjusting attitudes about a behavior to reduce psychic discomfort). The present research, in contrast, emphasizes a series of binary responses—Are you having a musical imagery experience or not? Can you change the song to a different one or not?—and such shifts are likely more salient and concrete than the processes studied by Nisbett and Wilson (1977). This difference does not preclude the possibility of people intentionally lying in their responses; however, the associations were in the expected direction—people who performed better in the lab tasks reported being able to control their imagery more frequently in daily life—and provide preliminary evidence for validity of these self-reports. Given the considerable differences between the lab and daily life assessments, the small associations between the behavioral and self-report measures do not necessarily mean that the use of self-reports in assessing mental control of musical imagery should be abandoned.

Limitations and Future Research

ESM manipulations. One strength of the present work is the variety of manipulations people were asked to make to their imagery in daily life. It is important to note, however, that some elements of musical imagery were reported as easier to control than others. This may be because the different manipulations people were asked to complete operated on different musicological features of their imagery. Since some musicological features tend to be more salient in imagery (e.g., tempo, instrument/voice, Bailes, 2015; lyrics, timbre, Bailes, 2007), manipulations involving more salient features could be seen as easier to do than manipulations operating on other factors. In future ESM research, it is important that we address why particular manipulations are seen as easier than others—the salience or vividness of particular imagery elements are top candidates for initial exploratory work on this topic.

Additionally, people may also choose to control their imagery in specific ways (vs. being asked to do so for the study). For instance, people report occasionally trying to alter the song that is being imagined (Williamson & Jilka, 2014; Williamson et al., 2014), and participants may have practice controlling their imagery, contributing to their greater reported success of doing so in the study. No prior work has asked people to describe ways in which they tend to control their musical imagery; instead, researchers identify particular manipulations of interest (e.g., changing the song, tempo) and only assess specific instances control. When using this approach, it is possible that researchers are omitting common ways in which people control their imagery. To better understand common ways in which people attempt to control their musical imagery, researchers should take a qualitative approach and allow participants to describe the ways in which they attempt to control their imagery. Given some limitations in asking people to provide such descriptions retrospectively (Cotter & Silvia, 2017), methods like Descriptive Experience Sampling (Hurlburt & Heavey, 2001), in which people are probed throughout the day about their musical imagery and subsequently have their reports interrogated, could deepen our understanding of musical imagery in daily life.

Environmental context and mental control. A major difference between ESM and lab-based measures of control is the considerable differences between the environments. In the lab, people are seated in a quiet room wearing headphones with limited external distractions and can focus their attention on the tasks they are completing. In everyday life, there are factors in the external environment that compete for our attention and may influence our ability to control musical imagery. Prior work has not extensively examined how environmental factors—external or internal—relate to the qualities of musical imagery. Some work indicates that, in daily life, negative moods (e.g., sadness, irritation) were associated with attempting to exert more control over musical imagery in daily life and lower enjoyment of the experience, whereas positive moods (e.g., happiness, excitement) were related with greater enjoyment and vividness of the episode (Cotter, 2017). Other work suggests that musical imagery can be related to our personal worries and concerns (Floridou et al., 2015), and higher endorsement of imagery being related to personal concerns is in turn related to more frequent musical imagery (Floridou et al., 2015) or the length of a repetitive section of music (Cotter et al., 2016). Associations between musical imagery qualities—especially mental control—have not been systematically examined and should be addressed in future work.

To address the relations between environmental factors and mental control specifically, it is important to recognize a limitation in the present research that can be addressed in future work. Because all control measures in the lab were behavioral and all control measures in daily life were self-report, and thus represented perceptions of control, the present research cannot speak to how the different contexts related to control of musical imagery. Given the continually increasing technological capabilities of ESM software, it is possible to incorporate behavioral assessments, similar to those used in the lab, to understand how objective control performance may differ in daily life contexts (e.g., programming ESM applications with behavioral lab tasks). Further, it is also possible to collect non-self-report information about people's environment (e.g., sampling the degree of ambient noise in the environment) to understand how both subjectively and objectively recorded features of the environment may relate to both behavioral and self-reported assessments of control in daily life. Conversely, future research should also include self-reported measures of control, similar to those used in ESM research, to assess reported ability to control imagery in the reduced-distraction environment of the lab. Collecting both behavioral and self-report measures of control in the lab and in daily life will enable future research to comment on the relations between mental control of musical imagery and environmental context.

Control ability versus self-reported control. One of the interesting findings from the present study was the disconnect between behavioral assessments of control and the self-reported perceptions of control. When examining the network, however, it does suggest there may be more widespread association between lab-based and daily life measures of control that the latent growth curve models would suggest—future research should further explore this relationship using different analytic approaches.

But it is still likely that people are not perfectly attuned to their control abilities. Future work should explore the consequences of having relatively accurate as compared to inaccurate perceptions of control over musical imagery. A starting point could be how the degree of discrepancy between self-reported perceptions and behavioral assessments relates to other fundamental dimensions of musical imagery (Cotter et al., 2019). For instance, people who overestimate their ability to control their imagery may show a greater negative impact on the affective valence of an episode of musical imagery when they unexpectedly fail to control their imagery than people with a smaller discrepancy between their control ability and perceptions of imagery control.

Additionally, future research should explore how qualities of a musical imagery episode relate to both behavioral control ability and perceptions of control. For example, the vividness of an episode may influence perceptions or ability to control that image. The Bucknell Auditory Imagery Scale (Halpern, 2015) is a self-report measure of auditory imagery vividness and ease of control, and prior research with this scale indicates that people who report more vivid auditory imagery also tend to report being able to alter their auditory imagery with greater ease. Although this scale assesses auditory imagery in general, not just musical imagery, this may suggest that vividness is an important factor in controlling imagery. Further, Cotter et al. (2019) found that vividness of individual episodes of musical imagery in daily life were modestly associated with feelings of being able to control the episode. In this study participants were not asked to attempt to control their imagery, but this does provide some preliminary evidence that vividness and self-reported control may be related. Behavioral studies of musical imagery control have not included measures of imagery vividness, making the relation between vividness and behavioral measures of musical imagery control an open question. Studying mental control of musical imagery from both behavioral and self-report methods to assess control ability and perceptions of control will contribute to a broader understanding of this dimension of musical imagery.

Conclusion

Musical imagery is a nearly universal, salient experience that serves as a good model for the understanding of auditory imagery more broadly. It is a good context to address broader questions about control over imagery because musical imagery is common, easy to explain to participants, and interesting to a large community. The present research provides new insights into how different measures of control and different imagery contents related to people's abilities to control musical imagery. These findings suggest that people can initiate and manage their musical imagery and report the ability to initiate and manage their imagery, although people's perceptions are related to, but not closely aligned with, their lab-based task performance. This project also provides several avenues for future research to better understand this underlying process in musical imagery experiences. Overall this work suggests that, contrary to the earworm stereotype, that we are not completely at the mercy of our musical imagery.

REFERENCES

Aleman, A., Nieuwenstein, M. R., Böcker, K. B. E., & Haan, E. H. F. (2000). Music training and mental imagery ability. *Neuropsychologia, 38*, 1664-1668.

Bailes, F. (2006). The use of experience-sampling methods to monitor musical imagery in everyday life. *Musicae Scientiae, 10*(2), 173-190.

Bailes, F. (2007). The prevalence and nature of imagined music in the everyday lives of music students. *Psychology of Music, 35*(4), 555-570.

Bailes, F. (2009). Translating the musical image: Case studies of expert musicians. In A. Chan & A. Noble (Eds.), *Sounds in translation: Intersections of music, technology and society* (pp. 41-59). Canberra, Australia: ANU E Press.

Bailes, F. (2015). Music in mind? An experience sampling study of what and when, towards an understanding of why. *Psychomusicology: Music, Mind, and Brain, 25*(1), 58-68.

Bailes, F., & Bigand, E. (2004, August). A tracking study of mental imagery for popular classical music. In S. D. Lipscomb, R. Ashley, R. O. Gjerdingen, & P. Webster (Eds.), *Proceedings of the 8th International Conference on Music Perception & Cognition.* Paper presented at the 8th International Conference on Music Perception & Cognition, Evanston, IL, 3-7 August (pp. 462-465).

Bailes, F., & Bishop, L. (2012). Musical imagery in the creative process. In D. Collins (Ed.), *The act of musical composition: Studies in the creative process* (pp. 53-78). New York, NY: Routledge.

Bailes, F., Bishop, L., Stevens, C. J., & Dean, R. T. (2012). Mental imagery for musical changes in loudness. *Frontiers in Psychology, 3*, 1-9.

Beaman, C. P., & Williams, T. I. (2010). Earworms ('stuck song syndrome'): Towards a natural history of intrusive thoughts. *British Journal of Psychology, 101*, 637-653.

Beaman, C. P., & Williams, T. I. (2013). Individual differences in mental control predict involuntary musical imagery. *Musicae Scientiae, 17*(4), 398-409.

Beaty, R. E., Burgin, C. J., Nusbaum, E. C., Kwapil, T. R., Hodges, D. A., & Silvia, P. J. (2013). Music to the inner ears: Exploring individual differences in musical imagery. *Consciousness and Cognition, 22,* 1163-1173.

Bernardi, N. F., De Buglio, M., Trimarchi, P. D., Chielli, A., & Bricolo, E. (2013). Mental practice promotes motor anticipation: Evidence from skilled music performance. *Frontiers in Human Neuroscience, 7*, 1-14.

Bishop, L., Bailes, F., & Dean, R. T. (2013). Musical expertise and the ability to imagine loudness. *PLoS One, 8*(2), 1-12.

Block, R. A., & Gruber, R. P. (2014). Time perception, attention, and memory: A selective review. *Acta Psychologica, 149*, 129-133.

Block, R. A., & Zakay, D. (1997). Prospective and retrospective duration judgments. A meta-analytic review. *Psychonomic Bulletin & Review, 4*(2), 184-197.

Bolger, N., & Laurenceau, J. P. (2013). *Intensive longitudinal methods: An introduction to diary and experience sampling research.* New York: Guilford.

Bowes, P. L. (2009). *An exploratory study of the use of imagery by vocal professionals: Application of a sport psychology framework* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 3420534).

Brown, S. (2006). The perceptual music track: The phenomenon of constant musical imagery. *Journal of Consciousness Studies, 13*(6), 43-62.

Campbell, D. T. & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*(2), 81-105.

Campbell, S. M., & Margulis, E. H. (2015). Catching an earworm through movement. *Journal of New Music Research, 44*(4), 347-358.

Christensen, A. P. (2018). NetworkToolbox: Methods and measures for brain, cognitive, and psychometric network analysis in R. *The R Journal, 10*, 422-439.

Cotter, K. N. (2017). *Understanding inner music: A dimensional approach to musical imagery* (Master's Thesis). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 10259933).

Cotter, K. N. (2019). Mental control in musical imagery: A dual component model. *Frontiers in Psychology, 10*(1904), 1-13.

Cotter, K. N., Christensen, A. P., & Silvia, P. J. (2016). Musical minds: Personality, schizotypy, and involuntary musical imagery. *Psychomusicology: Music, Mind, and Brain, 26*(3), 220-225.

Cotter, K. N., Christensen, A. P., & Silvia, P. J. (2019). Understanding inner music: A dimensional approach to musical imagery. *Psychology of Aesthetics, Creativity, and the Arts, 13*(4), 489-503.

Cotter, K. N. & Silvia, P. J. (2017). Measuring mental music: Comparing retrospective and experience sampling methods for assessing musical imagery. *Psychology of Aesthetics, Creativity, and the Arts, 11*(3), 335-343.

Cotter, K. N., & Silvia, P. J. (in press). Tuning the inner radio: Mental control of musical imagery. *Psychology of Music.*

Crowder, R. G., & Pitt, M. A. (1992). Research of memory/imagery for musical timbre. In D. Reisberg (Ed.) *Auditory imagery* (pp. 29-44). New York: Psychology Press.

Cumming, G. (2012). *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis*. New York: Routledge.

Duncan, T. E., Duncan, S. C., & Strycker, L. A. (2006). *An introduction to latent variable growth curve modeling: Concepts, issues, and applications* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Epskamp, S., & Fried, E. I. (2018). A tutorial on regularized partial correlation networks. *Psychological Methods, 23*(4), 617–634.

Farah, M. J., & Smith, A. F. (1983). Perceptual interference and facilitation with auditory imagery. *Perception & Psychophysics, 33*(5), 475-478.

Fine, P. A., Wise, K. J., Goldemberg, R., & Bravo, A. (2015). Performing musicians' understanding of the terms "mental practice" and "score analysis." *Psychomusicology: Music, Mind, and Brain, 25*(1), 69-82.

Floridou, G. A., & Müllensiefen, D. (2015). Environmental and mental conditions predicting the experience of involuntary musical imagery: An experience sampling method study. *Consciousness and Cognition, 33*, 472-486.

Floridou, G. A., Williamson, V. J., Stewart, L., & Müllensiefen, D. (2015). The involuntary musical imagery scale (IMIS). *Psychomusicology: Music, Mind, and Brain, 25*(1), 28-36.

Foster, N. E. V. & Zatorre, R. J. (2010). Cortical structure predicts success in performing musical transformation judgments. *NeuroImage, 53*, 26-36.

Foster, N. E. V., Halpern, A. R., & Zatorre, R. J. (2013). Common parietal activation in musical mental transformations across pitch and time. *NeuroImage, 75,* 27-35.

Gelding, R. W., Thompson, W. F., & Johnson, B. W. (2015). The pitch imagery arrow task: Effects of musical training, vividness, and mental control. *PLoS One, 10*(3), 1-20.

Godøy, R. I., Haga, E., & Jensenius, A. R. (2006). Playing "air instruments": Mimicry of sound-producing gestures by novices and experts. In S. Gibet, N. Courty, & J.-F. Kamp (Eds.) *Gesture in Human-Computer Interaction and Simulation* (pp. 256-267). Berlin: Springer-Verlag.

Golino, H. F., & Christensen, A. P. (2019). EGAnet: Exploratory Graph Analysis – A framework for estimating the number of dimensions in multivariate data using network psychometrics. R package version 0.8.0.

Golino, H. F., & Demetriou, A. (2017). Estimating the dimensionality of intelligence like data using Exploratory Graph Analysis. *Intelligence, 62*, 54-70.

Golino, H. F., & Epskamp, S. (2017). Exploratory graph analysis: A new approach for estimating the number of dimensions in psychological research. *PloS one*, *12*(6), e0174035.

Gregg, M. J., Clark. T. W., & Hall, C. R. (2008). Seeing the sound: An exploration of the use of mental imagery by classical musicians. *Musicae Scientiae, 12*(2), 231-247.

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

Halpern, A. R. (1992). Musical aspects of auditory imagery. In D. Reisberg (Ed.) *Auditory imagery* (pp. 1-28). New York: Psychology Press.

Halpern, A. R. (2015). Differences in auditory imagery self-report predict neural and behavioral outcomes. *Psychomusicology: Music, Mind, and Brain*, *25*(1), 37-47.

Halpern, A. R., & Zatorre, R. J. (1999). When that tune runs through your head: A PET investigation of auditory imagery for familiar melodies. *Cerebral Cortex*, *9*(7), 697-704.

Herholz, S. C., Lappe, C., Knief, A., & Pantev, C. (2008). Neural basis of music imagery and the effect of musical expertise. *European Journal of Neuroscience, 28,* 2352-2360.

Holmes, P. (2005). Imagination in practice: A study of the integrated roles of interpretation, imagery and technique in the learning and memorisation processes of two experienced solo performers. *British Journal of Music Education, 22*(3), 217-235.

Hubbard, T. L. (2013). Auditory aspects of auditory imagery. In S. Lacey & R. Lawson (Eds.), *Multisensory imagery* (pp. 51-76). New York: Springer.

Hubbard, T. L. (2018). Some methodological and conceptual considerations in studies of auditory imagery. *Auditory Perception & Cognition, 1*, 6-41.

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.

Hurlburt, R. T., & Heavey, C. L. (2001). Telling what we know: Describing inner experience. *Trends in Cognitive Sciences, 5*(9), 400-403.

Jakubowski, K., Bashir, Z., Farrugia, N., & Stewart, L. (2018). Involuntary and voluntary recall of musical memories: A comparison of temporal accuracy and emotional responses. *Memory & Cognition, 46,* 741-756.

Jakubowski, K., Farrugia, N., & Stewart, L. (2016). Probing imagined tempo for music: Effects of motor engagement and musical experience. *Psychology of Music, 44*(6), 1274-1288.

Jakubowski, K., Halpern, A. R., Grierson, M., & Stewart, L. (2015). The effect of exercise-induced arousal on chosen tempi for familiar melodies. *Psychonomic Bulletin and Review, 22*, 559-565.

Janata, P., & Paroo, K. (2006). Acuity of auditory images in pitch and time. *Perception & Psychophysics*, *68*(5), 829-844.

Johnson, R. B. (2011). Musical tempo stability in mental practice: A comparison of motor and non-motor imagery techniques. *Research Studies in Music Education, 33*(1), 3-30.

Johnson, R. E., Rosen, C. C., & Djurdjevic, E. (2011). Assessing the impact of common metho variance on higher order multidimensional constructs. *Journal of Applied Psychology, 96*(4), 744-761.

Kvavilashvili, L., & Anthony, S. H. (2012). *When do Christmas songs pop into your mind?: Testing a long-term priming hypothesis.* Poster presented at the 53rd Annual Meeting of the Psychonomic Society, Minneapolis, Minnesota.

Liikkanen, L. A. (2008). *Music in every mind: Commonality of involuntary musical imagery.* Paper presented at the 10th International Conference on Music Perception and Cognition, Sapporo, Japan.

Liikkanen, L. A. (2011). Musical activities predispose to involuntary musical imagery. *Psychology of Music, 40*(2), 236-256.

Lipson, C. T. (2006). The meanings and functions of tunes that come into one's head. *Psychoanalytic Quarterly, 75,* 859-878.

Little, T. D. (2013). *Longitudinal structural equation modeling.* New York, NY: Guilford.

Maniaci, M. R., & Rogge, R. D. (2014). Caring about carelessness: Participant inattention and its effects on research. *Journal of Research in Personality, 48*, 61-83.

Margulis, E. H. (2014). *On repeat: How music plays the mind.* New York, NY: Oxford University Press.

McKibben, W. B., & Silvia, P. J. (2016). Inattentive and socially desirable responding: Addressing subtle threats to validity in quantitative counseling research. *Counseling Outcome Research and Evaluation, 7,* 53-64.

McKibben, W. B., & Silvia, P. J. (2017). Evaluating the distorting effects of inattentive responding and social desirability on self-report scales in creativity and the arts. *Journal of Creative Behavior, 51,* 57-69.

McNally-Gagnon, A. (2016). *Imagerie Musicale Involontaire: Caractérstiques Phénoménologiques et Mnésiques.* Doctoral dissertation, University of Montreal, Montreal.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84*(3), 231-259.

Pons, P., & Latapy, M. (2006). Computing communities in large networks using random walks. *Journal of Graph Algorithms and Applications 10*(2), 191-218.

Pozzi, F., Di Matteo, T., & Aste, T. (2013). Spread of risk across financial markets: better to invest in the peripheries. *Scientific reports*, *3*, 1665.

Preacher, K. J., Wichman, A. L., MacCallum, R. C., & Briggs, N. E. (2008). *Latent growth curve modeling*. Thousand Oaks, CA: Sage.

R Core Team (2018). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research, 47,* 667-696.

Rubin-Rabson, G. (1941). Studies in the psychology of memorizing piano music. VI: A comparison of two forms of mental rehearsal and keyboard overlearning. *Journal of Educational Psychology*, *32*(8), 593-602.

Saintilan, N. (2014). The use of imagery during the performance of memorized music. *Psychomusicology: Music, Mind, and Brain*, *24*(4), 309-315.

Selya, A. S., Rose, J. S., Dierker, L. C., Hedeker, D., & Mermelstein, R. J. (2012). A practical guide to calculating Cohen's f^2, a measure of local effect size, from PROC MIXED. *Frontiers in Psychology*, *3*(111), 1-6.

van Borkulo, C. D., Boschloo, L., Kossakowski, J., Tio, P., Schoevers, R. A., Borsboom, D., & Waldorp, L. J. (2018). Comparing network structures on three aspects: A permutation test. Manuscript under review. pp. 1–34. Available from: https://www.researchgate.net/profile/Claudia_Van_Borkulo/publication/314750838_Comparing_network_structures_on_three_aspects_A_permutation_test/links/58c55ef145851538eb8af8a9/Comparing-network-structures-on-three-aspects-A-permutation-test.pdf

Vuvan, D. T., & Schmuckler, M. A. (2011). Tonal hierarchy representations in auditory imagery. *Memory and Cognition*, *39*, 477-490.

Weber, R. J., & Brown, S. (1986). Musical imagery. *Music Perception, 3*(4), 411-426.

Weir, G., Williamson, V. J., & Müllensiefen, D. (2015). Increased involuntary musical mental activity is not associated with more accurate voluntary musical imagery. *Psychomusicology: Music, Mind, and Brain*, *25*(1), 48-57.

Williamson, V. J., & Jilka, S. R. (2014). Experiencing earworms: An interview study of involuntary musical imagery. *Psychology of Music, 42*(5), 653-670.

Williamson, V. J., Jilka, S. R., Fry, J., Finkel, S., Müllensiefen, D., & Stewart, L. (2011). How do “earworms” start? Classifying the everyday circumstances of involuntary musical imagery. *Psychology of Music, 40*(3), 259-284.

Williamson, V. J., Liikkanen, L. A., Jakubowski, K., & Stewart, L. (2014). Sticky tunes: How do people react to involuntary musical imagery? *Plos One, 9*(1), 1-9.

Williams, T. I. (2015). The classification of involuntary musical imagery: The case for earworms. *Psychomusicology: Music, Mind, and Brain, 25*(1), 5-13.

Wöllner, C., & Williamon, A. (2007). An exploratory study of the role of performance feedback and musical imagery in piano playing. *Research Studies in Music Education, 29,* 39-54.

Zatorre, R. J., & Halpern, A. R. (1993). Effect of unilateral temporal-lobe excision on perception and imagery of songs. *Neuropsychologia, 31*(3), 221-232.

APPENDIX A

TABLES

Table 1

Descriptions of Auditory Imagery Tasks

Methodology	Description	Studies
Pitch Discrimination	Participants are presented with auditory stimuli (e.g., tones, song excerpts) and imagine music related to the initial stimuli, such as replicating it or imagining the continuation of the excerpt. People's images are then probed for pitch accuracy by determining whether a target tone or musical notation matches their imagery.	Bailes et al. (2012); Herholz et al. (2008); Janata & Paroo (2006); Weir et al. (2015)
Timing Judgment	Participants listen to the beginning of a song excerpt and imagine the continuation of the excerpt. People's images are then probed for timing accuracy—participants are presented with music from the same excerpt and determine whether it is in time with their image or is appearing too early or late.	Bailes & Bigand (2004); Janata & Paroo (2006); Weir et al. (2015)
Temporal Accuracy	Participants are instructed to imagine music excerpts of varying lengths. For each excerpt, participants indicate when they have imagined the full excerpt.	Halpern (1988); Halpern & Zatorre (1999)
Lyric Comparison	People are shown lyrics from well-known songs with two of the lyrics capitalized (e.g., happy BIRTH-day to YOU). Participants then determine whether the second capitalized lyric is on a pitch higher or lower than the first capitalized lyric.	Aleman et al. (2001); Zatorre & Halpern (1993)
Loudness Profile	People listen to a musical excerpt that varies in loudness during the passage. Participants then imagine the same excerpt, including its loudness profile, and use a slider to indicate the loudness profile of their image.	Bailes et al. (2012); Bishop et al. (2013a)

Contour Tracking	People hear short melodies. People imagine each melody and indicate whether the pitch of a note was higher, lower, or the same as the prior note.	Weber & Brown (1986)
Tempo Judgment	People listen to or imagine excerpts of well-known and familiar songs. People then indicate the tempo of the music by tapping with their finger to the beat or by using a dial to adjust the speed of a click track so it matches the beat of the music.	Jakubowski et al. (2016); Jakubowski et al. (2015)
Pitch Manipulation	Participants are presented with initial tone(s) and manipulate the pitch of the tones to be higher or lower as specified. People then complete a pitch discrimination task.	Gelding et al. (2015); Hubbard & Stoeckig (1988)
Melody Transformation	Participants hear a melody and are presented with a test melody that has been transformed—in a new key or reversed—or an untransformed control melody. People indicate if the test melody, when transformed, matches the first melody.	Foster et al. (2013)

Table 2

Details for Frequently Played Song Stimuli

Song Title	Artist	Start (s)	Cut Out (s)	Return (s)	End (s)	Clip Length (s)	BPM	Task
ABC	The Jackson 5	14.65	27.86	38.11	42.54	27.89	94	Pitch Discrimination
All Star	Smash Mouth	28.54	39.87	50.59	55.95	27.41	104	Pitch Discrimination
Another One Bites the Dust	*Queen*	*61.79*	*85.73*	*96.17*	*101.56*	*39.77*	*110*	*Timing Judgment*
Baby One More Time	**Britney Spears**	**20.50**	**37.59**	**50.06**	**57.80**	**37.30**	**93**	**Pitch Discrimination**
Beat it	**Michael Jackson**	**14.40**	**31.70**	**41.20**	**48.40**	**34.00**	**139**	**Pitch Discrimination**
Billie Jean	Michael Jackson	78.18	92.76	104.14	111.35	33.17	117	Pitch Discrimination
Call Me Maybe	Carly Rae Jepsen	26.11	44.50	53.31	60.51	34.40	120	Pitch Discrimination
Don't Stop Believin'	Journey	77.01	89.07	101.77	105.74	28.73	119	Pitch Discrimination
Eye of the Tiger	Survivor	66.78	83.47	92.39	96.78	30.00	109	Pitch Discrimination
Hallelujah	*Jeff Buckley*	*73.31*	*86.75*	*97.63*	*105.15*	*31.84*	*101*	*Pitch Discrimination*

Hey Ya	**Outkast**	**18.01**	**26.20**	**34.40**	**38.88**	**20.87**	**80**	**Timing Judgment**
Hips Don't Lie	Shakira feat. Wyclef Jean	8.39	24.43	33.98	38.90	30.51	100	Timing Judgment
I Gotta Feeling	**Black Eyed Peas**	**26.61**	**33.40**	**44.80**	**51.83**	**25.22**	**128**	**Pitch Discrimination**
I Will Always Love You	Whitney Houston	188.25	205.09	212.69	220.37	32.12	67	Timing Judgment
I Will Survive	*Gloria Gaynor*	*47.71*	*58.93*	*67.65*	*72.88*	*25.17*	*117*	*Pitch Discrimination*
Jingle Bell Rock	Bobby Helms	7.78	27.17	35.77	38.60	30.82	120	Pitch Discrimination
Just Dance	Lady Gaga	32.75	47.83	58.63	66.61	33.86	119	Timing Judgment
Last Christmas	*Wham!*	*17.54*	*36.06*	*44.85*	*53.81*	*36.27*	*108*	*Pitch Discrimination*
Let it Snow, Let it Snow, Let it Snow	Dean Martin	20.92	34.23	42.23	46.35	25.43	134	Pitch Discrimination
Moves Like Jagger	Maroon 5 feat. Christina Aguilera	42.22	49.52	59.97	63.81	21.59	128	Pitch Discrimination

Poker Face	Lady Gaga	40.39	48.49	56.65	62.31	21.92	119	Pitch Discrimination
Respect	*Aretha Franklin*	*88.12*	*109.20*	*117.39*	*121.52*	*33.40*	*115*	*Timing Judgment*
Rolling in the Deep	Adele	22.96	32.08	41.66	47.07	24.11	105	Timing Judgment
Silent Night	Bing Crosby	10.03	21.74	32.80	44.39	34.36	76	Timing Judgment
Single Ladies (Put a Ring on it)	Beyoncé	21.63	31.27	41.23	50.88	29.25	193	Pitch Discrimination
Stand By Me	Ben E. King	30.12	51.69	61.41	64.38	34.26	118	Timing Judgment
Stayin' Alive	Bee Gees	23.07	37.39	46.51	51.44	28.37	104	Timing Judgment
Thriller	Michael Jackson	81.60	89.53	97.83	102.37	20.77	118	Timing Judgment
Tik Tok	Ke$ha	31.80	39.82	47.83	51.82	20.02	120	Pitch Discrimination
Toxic	Britney Spears	56.54	62.45	73.59	78.16	21.62	143	Timing Judgment
Umbrella	Rihanna	54.61	67.10	75.41	81.59	26.98	174	Timing Judgment
Uptown Funk	Mark Ronson	16.92	37.51	45.80	50.33	33.41	115	Timing Judgment

	and Bruno Mars							
Wannabe	Spice Girls	31.62	40.59	49.31	58.24	26.62	110	Timing Judgment
What a Wonderful World	***Louis Armstrong***	***22.70***	***34.78***	***46.01***	***54.20***	***31.50***	***77***	***Timing Judgment***
Y.M.C.A.	Village People	51.65	59.99	68.53	75.20	23.55	127	Timing Judgment
Yeah!	Usher feat. Lil Jon and Ludacris	44.90	54.73	64.00	68.60	23.70	105	Timing Judgment

Note. BPM = Beats Per Minute. Songs in bold were originally used in Weir et al. (2015); songs in italics were used as practice trials.

Table 3

Details for Infrequently Played Song Stimuli

Song Title	**Artist**	**Start (s)**	**End (s)**	**Clip Length (s)**	**BPM**	**Task**
A Marshmallow World	Dean Martin	42.94	47.92	4.98	111	Tempo Change
Another Christmas Without You	Bobby Helms	38.96	44.37	5.41	107	Key Change
Anymore	Whitney Houston	86.76	94.75	7.99	114	Tempo Change
Baby Be Mine	Michael Jackson	83.48	89.69	6.21	110	Tempo Change
Beautiful, Dirty, Rich	Lady Gaga	62.83	68.36	5.53	120	Tempo Change
Boots & Boys	Ke$ha	53.38	59.61	6.23	126	Tempo Change
Breakout	Bee Gees	15.20	21.30	6.10	125	Key Change
Coming Soon	*Queen*	*70.24*	*76.93*	*6.69*	*137*	*Tempo Change*
Dead or Alive	Journey	92.27	97.70	5.43	190	Tempo Change
Don’t Know Nothing	Maroon 5	14.81	21.93	7.12	130	Tempo Change
Don’t Let Me Lose This Dream	*Aretha Franklin*	*42.63*	*49.23*	*6.60*	*126*	*Key Change*
Drive	Carly Rae Jepsen	49.89	56.60	6.71	123	Tempo Change

E-Mail My Heart	Britney Spears	88.79	93.39	4.60	142	Tempo Change
Home	Smash Mouth	11.73	16.66	4.93	165	Tempo Change
I Guess I'll Get The Papers and Go Home	*Louis Armstrong*	*78.68*	*84.01*	*5.33*	*80*	*Key Change*
I Like It Rough	Lady Gaga	72.20	77.12	4.92	120	Key Change
If You Were There	*Wham!*	*43.97*	*49.74*	*5.77*	*131*	*Key Change*
I'll Bet You	The Jackson 5	57.48	62.78	5.30	82	Key Change
Is Christmas Only A Tree	Bing Crosby	58.44	64.51	6.07	83	Tempo Change
Love Me Real	*Gloria Gaynor*	*31.63*	*38.90*	*7.27*	*126*	*Tempo Change*
Money Make Her Smile	Bruno Mars	167.00	172.18	5.18	93	Key Change
Naked	Spice Girls	36.43	42.47	6.04	160	Key Change
Now Generation	Black Eyed Peas	89.66	95.02	5.36	145	Key Change
On The Horizon	Ben E. King	30.07	37.28	7.21	104	Key Change
Parchman Farm Blues/Preachin' Blues (Up Jumped the Devil)	*Jeff Buckley*	*33.58*	*40.80*	*7.22*	*135*	*Tempo Change*
Question Existing	Rihanna	15.18	22.0	6.90	178	Tempo Change
Radio	Beyoncé	43.00	49.79	6.79	136	Tempo Change
Shadow	Britney Spears	10.95	15.89	4.94	141	Tempo Change
She's Alive	Outkast	16.94	23.91	6.97	144	Key Change

Silver Girl	Survivor	41.28	46.78	5.50	179	Key Change
Tabloid Junkie	Michael Jackson	72.81	79.25	6.44	111	Key Change
Take Your Hand	Usher	27.84	33.42	5.58	96	Key Change
The Lady in My Life	Michael Jackson	20.08	27.22	7.14	146	Key Change
Timor	Shakira	31.63	36.36	4.73	141	Tempo Change
Tired	Adele	19.61	24.57	4.96	194	Key Change
Ups and Downs	Village People	17.17	22.03	4.86	123	Key Change

Note. BPM = Beats Per Minute. Songs in italics were used as practice trials.

Table 4

Auditory Imagery Task Exclusions

Task	**Participants excluded**	**Did not complete task**	**Number of valid trials**				**Trials excluded for RT < 200 ms**				**Trials excluded for familiarity**			
			Med.	*M*	*SD*	*Range*	*Med.*	*M*	*SD*	*Range*	*Med.*	*M*	*SD*	*Range*
Tone Initiation	2	2	12.00	11.97	.18	11, 12	0.00	.03	.18	0, 1	N/A			
Chord Initiation	5	2	12.00	11.77	.57	8, 12	0.00	.23	.57	0, 1	N/A			
Tonal Pitch Discrim.	5	2	12.00	11.86	.56	8, 12	0.00	.14	.56	0, 1	N/A			
Tonal Timing Judgment	6	3	12.00	11.81	.65	7, 12	0.00	.19	.65	0, 5	N/A			
Simple Melody Comp.	3	2	15.00	14.72	.86	6, 15	0.00	.28	.86	0, 9	N/A			
Trans. Melody Comp.	3	2	15.00	14.63	.86	8, 15	0.00	.37	.86	0, 7	N/A			
Rhythm Judgment	2	2	15.00	14.73	.65	11, 15	0.00	.27	.65	0, 4	N/A			
Song Pitch Discrim.	22	2	14.00	12.92	2.62	3, 15	0.00	.03	.21	0, 2	1.00	2.06	2.61	0, 12
Song Timing Judgment	40	2	13.00	11.93	3.26	0, 15	0.00	.07	.42	0, 5	2.00	3.03	3.23	0, 15
Key Change	6	1	15.00	14.00	1.71	4, 15	0.00	.18	.50	0, 3	0.00	.82	1.61	0, 11
Tempo Change	8	4	15.00	14.03	1.95	1, 15	0.00	.17	.47	0, 3	0.00	.56	1.07	0, 6

Note. The "Participants excluded" column represents the total number of participants excluded for not completing tasks or have fewer than 10 valid trials. For the Song Pitch Discrimination and Song Timing Judgment tasks, exclusions for familiarity were due to ratings of 1 or 2; for the Key Change and Tempo Change tasks, exclusions for familiar were due to ratings of 3 or 4. For the Song Pitch Discrimination, Song Timing Judgment, Key Change, and Tempo Change tasks it is possible for trials to be excluded just due to RT, just due to familiarity, or due to both RT and familiarity. Descriptive statistics for number of valid trials, trials excluded for RT, and trials excluded for familiarity were calculated without participants who did not complete the task but do include people who were excluded from the task due to having less than 10 valid trials.

Table 5

Descriptive Statistics for Imagery Tasks and *t*-tests Testing Task Performance against Chance Levels

Task	**N**	**Chance**	**Median**	**Mean [95% CI]**	**SD**	**Range**	***t*-test, Cohen's *d* [95% CI]**
Tone Initiation	183	.33	.50	.54 [.51, .57]	.22	.08 – 1.00	*t* = 12.85***, *d* = .95 [.77, 1.12]
Chord Initiation	180	.33	.55	.55 [.52, .59]	.21	.09 – 1.00	*t* = 14.00***, *d* = 1.04 [.86, 1.22]
Tonal Pitch Discrimination	180	.33	.42	.45 [.42, .48]	.20	.08 – 1.00	*t* = 7.85***, *d* = .59 [.43, .74]
Tonal Timing Judgment	179	.33	.36	.37 [.35, .39]	.15	.08 – .70	*t* = 3.10**, *d* = .23 [.08, .38]
Simple Melody Comparison	182	.50	.67	.68 [.65, .70]	.17	.27 – 1.00	*t* = 14.28***, *d* = 1.06 [.88, 1.24]
Transposed Melody Comparison	182	.50	.53	.55 [.53, .57]	.14	.14 – .87	*t* = 4.59***, *d* = .34 [.19, .49]
Rhythm Judgment	183	.50	.53	.50 [.48, .52]	.13	.13 – .80	*t* = 0.01, *d* = .00 [-.14, .15]
Song Pitch Discrimination	162	.33	.65	.62 [.59, .66]	.20	.13 – 1.00	*t* = 8.27***, *d* = 1.44 [1.21, 1.65]
Song Timing Judgment	144	.33	.39	.39 [.37, .41]	.13	.31 – .77	*t* = 5.11***, *d* = .43 [.25, .60]
Key Change	179	.33	.61	.62 [.59, .65]	.21	.13 – 1.00	*t* = 17.83***, *d* = 1.33 [1.13, 1.53]
Tempo Change	177	.33	.53	.52 [.50, .54]	.16	.08 – .93	*t* = 15.97***, *d* = 1.20 [1.00, 1.39]

Note. Descriptive statistic values represent the proportion of trials with correct responses. One-sample *t*-test were used to assess whether mean task performance different from chance levels. ***p* < .01, ****p* < .001. Cohen's *d* is reported for effect size with values *d* > .20, *d* > .50, and *d* > .80 indicating small, medium, and large effect sizes, respectively.

Table 6

Correlations among Imagery Tasks

Task	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.
1. Tone	1										
2. Chord	.55***	1									
3. Tonal Pitch Discrimination	.46***	.42***	1								
4. Tonal Timing Judgment	.22**	.24**	.22**	1							
5. Simple Melody Comparison	.38***	.40***	.31***	.11	1						
6. Trans. Melody Comparison	.06	.06	.08	.08	.09	1					
7. Rhythm Comparison	.15*	.13	.10	.04	.04	.01	1				
8. Song Pitch Discrimination	.35***	.35***	.44***	.11	.31***	.10	.11	1			
9. Song Timing Judgment	.15	.08	.09	-.11	.27**	.02	.04	.09	1		
10. Key Change	.39***	.43***	.38***	.04	.50***	.09	.15*	.49***	.29***	1	
11. Tempo Change	.30***	.40***	.18*	.08	.32***	-.05	.10	.25**	.12	.34***	1

Note. * $p < .05$, ** $p < .01$, *** $p < .001$. These correlations can be interpreted as effect sizes using the following guidelines: small effect, $r > .10$; medium effect, $r > .30$; and large effect, $r > .50$ (Cumming, 2012).

Table 7

Descriptive Statistics and Correlations for Daily Life Musical Imagery Items

	M (range)	*SD*	1	2	3	4	5	6	7	8	9	10	11	12	13	14
M (range)			.45 (0, 1)	2.80 (1, 7)	.61 (0, 1)	2.07 (1, 7)	.79 (0, 1)	2.14 (1, 7)	.58 (0, 1)	2.56 (1, 7)	.67 (0, 1)	2.53 (1, 7)	.66 (0, 1)	2.44 (1, 7)	.79 (0, 1)	2.24 (1, 7)
SD			.50	2.14	.49	1.46	.41	1.42	.49	1.62	.47	1.60	.47	1.61	.41	1.45
1. Frequency	.42 (0, .96)	.24	1	---	---	---	-.01	-.04	-.03	.12	-.01	.10	-.01	-.03	-.02	.04
2. Purpose	2.91 (1.00, 7.00)	1.58	-.07	1	---	---	.05	-.02	.03	.02	.01	.03	.01	.08	.04	.07
3. Initiate Ability	.67 (0, 1.00)	.30	.49	-.32	1	---	.08	-.24	-.04	.26	.04	.26	.05	-.48	.18	.04
4. Initiate Difficulty	2.14 (1.00, 5.50)	.96	-.14	.27	-.27	1	-.12	.31	-.11	.16	-.07	.16	-.13	.24	-.08	.28
5. Speed Ability	.73 (0, 1.00)	.28	.42	-.14	.65	-.46	1	---	.24	-.10	.19	-.05	.13	-.05	.15	-.08
6. Speed Difficulty	2.40 (1.00, 7.00)	1.13	-.22	.28	-.32	.83	-.54	1	-.05	.27	-.02	.19	-.14	.39	-.07	.21
7. Key Ability	.53 (0, 1.00)	.35	.44	-.04	.45	-.34	.73	-.42	1	---	.25	-.09	.18	-.01	.17	-.05
8. Key Difficulty	2.88 (1.00, 7.00)	1.33	-.11	.16	-.09	.69	-.45	.82	-.49	1	-.09	.33	.00	.22	-.01	.21
9. Vocal Ability	.58 (0, 1.00)	.34	.43	-.09	.54	-.43	.78	-.51	.74	-.42	1	---	---	---	.13	-.04
10. Vocal Difficulty	2.73 (1.00, 7.00)	1.22	-.17	.17	-.09	.70	-.39	.80	-.35	.86	-.54	1	---	---	-.01	.25
11. Instrument Ability	.56 (0, 1.00)	.40	.44	-.06	.51	-.43	.71	-.46	.70	-.45	.72	-.43	1	---	.07	-.05
12. Instrument Difficulty	2.94 (1.00, 7.00)	1.50	-.24	.21	-.29	.67	-.35	.72	-.29	.75	-.41	.76	-.57	1	-.13	.29
13. Change Ability	.74 (0, 1.00)	.29	.42	-.18	.79	-.41	.75	-.39	.61	-.27	.71	-.25	.59	-.35	1	---

14. Change Difficulty	2.48 (1.00, 7.00)	1.08	-.23	.25	-.35	.89	-.49	.84	-.39	.72	-.50	.78	-.50	.80	-.52	1

Note. Within-person descriptive statistics (row) and correlations are presented above the diagonal; between-person descriptive statistics (column) and correlations are below the diagonal. These correlations can be interpreted as effect sizes using the following guidelines: small effect, $r > .10$; medium effect, $r > .30$; and large effect, $r > .50$ (Cumming, 2012). The *p*-values for within-person correlations based on multilevel data are more variable than for a cross-sectional, between-person design because the clusters have different variances and numbers of observations. Monte Carlo power simulations, however, indicate that within-person *r*s > .10 are significant at the $p < .05$ level. Some correlations are undefined due to survey branching and are designated by "---".

Table 8

Correlations Between ESM Responses and Lab Imagery Control Tasks

	Tone	Chord	Tonal Pitch Discrim.	Tonal Timing Judgment	Simple Melody Comp.	Trans. Melody Comp.	Rhythm Comp.	Song Pitch Discrim.	Song Timing Judgment	Key Change	Tempo Change
Frequency	.24**	.24**	.29***	.09	.32***	.07	.09	.36***	.24**	.38***	.19*
Purpose	-.12	-.06	-.18**	-.12	-.05	.01	.06	-.12	-.06	-.00	-.02
Initiate Ability	.25**	.20**	.20**	.10	.15	-.12	-.01	.17*	.06	.18*	.01
Initiate Difficulty	-.09	-.05	.02	.00	-.13	.00	.02	-.11	-.02	-.09	-.12
Speed Ability	.24***	.25***	.16**	.04	.31***	-.07	.03	.25***	.15	.23***	.12
Speed Difficulty	-.24***	-.25***	-.17*	-.09	-.26***	-.05	.02	-.31***	-.06	-.17*	-.16*
Key Ability	.23***	.31***	.24***	.03	.32***	-.03	.03	.33***	.15	.34***	.28***
Key Difficulty	-.12	-.18**	-.14*	-.07	-.19*	-.13	.08	-.16	-.11	-.12	-.20**
Vocal Ability	.23***	.25***	.16*	.08	.34***	-.11	-.09	.29***	.15	.32***	.24***
Vocal Difficulty	-.08	-.15*	-.13	-.10	-.20**	-.09	.11	-.16	-.04	-.13	-.16*
Instrument Ability	.21*	.37***	.22**	.07	.36***	-.02	.01	.27**	.16	.40***	.36***
Instrument Difficulty	-.23**	-.22*	-.25**	-.22*	-.18	-.06	.15	-.16	-.05	-.17	-.25**
Change Ability	.25***	.23***	.16*	.16	.15*	-.02	.05	.20**	.05	.21**	.03
Change Difficulty	-.27***	-.17*	-.18*	-.11	-.22**	-.12	.05	-.23**	.01	-.13	-.17*

Note. * $p < .05$, ** $p < .01$, *** $p < .001$. These correlations can be interpreted as effect sizes using the following guidelines: small effect, $r > .10$; medium effect, $r > .30$; and large effect, $r > .50$ (Cumming, 2012).

Table 9

Correlations Between ESM Responses and Lab Imagery Control Tasks

	Initiation vs. Management		Tonal vs. Song	
	Intercept	Slope	Intercept	Slope
Frequency	.38***	-.30***	.43***	.29***
Purpose	-.12	.11	-.11	.01
Initiate Ability	.24**	-.25***	.22**	.05
Initiate Difficulty	-.11	.10	-.11	-.13
Speed Ability	.30***	-.28***	.30***	.15*
Speed Difficulty	-.30***	.28***	-.30***	-.12
Key Ability	.38***	-.33***	.40***	.28***
Key Difficulty	-.22***	.19**	-.24***	-.13
Vocal Ability	.34***	-.29***	.35***	.29***
Vocal Difficulty	-.18*	.14*	-.19**	-.12
Instrument Ability	.42***	-.36***	.44***	.32***
Instrument Difficulty	-.32***	.31***	-.30***	-.10
Change Ability	.28***	-.28***	.26***	.07
Change Difficulty	-.28***	.26***	-.28***	-.08

Note. * $p < .05$, ** $p < .01$, *** $p < .001$. These correlations can be interpreted as effect sizes using the following guidelines: small effect, $r > .10$; medium effect, $r > .30$; and large effect, $r > .50$ (Cumming, 2012).

Table 10

Performance on Lab Musical Imagery Tasks Predicting Self-Reported Mental Control Abilities in Daily Life

	Initiation vs. Management		Tonal vs. Song	
	Intercept	Slope	Intercept	Slope
Frequency	1.04*** (.22)	.70** (.23)	.37*** (.08)	.12 (.08)
	[.61, 1.47]	[.25, 1.14]	[.21, .53]	[-.04, .27]
	f^2 = .13	f^2 = .07	f^2 = .13	f^2 = .01
	p < .001	p = .002	p < .001	p = .143
Purpose	-.08 (.26)	.03 (.27)	-.15 (.09)	.09 (.09)
	[-.59, .44]	[-.50, .57]	[-.31, .02]	[-.09, .27]
	f^2 = .00	f^2 = .00	f^2 = .02	f^2 = .01
	p = .762	p = .902	p = .084	p = .323
Initiation Ability	-.10 (.27)	-.36 (.27)	.25** (.09)	-.07 (.09)
	[-.64, .35]	[-.90, .17]	[.06, .43]	[-.24, .10]
	f^2 = .00	f^2 = .01	f^2 = .05	f^2 = .01
	p = .708	p = .184	p = .009	p = .428
Initiation Difficulty	-.12 (.26)	-.02 (.24)	-.07, (.09)	-.07 (.09)
	[-.62, .38]	[-.50, .46]	[-.24, .11]	[-.25, .11]
	f^2 = .00	f^2 = .00	f^2 = .00	f^2 = .00
	p = .644	p = .931	p = .475	p = .429
Tempo Ability	.16 (.12)	-.09 (.13)	.25*** (.06)	-.04 (.14)
	[-.07, .39]	[-.34, .17]	[.14, .36]	[-.32, .23]
	f^2 = .01	f^2 = .02	f^2 = .06	f^2 = .03
	p = .172	p = .498	p < .001	p = .751
Tempo Difficulty	-.34 (.21)	-.05 (.21)	-.30*** (.08)	.03 (.08)
	[-.87, .06]	[-.47, .37]	[-.44, -.16]	[-.13, .19]
	f^2 = .01	f^2 = .00	f^2 = .08	f^2 = .00
	p = .095	p = .814	p < .001	p = .725

Key Ability	.01 (.10) [-.18, .20] f^2 = .00 p = .928	-.17* (.08) [-.33, -.01] f^2 = .00 p = .043	.16 (.09) [-.01, .33] f^2 = .02 p = .061	-.07 (.22) [-.50, .37] f^2 = .00 p = .768
Key Difficulty	-.43 (.25) [-.92, .07] f^2 = .02 p = .090	-.23 (.26) [-.72, .28] f^2 = .01 p = .377	-.22** (.07) [-.36, -.08] f^2 = .04 p = .003	-.01 (.09) [-.18, .16] f^2 = .00 p = .922
Vocal Ability	.13 (.20) [-.26, .53] f^2 = .01 p = .576	-.10 (.18) [-.45, .25] f^2 = .01 p = .562	.23** (.07) [.09, .38] f^2 = .04 p = .002	-.06 (.18) [-.41, .30] f^2 = .00 p = .752
Vocal Difficulty	-.44 (.24) [-.91, .02] f^2 = .02 p = .061	-.30 (.23) [-.76, .15] f^2 = .01 p = .191	-.18* (.08) [-.34, -.01] f^2 = .02 p = .036	.01 (.09) [-.17, .19] f^2 = .00 p = .930
Instrumental Ability	.91*** (.22) [.48, 1.3] f^2 = .15 p < .001	.40 (.24) [-.08, .88] f^2 = .00 p = .103	.54*** (.11) [.32, .77] f^2 =.40 p < .001	.03 (.11) [-.18, .24] f^2 = .00 p = .804
Instrumental Difficulty	-.31 (.27) [-.84, .21] f^2 = .01 p = .245	.00 (.25) [-.49, .50] f^2 = .00 p = .994	-.32*** (.09) [-.49, -.15] f^2 = .09 p < .001	.04 (.11) [-.17, .25] f^2 = .00 p = .711
Change Song Ability	-.22 (16) [-.53, .10] f^2 = .00 p = .185	-.35** (.13) [-.61, -.09] f^2 = .00 p = .009	.10 (.08) [-.06, .27] f^2 = .01 p = .221	-.06 (.09) [-.24, .11] f^2 = .00 p = .479
Change Song Difficulty	-.37 (.23) [-.82, .09]	-.09 (.24) [-.56, .37]	-.30*** (.07) [-.44, -.16]	.06 (.08) [-.10, .22]

	$f^2 = .01$	$f^2 = .00$	$f^2 = .08$	$f^2 = .00$
	$p = .116$	$p = .689$	$p < .001$	$p = .470$

Note. Standardized betas (Standard Error), [95% Confidence Interval]. Coefficients for the Purpose and all difficulty items represent standardization for both the outcomes and predictors; coefficients for the Frequency and ability items represent standardization of only the predictors. * $p < .05$, ** $p < .01$, *** $p < .001$. Cohen's f^2 is reported as an effect size with $f^2 > .02$, $f^2 > .15$, and $f^2 > .35$ indicating small, medium, and large effects, respectively.

Table 11

Communities and Hybrid Centrality for Lab Auditory Imagery Tasks

Task	**Community**	**Hybrid Centrality**
Tone Initiation	1	.89
Chord Initiation	1	.85
Tonal Pitch Discrimination	1	.63
Tonal Timing Judgment	1	.17
Simple Melody Comparison	2	.63
Transposed Melody Comparison	N/A	.00
Rhythm Comparison	N/A	.00
Song Pitch Discrimination	1	.47
Song Timing Judgment	2	.17
Key Change	2	.97
Tempo Change	1	.38

Table 12

Communities and Hybrid Centrality for Between-Person Experience-Sampling Network

Item	**Dimension**	**Hybrid Centrality**
Initiation Ability	1	.16
Initiation Difficulty	2	.55
Speed Ability	1	.92
Speed Difficulty	2	.60
Key Ability	1	.40
Key Difficulty	2	.75
Vocal Ability	1	.66
Vocal Difficulty	2	.55
Instrument Ability	1	.27
Instrument Difficulty	2	.13
Change Song Ability	1	.48
Change Song Difficulty	2	.93

Table 13

Communities and Hybrid Centrality for Lab Auditory Imagery Task and Experience-Sampling Network

Task/Item	Community	Hybrid Centrality
Tone Initiation	2	.75
Chord Initiation	2	.64
Tonal Pitch Discrimination	2	.51
Tonal Timing Judgment	2	.09
Simple Melody Comparison	2	.70
Transposed Melody Comparison	N/A	.00
Rhythm Comparison	2	.02
Song Pitch Discrimination	2	.51
Song Timing Judgment	2	.07
Key Change	2	.82
Tempo Change	2	.38
ESM Initiation Ability	1	.21
ESM Initiation Difficulty	3	.35
ESM Speed Ability	1	.72
ESM Speed Difficulty	3	.57
ESM Key Ability	1	.92
ESM Key Difficulty	3	.40
ESM Vocal Ability	1	.80
ESM Vocal Difficulty	3	.35
ESM Instrument Ability	1	.62
ESM Instrument Difficulty	3	.50
ESM Change Song Ability	1	.52
ESM Change Song Difficulty	3	.85

APPENDIX B

FIGURES

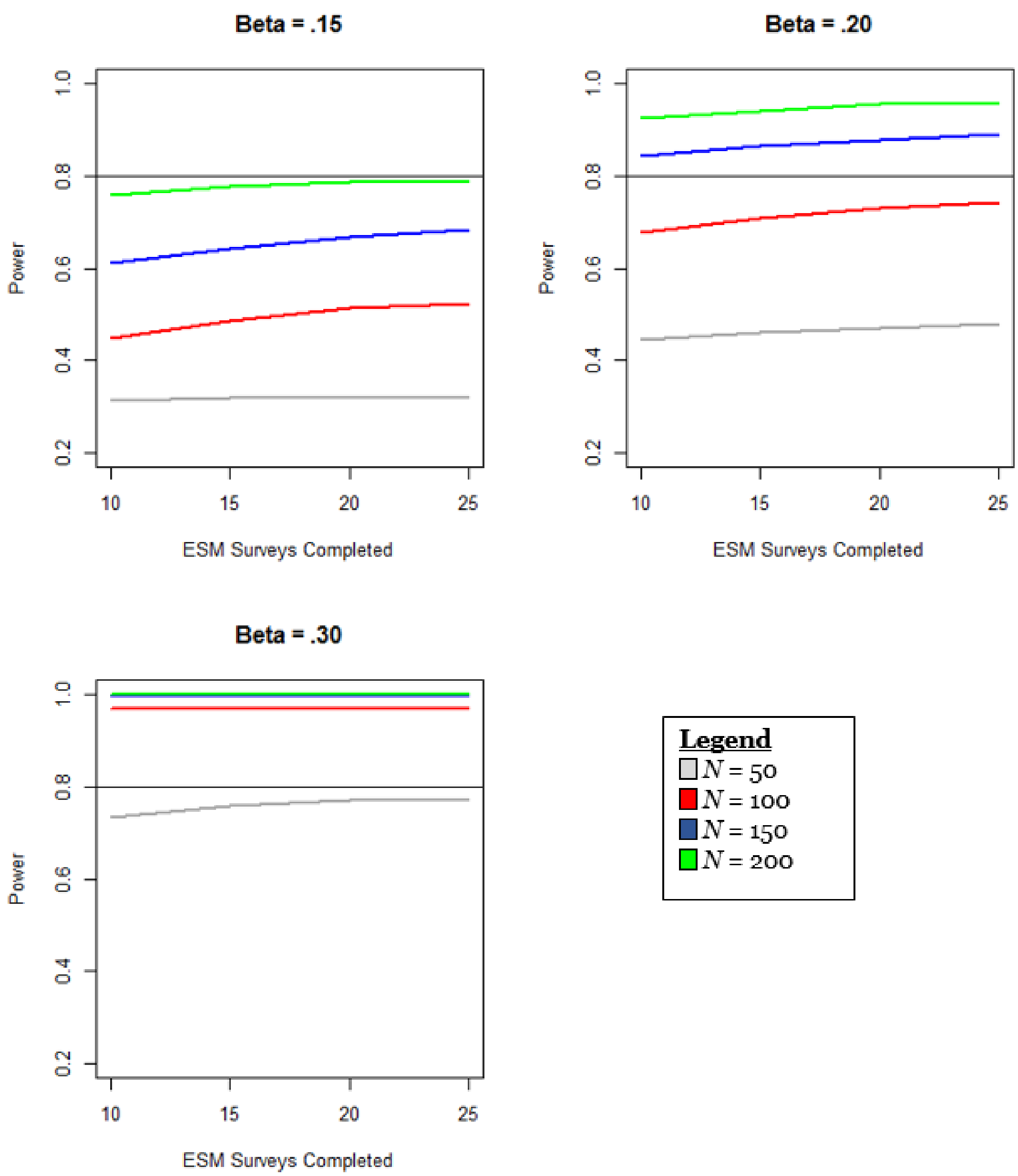
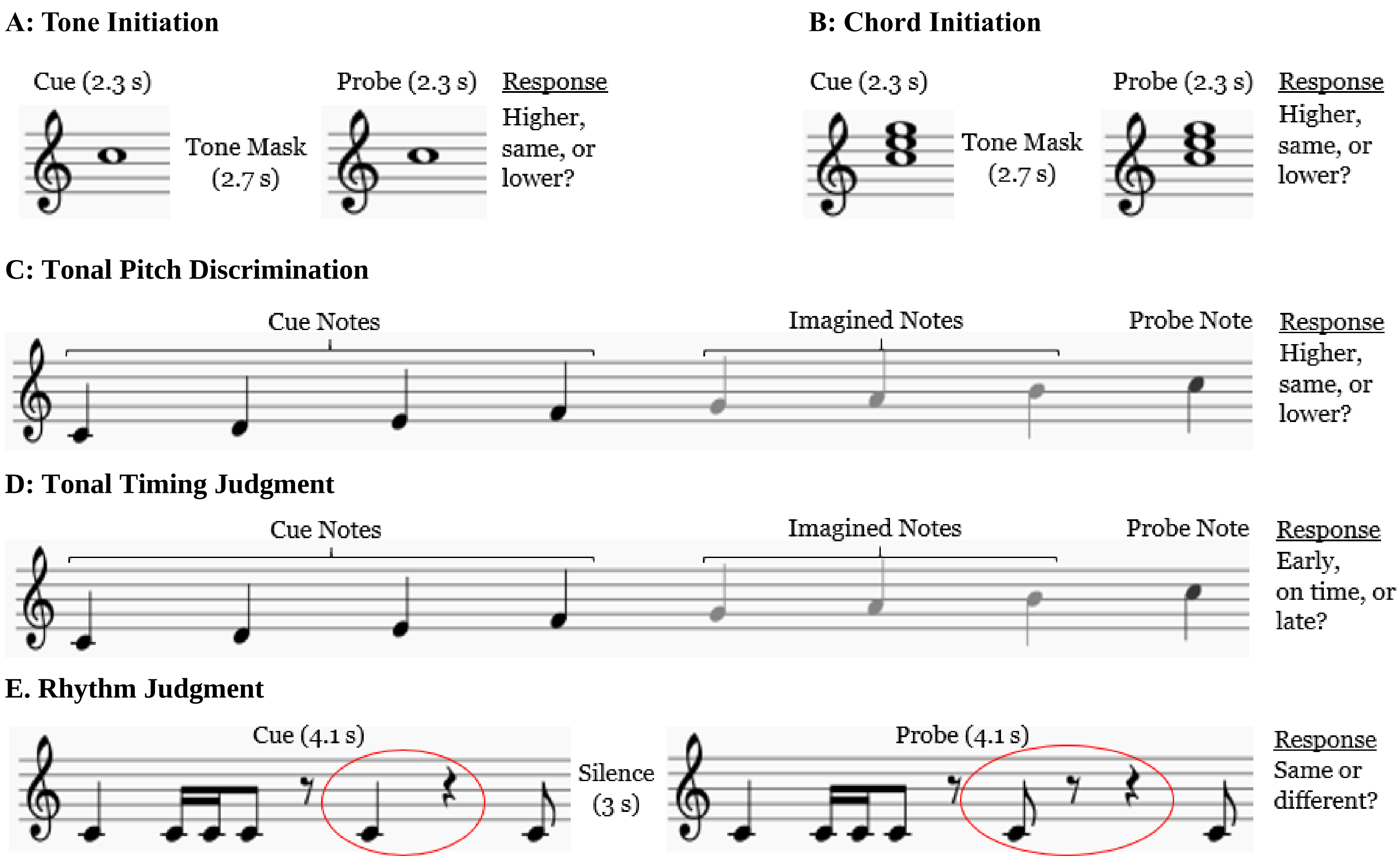
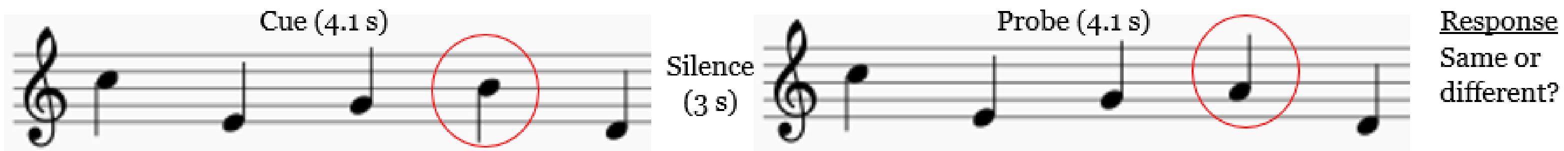


Figure 1. Power Curves based on Monte Carlo Simulations


A: Tone Initiation
Cue (2.3 s)
Tone Mask (2.7 s)
Probe (2.3 s)
Response
Higher, same, or lower?
B: Chord Initiation
Cue (2.3 s)
Tone Mask (2.7 s)
Probe (2.3 s)
Response
Higher, same, or lower?
C: Tonal Pitch Discrimination
Cue Notes
Imagined Notes
Probe Note
Response
Higher, same, or lower?
D: Tonal Timing Judgment
Cue Notes
Imagined Notes
Probe Note
Response
Early, on time, or late?
E. Rhythm Judgment
Cue (4.1 s)
Silence (3 s)
Probe (4.1 s)
Response
Same or different?

F: Simple Melody Comparison



G: Transposed Melody Comparison

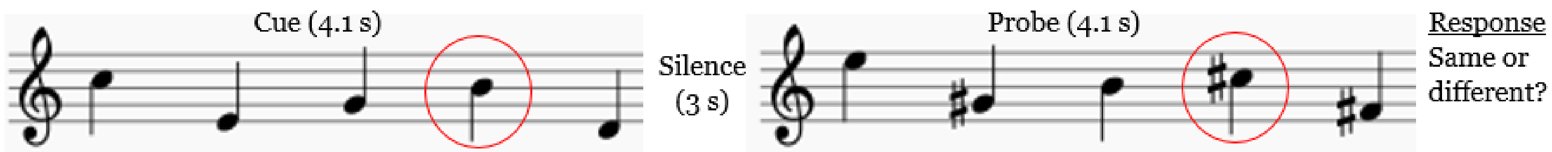


Figure 2. Sample Trials of Tonal Auditory Imagery Tasks

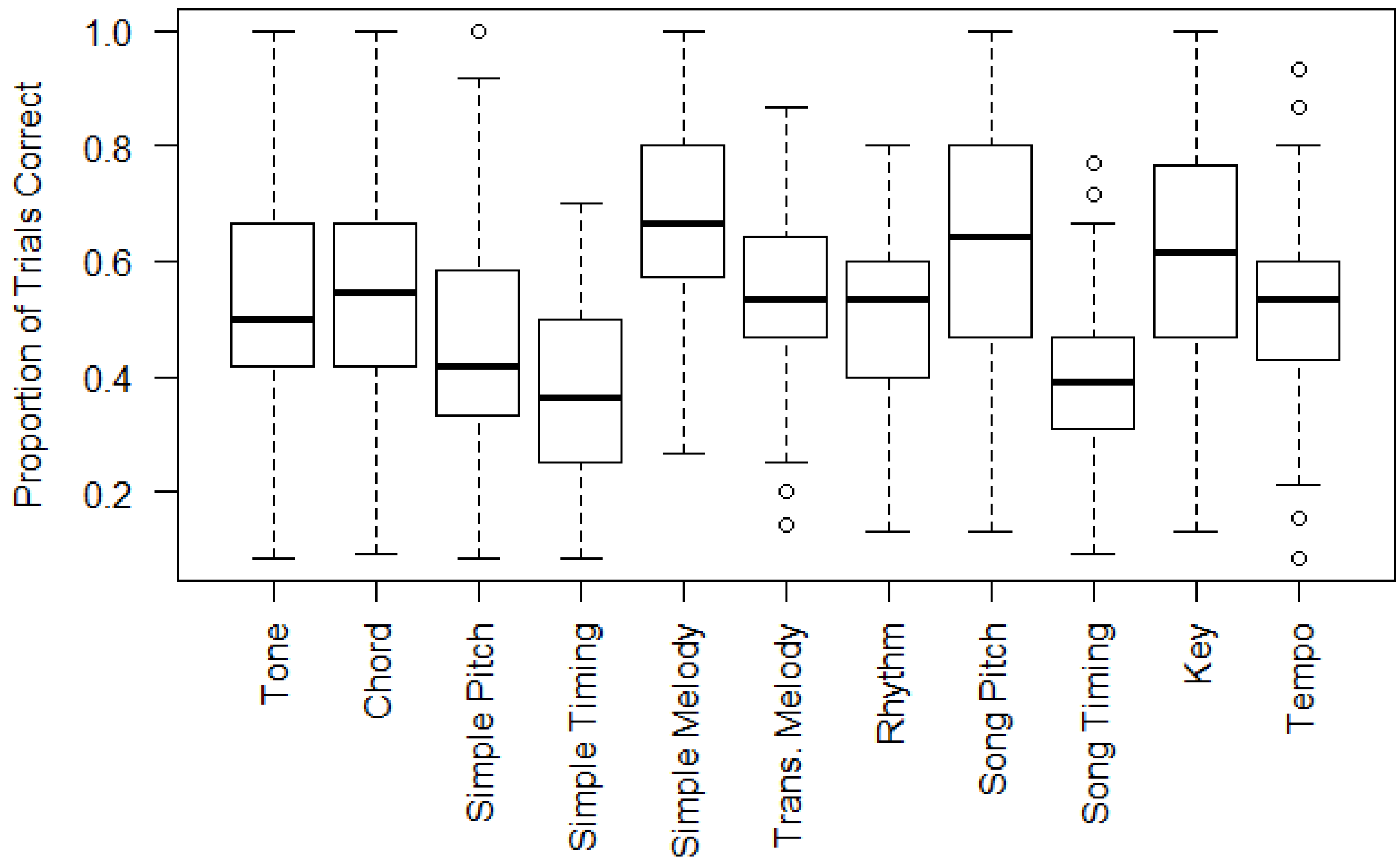


Figure 3. Distributions of Auditory Imagery Task Performance (Proportion of Trials Answered Correctly)

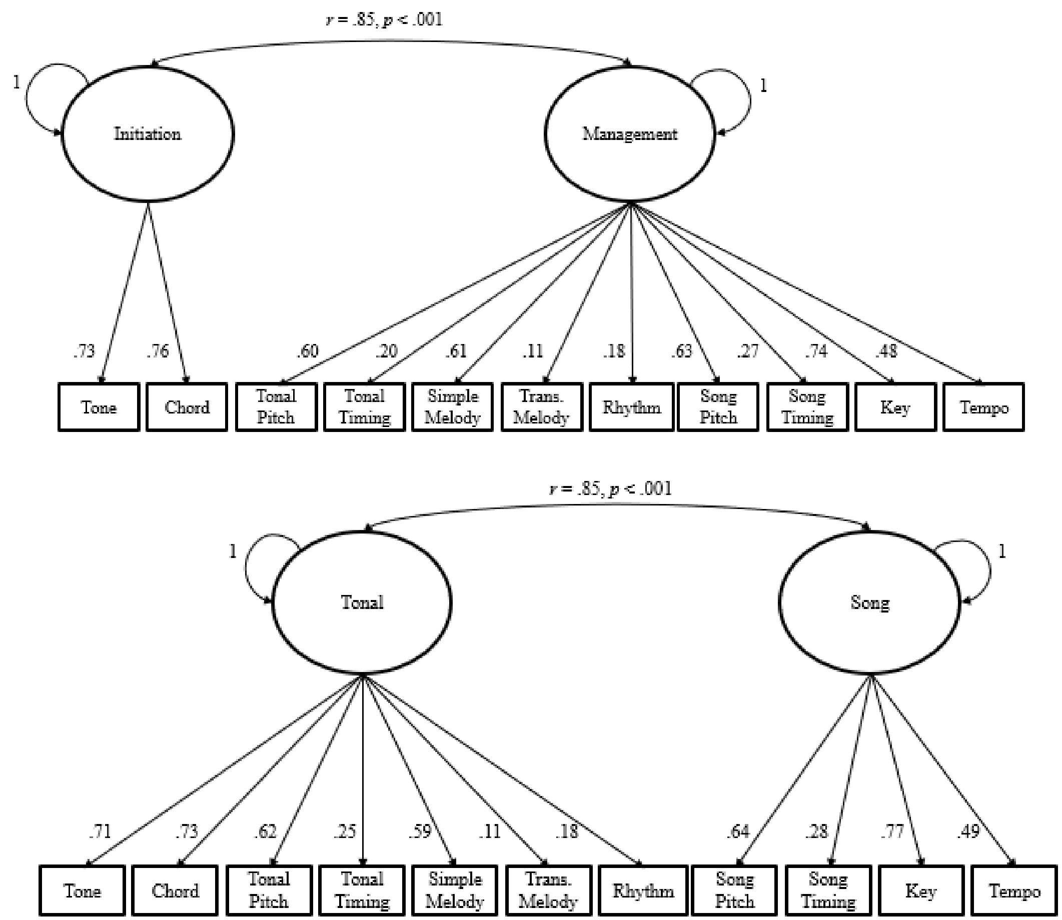


Figure 4. Confirmatory Factor Analyses of Auditory Imagery Tasks

Note. TOP: Initiation vs. management; BOTTOM: Tonal vs. song stimuli. Factor variances were fixed to 1; standardized factor loadings are reported.

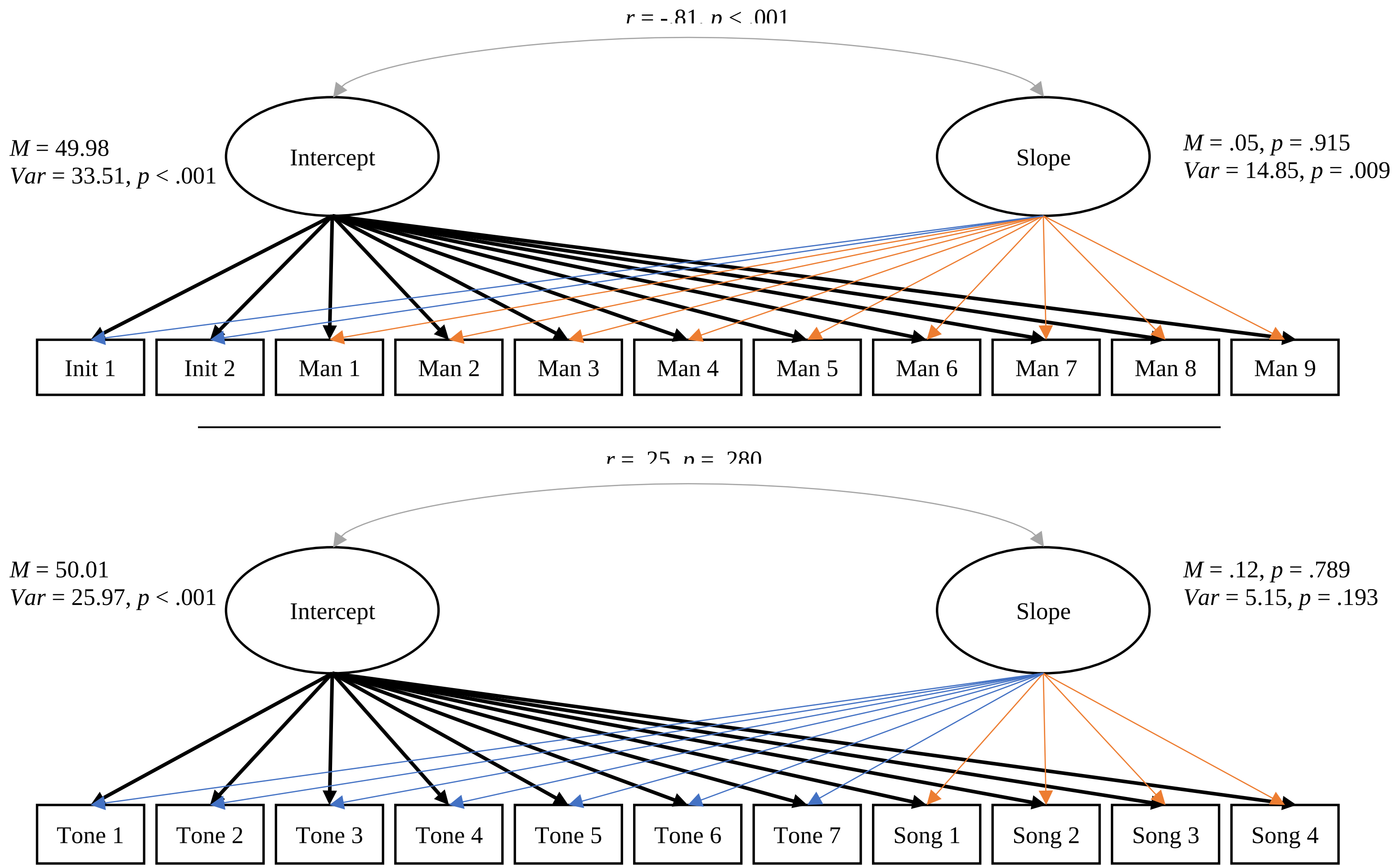


Figure 5. Latent Growth Curve Analyses of Auditory Imagery Tasks

Note. TOP: Modeling initiation vs. management tasks; BOTTOM: Modeling tonal vs. song stimuli. Black paths were fixed to 1, blue paths were fixed to -0.5, and orange paths were fixed to 0.5.

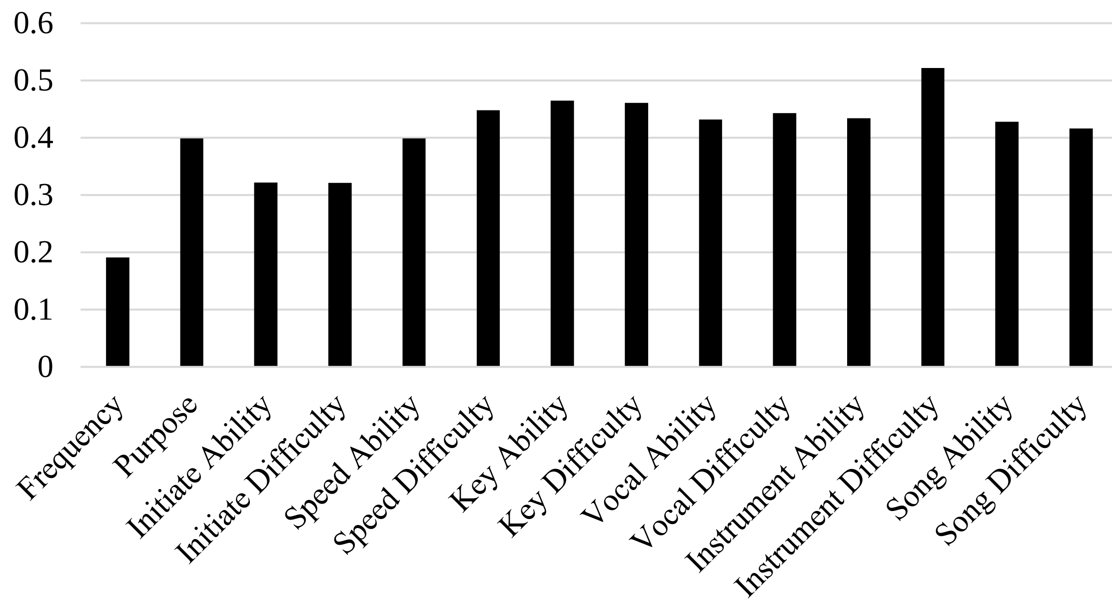


Figure 6. Intraclass Correlation Coefficients for Experience-Sampling Items

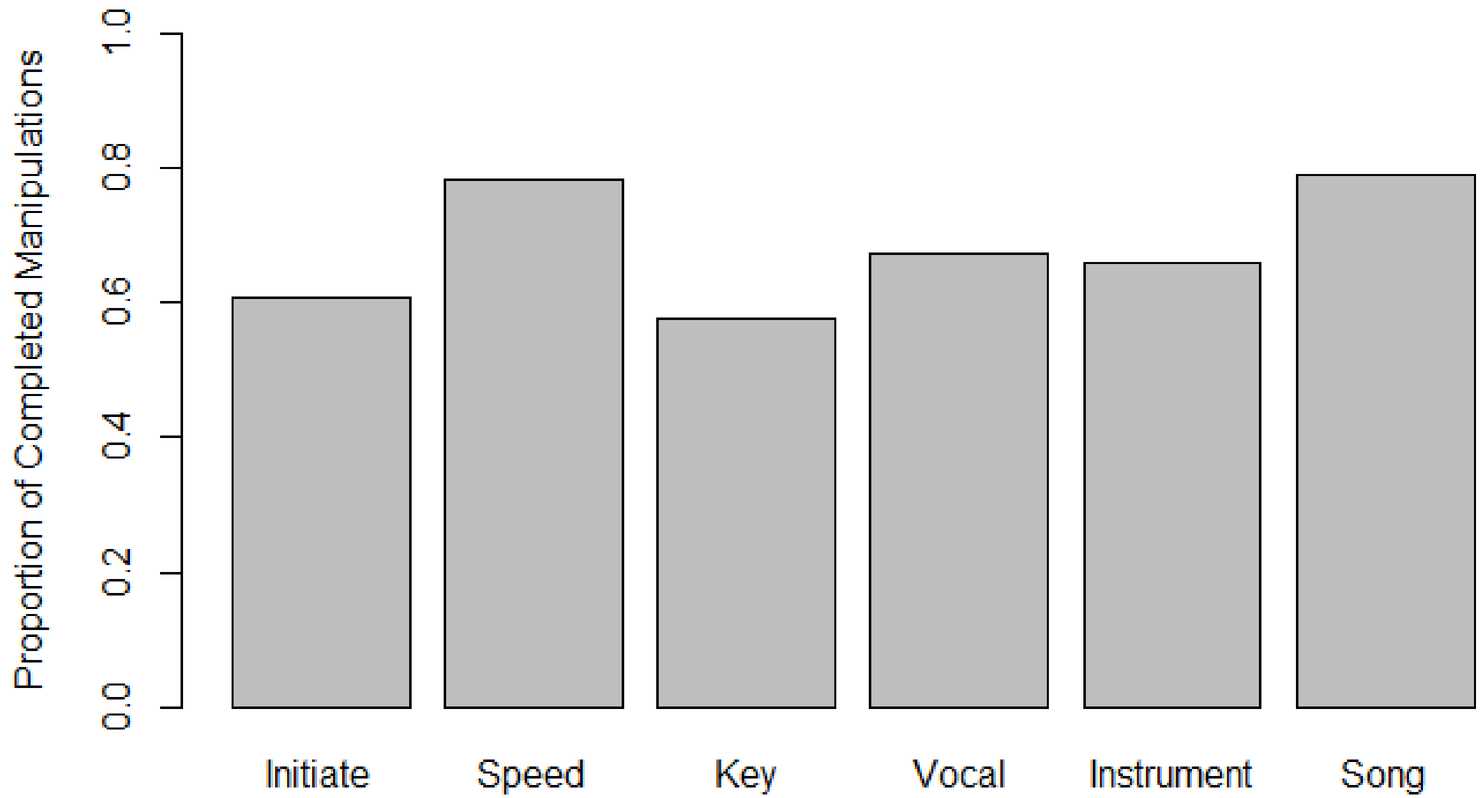


Figure 7. Within-Person Distributions of Experience-Sampling Mental Control Ability Items

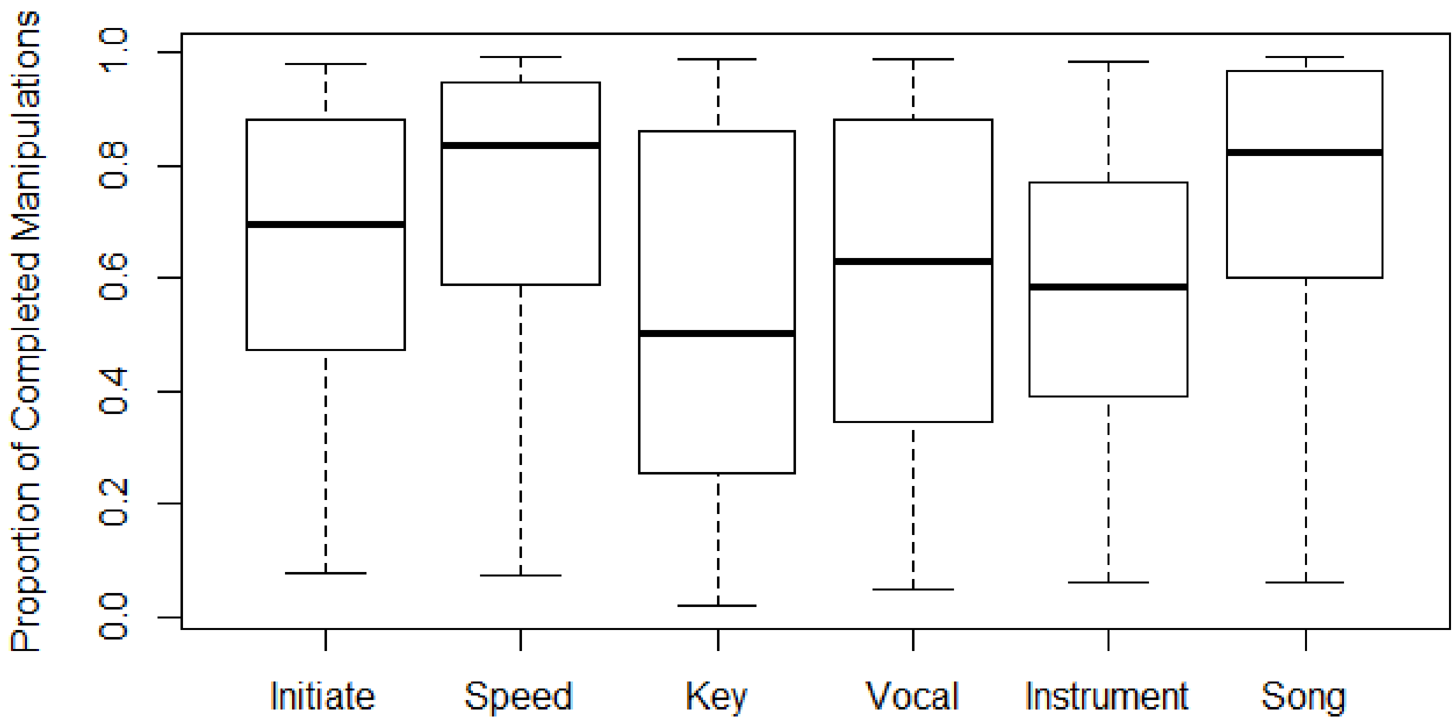


Figure 8. Between-Person Distributions of Experience-Sampling Mental Control Ability Items

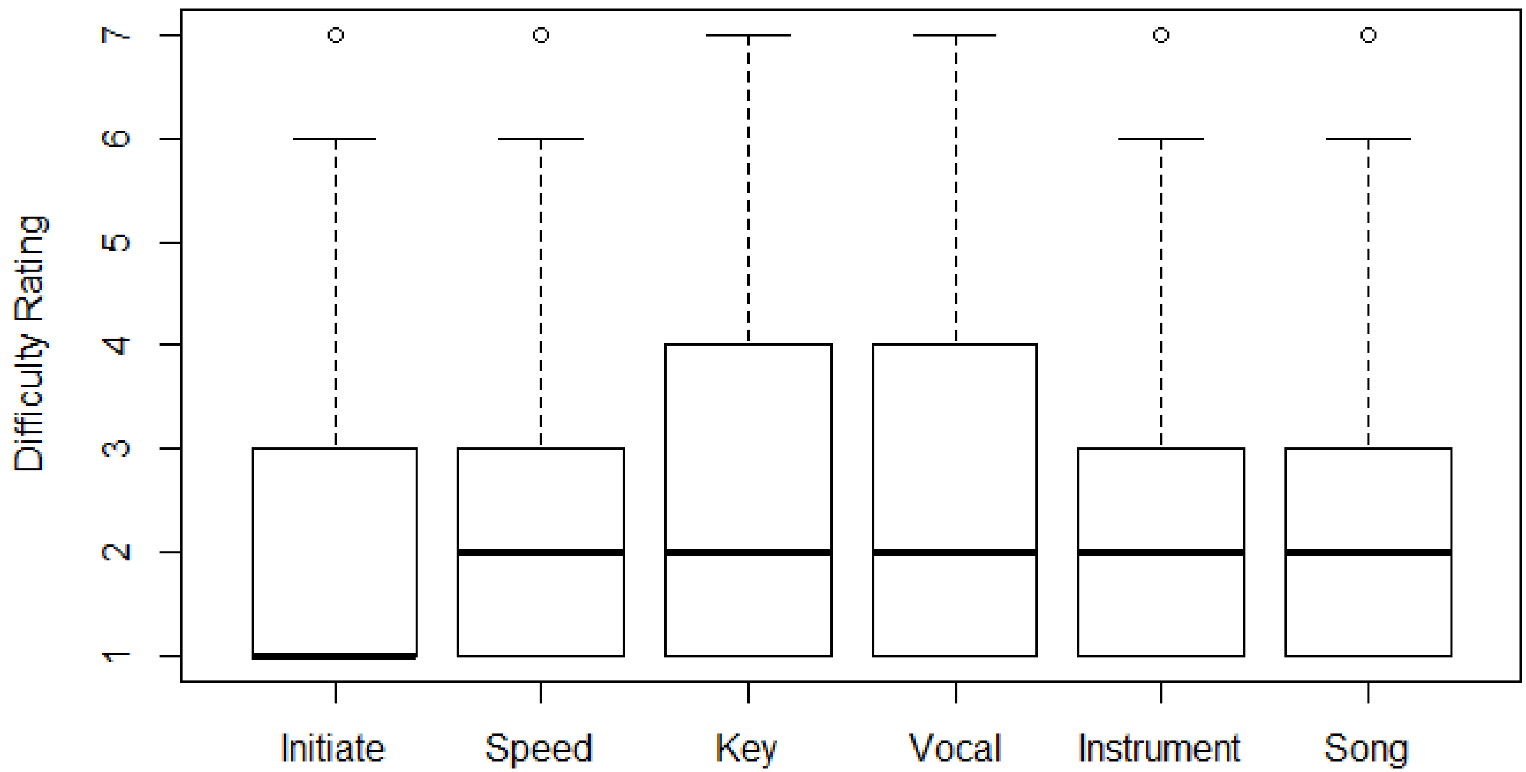


Figure 9. Within-Person Distributions of Experience-Sampling Mental Control Difficulty Items

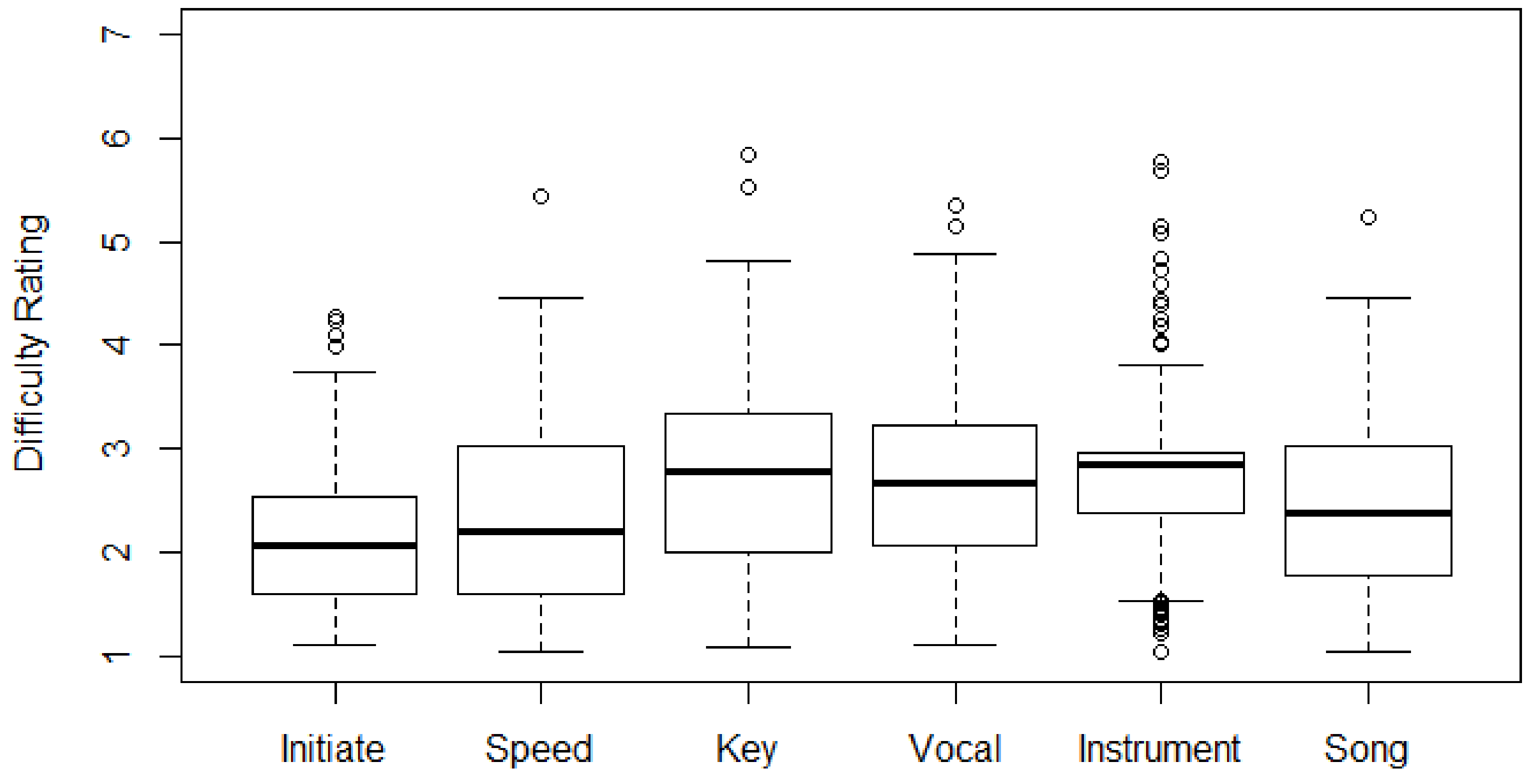


Figure 10. Between-Person Distributions of Experience-Sampling Mental Control Difficulty Items

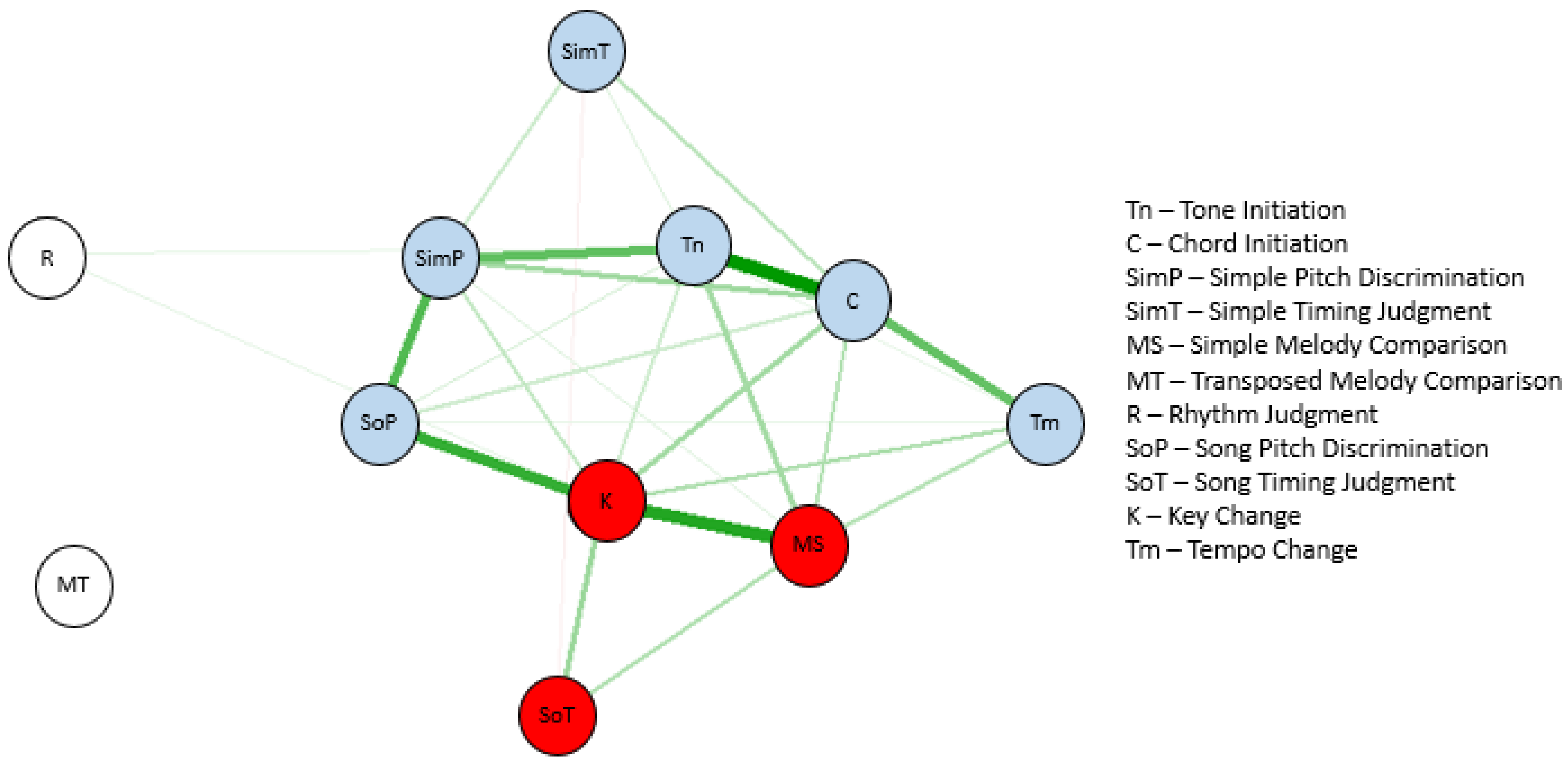


Figure 11. Network of Auditory Imagery Tasks (Clusters are indicated by Node Color)

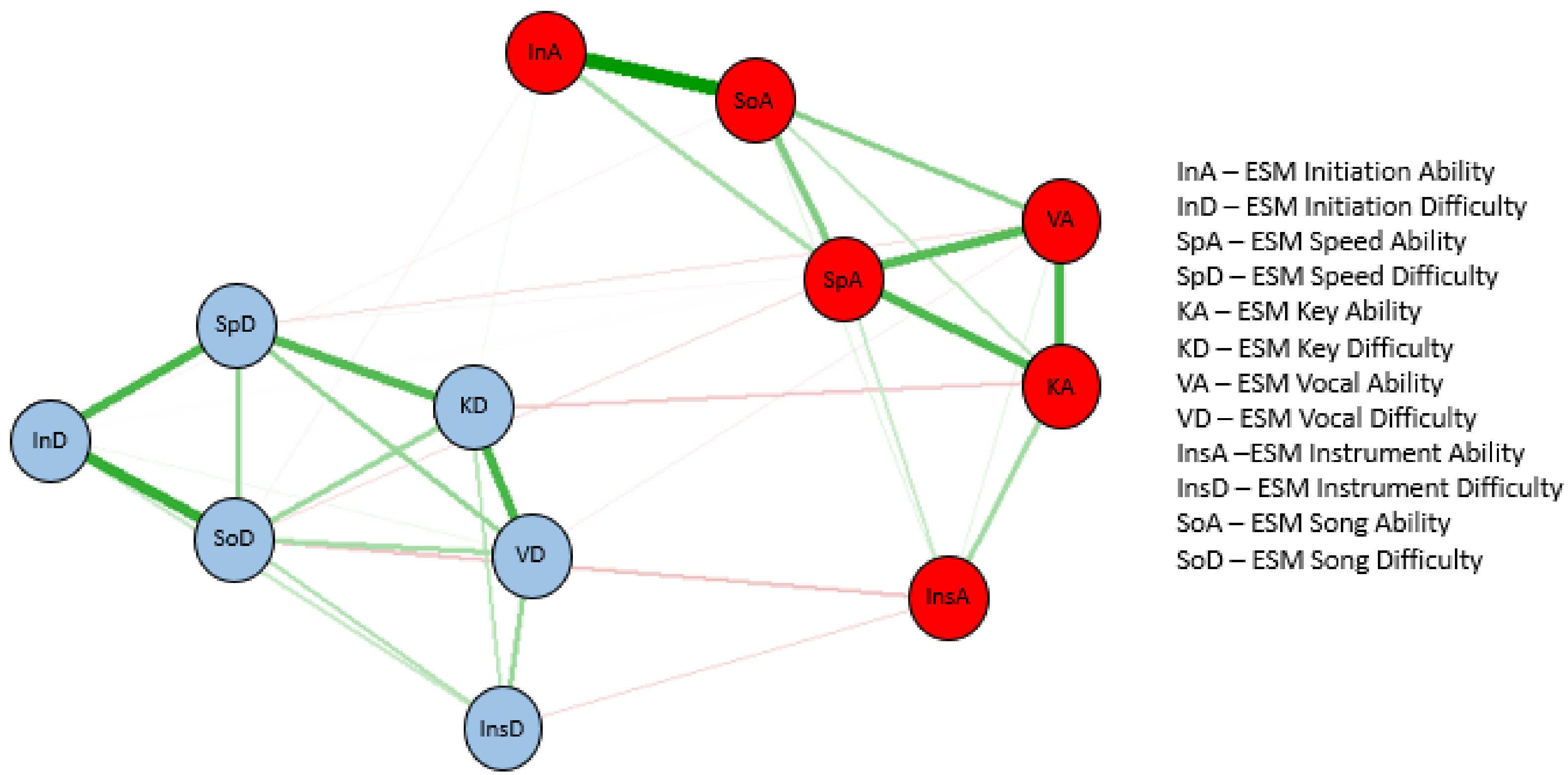


Figure 12. Network of Experience-Sampling Items at the Between-Person Level (Clusters are indicated by Node Color)

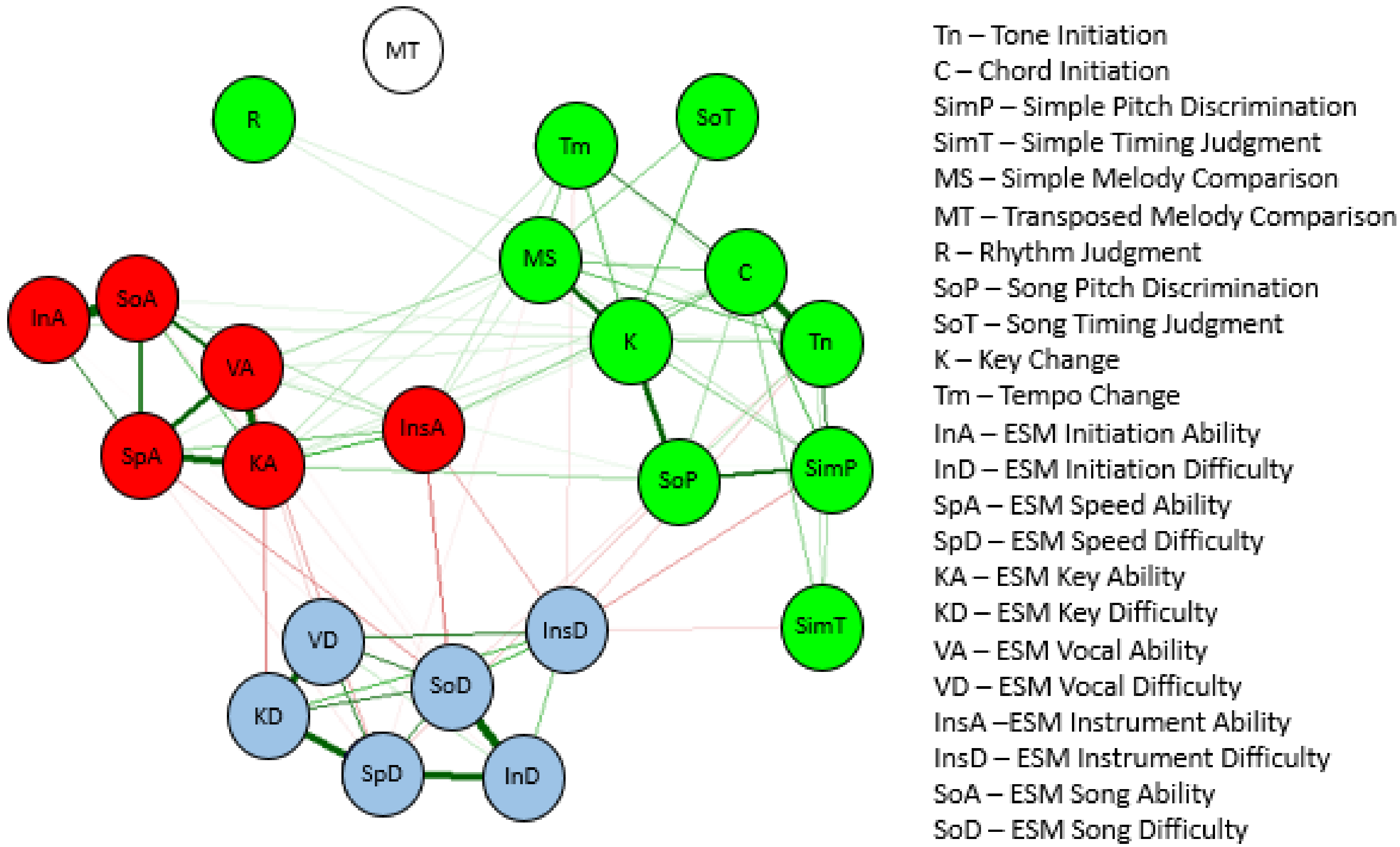


Figure 13. Network of Auditory Imagery Tasks and Experience-Sampling Items at the Between-Person Level (Clusters are indicated by Node Color)

APPENDIX C

MUSICAL IMAGERY EXPERIENCE-SAMPLING SURVEY

Dimension	Item	Response Scale
Frequency	**Hearing Musical Imagery:** Right now, are you hearing music in your head?	*Yes* or *No*
Initiation	**Start on Purpose:** I made the music in my mind start playing on purpose.	1 (*Strongly Disagree*) to 7 (*Strongly Agree*)
	Initiate Ability: Think of a song you have heard recently. Are you able to start playing this song in your head?	*Yes* or *No*
	Initiate Difficulty: How difficult was it to start playing this song?	1 (*Not difficult at all*) to 7 (*Very difficult*)
Management	**Speed Ability:** Try to increase the tempo (speed) of the music in your head. Are you able to do this?	*Yes* or *No*
	Speed Difficulty: How difficult was it to increase the tempo of the music?	1 (*Not difficult at all*) to 7 (*Very difficult*)
	Key Ability: Try to change the key of the music in your head. Are you able to do this?	*Yes* or *No*
	Key Difficulty: How difficult was it to change the key of the music?	1 (*Not difficult at all*) to 7 (*Very difficult*)
	Is the music in your head primarily vocal or instrumental?	*Vocal* or *Instrumental*
	Vocal Ability: Try to change the gender of the vocalist. Are you able to do this?	*Yes* or *No*
	Vocal Difficulty: How difficult was it to change the gender of the voice?	1 (*Not difficult at all*) to 7 (*Very difficult*)
	Instrument Ability: Try to change the primary instrument to a different instrument. Are you able to do this?	*Yes* or *No*
	Instrument Difficulty: How difficult was it to change the instrument?	1 (*Not difficult at all*) to 7 (*Very difficult*)
	Change Song Ability: Try to change the music in your head so that you are hearing a different song. Are you able to do this?	*Yes* or *No*
	Change Song Difficulty: How difficult was it to change to a different song?	1 (*Not difficult at all*) to 7 (*Very difficult*)

Mood and Environment	Right now, I feel happy.	1 (*Strongly disagree*) to 7 (*Strongly agree*)
	Right now, I feel relaxed.	1 (*Strongly disagree*) to 7 (*Strongly agree*)
	Right now, I feel bored.	1 (*Strongly disagree*) to 7 (*Strongly agree*)
	Right now, I feel sad.	1 (*Strongly disagree*) to 7 (*Strongly agree*)
	Right now, I feel irritated.	1 (*Strongly disagree*) to 7 (*Strongly agree*)
	Right now, I feel excited.	1 (*Strongly disagree*) to 7 (*Strongly agree*)
	Right now, I feel tired.	1 (*Strongly disagree*) to 7 (*Strongly agree*)
	When I started the survey, I was:	*Alone* or *With other people*
	Are you interacting with other people?	*Yes* or *No*
Filler Items	Right now, my thoughts are pleasant.	1 (*Strongly disagree*) to 7 (*Strongly agree*)
	Right now, my thoughts are strange or unusual.	1 (*Strongly disagree*) to 7 (*Strongly agree*)
	Right now, my thoughts are clear.	1 (*Strongly disagree*) to 7 (*Strongly agree*)
	Right now, I can hardly control my thoughts.	1 (*Strongly disagree*) to 7 (*Strongly agree*)
	Right now, my thoughts are racing.	1 (*Strongly disagree*) to 7 (*Strongly agree*)
	Right now, I am thinking about a lot of things.	1 (*Strongly disagree*) to 7 (*Strongly agree*)
	Right now, I am having trouble concentrating.	1 (*Strongly disagree*) to 7 (*Strongly agree*)

APPENDIX D

INDIVIDUAL DIFFERENCES MEASURES

Measure	**Description**
Bucknell Auditory Imagery Scale (Halpern, 2015)	Survey measure (28 items) that assesses vividness and control of auditory imagery for voices, environmental sounds, and music.
NEO-FFI-3 (McCrae & Costa, 2010)	Survey measure (60 items) that assesses the Big Five personality traits.
Goldsmiths Musical Sophistication Index (Müllensiefen, Gingras, Musil, & Stewart, 2014)	Survey measure (38 items) that assesses five facets of musical expertise: active engagement, perceptual abilities, singing abilities, emotions, and musical training.
Cattell Culture Fair Tests (Cattell & Cattell, 1961/2008)	Based on a series of pictures arranged in a pattern, select the picture that follows next in the series.
Letter Sets	Based on five sets of four letters, determine which of five does not follow the pattern of the other four.
Number Series	Based on a series of numbers that follow a pattern, determine which number would appear next in the series.