Investigating Temporal and Melodic Aspects of Musical Imagery

Kelly Joan Jakubowski

Thesis submitted to the University of London for the degree of
Doctor of Philosophy

2015

Department of Psychology

Goldsmiths, University of London

Statement of Originality

This thesis represents my own original work towards this research degree.

Kelly Jakubowski

29 July 2015

Statement of Contributions

1) The Max/MSP application utilized for shifting the tempo of the stimuli in Studies 1 and 2 was developed by Dr. Mick Grierson (Goldsmiths, Computing Department). This application was adapted for further use in Study 3 by the present author and Dr. Nicolas Farrugia.

2) The tapping data analysis scripts used in Study 4 were constructed based on pre-existing tapping data analysis methods by Dr. Nicolas Farrugia and Sathish Sankarpandi, with feedback in the development stages from the present author.

3) In Study 4, the musical mode of each tune in the INMI dataset was coded independently by two research assistants, Suzanne Capps and Jessica McKenzie. Any discrepancies in coding between the two assistants was examined and resolved by the present author.

4) In Study 5, the manual extraction and verification of the data from the 3,000 online survey participants and the manual compilation of the 438 non-INMI "control" tunes dataset was conducted by the present author in conjunction with several interns to the Music, Mind, & Brain group: Katharina Bauer, Annick Odom, and Michelle Dickinson.

All other elements of design, testing, analysis, and writing are the work of the author.

Acknowledgements

I would like to thank Lauren Stewart for her advice as a supervisor over the past three years, particularly in terms of her detailed feedback on my work and dedication to carefully designed experiments. Thanks to Daniel Müllensiefen, my second supervisor, who first taught me how to think like a psychologist during my MSc and has offered a huge amount of advice and support ever since. I would also like to thank my officemates Georgina Floridou, Nicolas Farrugia, and Vicky Williamson for their friendship and fruitful discussions on musical imagery and music cognition in general. Thanks to several external collaborators who have given valuable long-distance feedback throughout my PhD, in particular Andrea Halpern.

Thanks to the participants in my studies, without whom this work would not be possible. Thank you to the Science Museum, London and Corrinne Burns for the opportunity to serve as a 'Scientist in Residence' at the museum to collect data for two experiments and to the Leverhulme Trust for providing funding for my PhD research and enabling it to go forward (grant number RPG-297 awarded to Lauren Stewart).

Finally, thanks to my family for all their support and encouragement over the years and to Steve for believing in me always.

Related Publications

A version of the research reported in Chapter 2 is published in:

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

A version of the research reported in Chapter 4 is published in:

Jakubowski, K., Farrugia, N., Halpern, A.R., Sankarpandi, S.K., & Stewart, L. (2015). The speed of our mental soundtracks: Tracking the tempo of involuntary musical imagery in everyday life. *Memory & Cognition.* doi: 10.3758/s13421-015-0531-5

A version of the research reported in Chapter 3 is under review in:

Jakubowski, K., Farrugia, N., & Stewart, L. (in review). Probing imagined tempo for music: Effects of motor engagement and musical background. *Psychology of Music.*

Additional publications related to the present work:

Farrugia, N., **Jakubowski, K.,** Cusack, R. & Stewart, L. (2015). Tunes stuck in your brain: The frequency and affective evaluation of involuntary musical imagery correlate with cortical structure. *Consciousness & Cognition, 35,* 66-77.

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): e86170.

Williamson, V. & **Jakubowski, K.** (2014). Earworms. In W. F. Thompson (Ed.), *Music in the Social and Behavioral Sciences.* Sage.

Liikkanen, L.A., **Jakubowski, K**. & Toivanen, J. (in press). Catching earworms on Twitter: Using big data to study involuntary musical imagery. *Music Perception.*

Abstract

Musical imagery, the mental replay of music in the absence of a perceived stimulus, is a common experience in the everyday lives of both musicians and non-musicians. The present research investigated aspects of the temporal and melodic content of musical imagery in an attempt to better understand factors underlying the genesis, stability, and potential functions of imagined music. Specifically, this research explored how factors such as one's current state of mood and arousal, level of musical experience, self-reported imagery abilities, and the musical properties of a song itself can influence the generation and stability of a musical image. The first three studies of the thesis were conducted in a laboratory context, in which participants deliberately imagined familiar tunes. The second two studies extended existing quantitative methods to examine the experience of *involuntary* musical imagery—music that is recalled spontaneously and repetitively within the mind—in more naturalistic contexts. The findings of the research have revealed that factors known to affect time perception, such as physiological arousal, can affect tempo representations within both voluntarily and involuntary musical imagery. The recall of tempo for imagined songs that exist in definitive versions was found to be fairly veridical in both voluntary and involuntary imagery, although temporal veridicality was also influenced by motor engagement with the imagery, a participant's musical background, and the original tempo of a song stimulus itself. Additionally, the occurrence of an involuntarily generated musical image was predicted by both extra-musical factors, such as song popularity, as well as intra-musical features related to a song's melodic contour and tempo. The results of this research suggest several parallels between voluntarily and involuntarily generated musical imagery and contribute a variety of novel findings regarding features underlying the genesis and possible functions of musical imagery.

Table of Contents

List of Figures

List of Tables

Overview

Music is an omnipresent stimulus in Western culture. A walk down any city street within our society might include glimpses of teenagers immersed in their iPods oblivious to the surrounding world, runners using music to boost motivation and steady their pace, a passing car with the radio blasting as a clear showcase of the driver's musical preferences, and shops and restaurants playing music to attract customers from specific social groups or economic backgrounds. These few examples already provide an array of reasons for which we, as humans, use music –for social, motivational, emotional, entertainment, and identity-based functions. The diversity of functions for music may in turn help to explain why it is used in so many different contexts and situations within our lives. One potential consequence of this ubiquity of music is that we may often find ourselves experiencing a mental replay of a tune in our heads (Sacks, 2007), which can extend long beyond the initial hearing.

Musical imagery, the mental replay of music in the absence of a perceived stimulus, is a common experience. Approximately 90% of respondents to a large-scale, Internet survey reported experiencing spontaneous episodes of musical imagery (here termed *involuntary musical imagery* or *INMI*; also commonly called *earworms*) at least once per week (Liikkanen, 2012a). A study of over 80,000 incidences of "earworms" reported on Twitter included data from 173 different geographic regions across the world (Liikkanen, Jakubowski, & Toivanen, in press). The purposeful recall of music within imagery (here termed *voluntary musical imagery* or *VMI*) is also a common experience, at least for musicians, who use this form of imagery regularly in mental practice, composition, and online preparation for upcoming material while performing (Clark, Williamon, & Aksentijevic, 2011; Keller, 2012; Mountain, 2001).

Researchers have often studied voluntary and involuntary musical imagery as separate entities. This is likely due to the different means by which each type of imagery is triggered-- it is easy to ask a participant to voluntarily imagine the song "Happy Birthday" in a laboratory; it is much more difficult to wait for the song to pop into his/her head spontaneously. As such, various retrospective (e.g., questionnaires) and ecological (e.g., diary studies) methods have often been used to study INMI (Beaman & Williams, 2010; Halpern & Bartlett, 2011; Liikkanen, 2012a), although a handful of laboratory studies have also been conducted (Hyman et al., 2013; Liikkanen, 2012b).

The divergence of these two areas of study has meant that often very different types of research questions have been asked regarding the two types of musical imagery. For instance, laboratory-based studies of VMI have often focused on examining which dimensions of music are represented within imagery (e.g., pitch, tempo, timbre, loudness, etc.) and the precision with which these features are represented in comparison to the original (perceived) version of a song (Crowder, 1989; Halpern, 1989; Levitin & Cook, 1996). Conversely, studies of INMI have focused more on descriptive, self-report evidence on its phenomenology, as well many extra-musical features of the experience, including cues that trigger an INMI experience, personality traits and individual differences that increase its occurrence, and affective responses to INMI (Beaman & Williams, 2010; Beaman & Williams, 2013; Liikkanen, 2012a; Müllensiefen, Jones, Jilka, Stewart, & Williamson, 2014; Williamson & Jilka, 2013; Williamson et al., 2011; Williamson, Liikkanen, Jakubowski, & Stewart, 2014).

The present research takes a more unified approach to investigating deliberately and spontaneously recalled musical imagery. This research utilizes a combination of behavioural, computational, and naturalistic methods to examine the imagery experience from several angles. Specifically, the first three studies in this thesis aim to clarify previously under-investigated aspects of temporal features of VMI, which help to

inform the second two studies that examine specific musical features (tempo and melody) of INMI. The thesis concludes by discussing VMI and INMI, not as separate entities, but rather in terms of the similarities and differences that the present work has revealed between the two types of musical imagery and how these findings can help to inform musical memory research in general. The results of this work allow for a more comprehensive understanding of the experience of musical imagery, both when generated deliberately within a laboratory context and when occurring naturally and spontaneously within the context of everyday life.

Chapter 1. Current Perspectives in Musical Imagery Research

Abstract

An introduction to previous research on musical imagery is presented. In particular, existing literature on temporal and melodic features of voluntary musical imagery is reviewed, followed by a review of the literature on involuntary musical imagery. Limitations and gaps in the previous literature are highlighted throughout, and methods used to study the various facets of musical imagery are compared and contrasted. The chapter concludes with an overview of the empirical work that will be presented within the subsequent chapters of the thesis.

1.1 Psychological Research on Imagery

The ability to imagine a scene, sound, or future action is a crucial part of the human experience that once was considered impenetrable to psychological research. Since the late 1960s, research on mental imagery has increased exponentially, and many subsequent studies have revealed parallels between the experiences of perceiving and imagining visual, spatial, and auditory stimuli (Finke, 1980; Hubbard, 2010; Kosslyn, 1980; Reisberg & Heuer, 2005; Shepard & Cooper, 1982). This increase in mental imagery research in the last few decades has been closely succeeded by an upsurge of research into "self-generated thoughts" (Callard, Smallwood, Golchert, & Margulies, 2013; Smallwood & Schooler, 2006, 2015), an umbrella term for concepts such as mind wandering, daydreaming, and spontaneous cognitions. It is evident that the study of the human mind is no longer constrained to measuring behavioural responses to perceivable, external stimuli, and innovative behavioural paradigms and the advent of neuroimaging techniques have allowed for a number of crucial advances in our understanding of inner mental experiences.

The study of mental imagery— whether it be visual, auditory, motor imagery, etc.—provides a unique type of insight into the human mind. As Neisser states, "if memory and perception are the two key branches of cognitive psychology, the study of imagery stands precisely at their intersection" (1972, pg. 233). In an MRI scan, perception and imagery of the same stimulus reveal very similar patterns of brain activations-- both visual perception and visual imagery activate the visual cortex, both auditory perception and auditory imagery activate the auditory cortex, etc. (Ganis, Thompson, & Kosslyn, 2004; Zatorre & Halpern, 2005). However, unlike pure perception of an external stimulus, imagery must also draw on memory mechanisms. That is, one must recreate an image based on one's own memory of the (previously) perceived stimulus' features. The preservation, or lack thereof, of these features can provide vital information about the degree to which certain features of the perceived stimulus are retained in memory. As such, imagery provides an important window into the intersection of our perceptual and memory systems.

1.1.1 Auditory Imagery

To date, imagery within the auditory domain has been less explored than visual imagery, but marks a nevertheless important part of our inner worlds. Auditory imagery can be defined as "the introspective persistence of an auditory experience, including one drawn from long-term memory, in the absence of direct sensory instigation of that experience" (Intons-Peterson, 1992, p. 46). Another similar definition is given by Baddeley and Logie (1992, p. 179): "What an auditory image involves is a conscious experience that resembles in certain, as yet unspecified ways, the experience of hearing the sound in question directly, but the image can be present in the absence of any auditory signal and it can be evoked intentionally by the subject." These definitions both implicate the two critical systems for imagery as described by Neisser (1972)—perception and memory. Both describe a perception-like experience in the absence

of direct perceptual input ("the introspective persistence of an auditory experience… in the absence of direct sensory instigation" and "the experience of hearing the sound in question directly, but the image can be present in the absence of any auditory signal") and also allude to memory recall mechanisms ("including one drawn from long-term memory" and "it can be evoked intentionally by the subject"). The issue of intentionality highlighted by Baddeley and Logie (1992) is also a key feature discussed within the musical imagery literature, in regard to voluntarily versus involuntarily generated musical images; this distinction will be discussed in detail throughout this chapter.

Research on auditory imagery has often explored the similarities and differences between perceived and imagined sounds, as well as the possible memory and maintenance mechanisms required to mentally rehearse sounds. A summary of the full scope of the auditory imagery literature is beyond the scope of this thesis, as the focus of this literature review will be to outline in detail previous findings within musical imagery research. However, it should be noted that a related body of literature pertaining to imagined speech (including the "inner voice") and imagined environmental sounds has accumulated in parallel to the literature on musical imagery (see Hubbard, 2010, for a review).

A main difference between auditory and visual imagery that has been noted by previous researchers (e.g., Halpern, 1992) is that auditory imagery extends in time rather than space. As such, several novel methods have been developed to study auditory imagery that extend or diverge from visual imagery paradigms. The temporal elements within auditory imagery bear perhaps more resemblance to features of motor imagery (Jeannerod, 1995) than visual imagery. As the goals of motor imagery often include planning and rehearsing future actions, mental preparation regarding the timing of these actions is a crucial element of this type of imagery (e.g., a baseball player imagining his actions before

hitting a ball). Paradigms used to study motor imagery can therefore be potentially informative to timing-related aspects of auditory imagery. Indeed, several researchers have investigated the crucial interactions between auditory and motor imagery implicated specifically in performing music (e.g., Brown & Palmer, 2012; Highben & Palmer, 2004; Keller, 2012).

1.1.2 Musical Imagery

Musical imagery research has comprised a seemingly large proportion of the research conducted within the auditory imagery domain. There are several reasons why musical imagery serves in many ways as an ideal stimulus for probing many of the questions asked by researchers of auditory imagery. Music combines a variety of dimensions of sound that are of interest to researchers of auditory imagery, including pitch, temporal and rhythmic aspects, timbre, loudness, and other expressive qualities. Perhaps the only other sound stimulus that reliably comprises all of these aspects is speech. However, features such as pitch and rhythmic regularity are much more salient and exaggerated in music than in speech (Patel, 2007). As such, music provides a valuable framework for studying these particular aspects of auditory imagery. Within a musical framework, it is also quite feasible and practical for researchers to manipulate these features (pitch, rhythm, timbre, etc.) in a systematic way for empirical research.

To extend Neisser's argument that "if memory and perception are the two key branches of cognitive psychology, the study of imagery stands precisely at their intersection" to the musical imagery domain, Kalakoski (2001, p. 43) adds "musical imagery not only stands at the intersection of memory and perception, but also at the intersection of several sense modalities." Indeed, researchers of musical imagery have more recently begun to focus on other modalities implicated within musical imagery besides the auditory dimension. Musical imagery in musicians often

comprises motor elements, which aid in action planning for performance (Keller, 2012). Interactions between multiple modalities are often seen as well; for instance, auditory imagery can be generated from a visual representation of a score ("notational audiation"; Brodsky, Henik, Rubinstein, & Zorman, 2003). A range of mental imagery techniques comprising several modalities are often implicated in musicians' preparations for musical performance, including visualization of a certain scene and rehearsal of the sounds and actions of an upcoming performance in one's mind (Connolly, 2002; Clark, Williamon, & Aksentijevic, 2011).

Previous research has often delineated two main categories of imagery—voluntary and involuntary musical imagery. Although a handful of studies have examined both categories of imagery without differentiation (e.g., Bailes, 2006, 2007a, 2015; Beaty et al., 2013), researchers have often focused on one type or the other. This is likely primarily due to the different types of methods required to successfully capture and examine each type of imagery. Most research on musical imagery to date has focused on voluntarily, deliberately recalled imagined music. For example, a researcher asks a participant in a laboratory experiment to purposefully recall a tune (e.g., "think of the song Happy Birthday") and asks the participant to describe features of this tune or to make some type of objective judgement based on the imagined representation of the song. More recently, researchers have begun to study involuntarily generated musical imagery, i.e., the experience of a tune appearing spontaneously in the mind and repeating. Various empirical methods for studying involuntary musical imagery have been developed to either covertly trigger an involuntary imagery experience or to study this type of imagery outside the traditional laboratory setting.

Due to the fact that most studies have examined voluntary and involuntary musical imagery separately, existing literature pertaining to

each of these types of imagery will be reviewed separately below. Conclusions and an overview of the thesis research goals that combine the two branches of musical imagery research are presented at the end of this chapter.

1.2 Voluntary Musical Imagery

1.2.1 Introduction

Voluntary musical imagery (hereafter abbreviated as *VMI*) is the conscious and deliberate mental replay of musical material. VMI is common amongst musicians, for example in mental rehearsal of familiar pieces of music or in composing new music (Brodsky et al., 2003; Mountain, 2001). VMI has most often been studied in controlled laboratory settings and is a type of imagery that can be purposefully generated by musical experts and non-experts alike (i.e., most people without any musical training can deliberately imagine a familiar tune playing in their heads). The study of VMI has provided valuable insights into how various dimensions of music, such as pitch, tempo, timbre, and loudness, are represented within imagery. Studying the veridicality of voluntarily generated musical images (as compared to the original version of a song) and their stability/consistency over time can also provide valuable insights into how precisely and vividly various features of music are reproduced within imagery and stored within long-term memory.

Literature pertaining specifically to the study of temporal and melodic aspects of VMI will be summarized in depth below, as these are the main points of interest to the present thesis. Although a number of studies have been conducted to examine additional musical dimensions of the VMI experience, such as imagined timbre (Bailes, 2007b; Crowder, 1989; Halpern, Zatorre, Bouffard, & Johnson, 2004; Pitt & Crowder, 1992), loudness (Bishop, Bailes, & Dean, 2013; Intons-Peterson, 1980; Pitt &

Crowder, 1992), and expressive features (Lucas, Schubert, & Halpern, 2010), this literature will not be reviewed in detail here as it is beyond the scope of the present work.

1.2.2 Temporal Aspects of VMI

In early research on temporal aspects of VMI, new paradigms were required that could take the time dimension into account, as many previous studies on imagery had focused on static images or images extending across space in the visual domain (e.g., Kosslyn, 1980). One new paradigm was developed by Halpern (1988a), who conducted a series of studies by adapting a mental scanning task from the visual imagery literature (Kosslyn, Ball, & Reiser, 1978) for use in the auditory domain. Participants in Halpern's studies were presented two song lyrics in succession and were asked to indicate if the second lyric was from the same song as the first lyric. An adaptation of this task in which participants were asked whether the second lyric was higher or lower in pitch than the first lyric was also tested. In both versions of the task, the amount of time to complete the task increased incrementally as the distance between the two lyrics increased. These results suggest that participants were mentally scanning through the song in order to make their judgements. This study provided evidence that temporal aspects of music were indeed represented within auditory imagery, thus forming a basis for future research into more detailed aspects of imagined musical tempo.

1.2.2.1 Parallels Between Perceived and Imagined Tempo

Several studies have highlighted the parallels between perceived and imagined musical tempo. Weber and Brown (1986) asked participants to track the contour of melodies as fast as possible by tracing the ups and downs of each melody with a pencil on paper. Participants did this while singing each melody out loud (perception task) and also while

deliberately imagining each melody (imagery task). No significant difference was found in the amount of time taken to complete each of these two tasks. The authors concluded that this finding provides evidence that participants were utilizing a common mental representation in both the sung and imagined reproduction tasks. Halpern (1988b) also compared perceived and imagined tempo representations by asking participants to make judgements regarding the tempo of familiar, non-canonical songs, such as "London Bridge" and "Happy Birthday". Participants adjusted the tempo of the music to the speed that felt "correct" to them in 1) a perception condition while hearing a song aloud and 2) an imagery condition while imagining the same song in their heads. In the imagery condition, the tempo choice was made by selecting a setting on a metronome to match the beat of the imagined music. A high correlation emerged between perceived and imagined tempi across all trials ($r = .63$), indicating that a song was generally imagined at a similar tempo to a participant's preferred tempo for hearing the same song aloud.

It should be noted that, while the above studies provide general evidence for the implementation of shared mental resources in perceiving and imagining music, neither the lack of a significant difference in means between a perception and imagery condition (Weber & Brown, 1986) nor a significant correlation between perceived and imagined tempo ratings (Halpern, 1988b) provide evidence that perceived and imagined tempo judgements were exactly equivalent. These analyses also do not provide precise measures of how close tempo ratings in perceived and imagined conditions were to one another in absolute terms (e.g., bpm differences between the two conditions). The research conducted in this thesis aims to more objectively compare potential differences between perceived and imagined tempo judgements (Study 3), as well as to investigate possible factors that contribute to non-equivalence of perceived and imagined tempo judgements, such as stimulus familiarity, musical training, and auditory imagery abilities (Studies 1 and 3).

1.2.2.2 Consistency of Imagined Tempo

The consistency of imagined tempo across multiple trials and/or multiple days has been investigated in several contexts. Clynes and Walker (1982) asked musicians to complete several trials of tapping to an imagined version of Mozart's C Major Piano Concerto, K. 467 on two consecutive days. Tapping speed when imagining this piece of music was significantly more consistent than when the musicians tapped to a verbally imagined phrase ("Saturday, Sunday") or when they were asked to tap without any mental image. In a similar study, participants tapped to imagined versions of the same song across multiple trials within the same session and in different sessions that were separated by 2 to 5 days (Halpern, 1992). Musician participants were highly consistent in their tapped tempo for the same song within the same session and across different sessions, with average standard deviations of around 6 bpm. Non-musicians were very consistent within a single session, but were much less consistent in their performance across different sessions. Additionally, Johnson (2011) compared the temporal consistency of two imagery-based mental practice techniques in university music students: 1) a non-motor imagery condition, in which participants purely imagined a piece of music using auditory imagery, and 2) a motor imagery condition, in which participants mentally rehearsed the same piece using both auditory imagery and motor imagery of the movements involved. Contrary to the hypothesis of the study that the added motor imagery component would increase temporal accuracy, Johnson found no difference in temporal accuracy between these two imagery strategies as compared to a performed version of the same piece of music. However, the results indicated that participants became more consistent in terms of their imagined tempo as the experiment progressed.

The results of the studies summarized above provide some general evidence for the temporal consistency of imagined songs across multiple

trials and multiple days. However, Halpern's (1992) research highlights a key difference in terms of more long-term temporal consistency within musical imagery that may be at least partially accounted for by musical training. Johnson's (2011) work also sheds light on the role of practice effects, as the musician participants within his study tended to become more consistent at the imagined tempo task even within the relatively short duration of the testing session. These findings suggest that the generation of consistent tempo representations within musical imagery may be dependent on prior training or practice and that consistency of tempo recall within imagery can be developed and refined, at least to a point. The influence of musical training, as well as the influence of potentially related abilities such as engagement with music and auditory imagery abilities, will be investigated in further detail throughout this thesis in order to further clarify the role of such factors on the generation of temporal aspects of musical imagery.

It is also worth noting here that similar work to the aforementioned studies on temporal consistency in the imagined music domain has been conducted in the domain of musical performance. Recordings by the same performer over the course of months, and even years, have been compared in several studies. These studies revealed that, in general, performances were remarkably consistent in terms of tempo and/or overall performance duration (Bailes & Barwick, 2011; Clynes & Walker, 1982; Collier & Collier, 1994). This long-term temporal consistency for music was even found to extend to ensemble (string quartet) performances (Clynes & Walker, 1986). It is plausible that these two areas of study are highly related, i.e., that musicians may be generating and drawing upon the stable mental tempo representations revealed in studies of musical imagery in order to maintain highly consistent tempi across multiple musical performances, even over a period of many years.

1.2.2.3 Veridicality of Imagined Tempo

Levitin and Cook (1996) asked non-musician participants to sing two familiar pop songs from memory in order to investigate the veridicality of memory for musical tempo. The tempi of the sung productions were compared to the tempi of the original recordings of the songs. A total of 72% of trials within the study came within 8% of the original, recorded tempo. The authors cite several previous studies that estimate the JND (just-noticeable difference) for tempo to be around 8%. Thus, it was concluded that most participants exhibited a type of "absolute memory" for the tempo of familiar music.

It should be noted that the aforementioned work by Levitin and Cook might not be classified strictly as a study of musical imagery, as the task involved singing aloud rather than imagining music in one's head. Levitin and Cook did instruct participants to form a mental image of their chosen songs before singing them aloud, although no objective measurement was taken to affirm that participants actually formed mental images before singing. Additionally, the act of singing aloud may have influenced one's mental image and participants may have adapted the tempo of their performances while singing based on auditory feedback cues. It could be presumed that participants were using a form of *anticipatory* imagery (Keller, Dalla Bella, & Koch, 2010; Keller, 2012), described in studies of musical performance as auditory and motor imagery of one's upcoming actions during performance, which "facilitates the planning and execution of musical actions" (Keller, 2012, p. 209). However, it is not possible to discern from the data presented by Levitin and Cook whether anticipatory imagery was indeed being generated, or whether other strategies underlie these highly veridical sung productions. More work is needed to understand the mechanisms underlying the veridicality of tempo recall, and comparisons of tempo recall in pure imagery versus perceptual tasks can be useful in this regard. Such questions will be addressed in Study 3 of the present thesis.

1.2.2.4 Expressive Timing

A less commonly investigated temporal dimension of imagined music is expressive timing. Repp (2001) examined timing profiles of perceived and imagined performances by pianists across a variety of conditions, including expressive performance on a keyboard, expressive performance on a keyboard without auditory feedback, and tapping to an imagined expressive performance. In the conditions without auditory feedback it was presumed that a musical image must be generated in order to complete the task. Very high intra-subject correlations (ranging from 0.84 to 0.98) were found between all conditions, suggesting expressive elements of timing were often vividly represented within musical imagery. The data also suggested that some participants were able to imagine expressive elements of timing more accurately than others, although the potential causes underlying these individual differences were not explored. It is possible that these differences in performance between pianists were related to duration of musical training or prior experience in musical imagery tasks, as these two factors have been implicated in other studies of imagined tempo (Halpern, 1992; Johnson, 2011), but further research is needed to verify such an account.

1.2.2.5 Limitations on Imagined Tempo Representations

Despite the variety of research highlighting the parallels between perceived and imagined tempo for music, some work has also reported certain degradations of temporal information within musical imagery. Janata and Paroo (2006) examined the acuity of both pitch and temporal aspects of musical imagery. In a perception condition, participants heard a full ascending diatonic scale. In an imagery condition, participants heard the beginning of the same scale played aloud and were asked to imagine either three or five notes immediately preceding the final note of the scale. In both conditions the final note of the scale was heard aloud and participants were asked to judge either a) whether the pitch of the

final note was in tune (pitch task) or b) whether the final note was on time relative to the beat of the music (timing task). Participants' performance in judging the intonation of the final note in the imagery condition did not significantly differ from their performance on the same task in the perception condition. However, judgements of timing were impaired in the imagined music condition relative to the perception condition. This result was replicated in a second experiment using a psychophysical staircase procedure, which allowed for estimations of thresholds for each condition for each participant. The authors suggest, based on these two experiments, that temporal aspects of musical imagery may be more susceptible to distortion than the pitch content of imagery. However, performance in the imagery condition (for both the pitch and timing tasks) was more accurate overall in participants with more musical training, suggesting that the precision of imagery can be developed with increasing expertise.

Two additional studies have provided support for the findings reported by Janata and Paroo (2006). A study by Weir, Williamson, and Müllensiefen (2015) used an imagined continuation paradigm, in which an excerpt of a pop song began playing aloud and was then muted for approximately 10 seconds. The music returned after the muted section and participants were required to make a judgement as to whether the return of the song was 1) in the right key and 2) whether it was on time. During the muted section, participants were asked to imagine the song in their heads, to aid their subsequent judgements of the pitch and timing of the song when it came back in. Although a main focus of the paper was a comparison of VMI abilities to the frequency of experiencing *involuntary* musical imagery, the result of particular note to the present discussion was that the pitch judgements were overall more accurate than the timing judgements in this task, although performance in both conditions was above chance level. Bailes (2002) reported similar findings in a study comparing perception and imagery conditions using familiar folk melodies. Her study utilized separate pitch and timing manipulations;

overall, the design was quite similar to that used by Janata and Paroo (2006). In Bailes' imagery condition, the pitch tasks were completed significantly more accurately than the timing tasks, despite the finding that reaction times for the pitch tasks were significantly slower than for the timing tasks. Bailes points out that the pitch and timing tasks were not specifically constructed to be analogous to one another in terms of difficulty (the two tasks were initially constructed to be examined separately), and the same might be true of the work by Weir and colleagues. However, taken together with the findings of Janata and Paroo (2006), these three studies all suggest that timing-related aspects of musical imagery appear to be more quickly distorted or less vividly represented within musical imagery than pitch aspects. The work in Studies 1 through 3 of the present thesis aims to provide some first explanations as to *why* temporal images may be susceptible to distortion, by testing the influence of factors known to influence time perception and individual differences in musical background on imagined tempo recall for familiar music.

Halpern (1988b) observed another notable difference between perceived and imagined musical tempi. Despite the fairly high correlation found between perceived and imagined tempi (as reported in Section 1.2.2.1), Halpern also reported evidence of a "regression to the mean" within imagined music. That is, songs for which participants chose quite slow tempi as the speed that sound "correct" to them in a perceived music condition tended to be imagined *faster* than their preferred perceived tempo. Songs for which the chosen tempo was set quite fast in the perceived music condition tended to be imagined *slower* than the tempo selected in the perceived music condition. One reason for this might be that there exist certain limits on imagined tempo capacity, such that it is difficult or impossible to imagine songs at very fast or very slow tempi. However, Study 2 of Halpern (1998b) examined specifically the tempo limits at which familiar songs could be purposefully imagined. The mean fastest imagined tempo was 164 bpm and the mean slowest imagined

tempo was 65 bpm, suggesting a wide range of tempi at which familiar songs could be imagined. Thus, it appears that participants have the capacity to imagine music at a fairly wide range of tempi, but may only reach the very extremes of this range when specifically instructed to do so.

Finally, Clynes and Walker (1982) reported research that compared durations of performed and imagined pieces of music. Musician participants were asked to complete three tasks: 1) perform a piece, 2) imagine the same piece, and 3) imagine the same piece while conducting. On average, both imagined versions of the pieces (imagining and imagining while conducting) took place over significantly longer durations than the performed version, although the duration of the conducted version was generally closer in duration to the performed version than the purely imagined version. These results are somewhat at odds with Halpern's (1988b), as there was no "regression to the mean" present for imagined tempi in Clynes and Walker's study. However, it is unclear from the presented results whether the pieces of music performed by the musicians encompassed a large tempo range. For instance, if all or most of the pieces of music within Clynes and Walker's study were relatively fast in terms of performance tempo, Halpern's findings would indeed predict that the pieces would be imagined at overall slower tempi than that at which they were performed.

1.2.2.6 Summary of Findings on Imagined Tempo

As is evident from the above literature, a variety of different tasks have been used to measure the tempo of imagined music. Table 1.1 summarizes the different tasks used, with potential benefits and limitations of each. Study 3 of this thesis will make direct comparisons between several of these tasks within the same participant, thus providing a sensitive measure of the extent to which methodological differences can influence task performance when all other factors are held constant.

In sum, early studies (Halpern, 1988a; Weber & Brown, 1986) provided evidence that tempo information is represented within music imagery. Subsequent findings indicate that imagined tempo representations are quite consistent over time (Clynes & Walker, 1982), although long-term representations appear to be more consistent for musicians than non-musicians (Halpern, 1992). Some findings suggest a high degree of veridicality within imagined tempo (Levitin & Cook, 1996), although this has not yet been tested using a method in which the imagery remains purely imagined, without the assistance of perceptual cues. Findings also suggest that representations of tempo may be more susceptible to degradations within imagery than mental representations of pitch (Bailes, 2002; Janata & Paroo, 2006; Weir et al., 2015), although explanations of why this is are still lacking. Finally, several studies suggest a range of individual differences in the ability to represent tempo information in a consistent and veridical form (Halpern, 1992; Janata & Paroo, 2006; Johnson, 2011; Repp, 2001). However, factors that can explain these performance differences are not well understood, aside from an influence of formal music training that has been reported in some work (Halpern, 1992; Janata & Paroo, 2006).

Table 1.1: *Tasks Previously Used to Measure Imagined Tempo*

Task	Studies using this task	Description	Advantages	Disadvantages
Entire duration measurement	Clynes & Walker (1982), Johnson (2011), Weber & Brown (1986)	Experimenter calculates the total duration of an imagined/ performed musical excerpt	-Useful for short excerpts of music -Less invasive: participants can focus on imagining the music without a concurrent task	-Rough, overall estimate of duration -Cannot account for tempo changes within the excerpt -No objective indication that participants completed the task
Mental scanning task	Halpern (1988a), Zatorre & Halpern (1993)*, Zatorre et al. (1996)*	Participant compares some aspect (e.g., pitch) of two elements of a piece of music and experimenter measures amount of time taken to compare these two aspects	-Participant responses required (e.g., was the second pitch higher or lower than the first?), correct responses serve as an indirect indicator that imagery occurred	-Difficult to derive an overall tempo calculation (bpm) -Cannot account for tempo changes within the excerpt
Tapping	Clynes & Walker (1982), Halpern (1992), Repp (2001)	Participants tap along to the beat of imagined music	-Can estimate both average tempo (bpm) and tempo variability -Increases or decreases in tempo can be detected	-Relies on motor constraints of participants -Relies on beat extraction abilities of participants

Metronome adjustment	Halpern (1988b)	Participants adjust a metronome to the beat of imagined music	-Provides a measure of exact tempo in bpm	-Relies on beat extraction abilities of participants
			-Does not rely on motor constraints of participants (as in tapping)	-Possible auditory interference of metronome sound with imagery
Imagined continuation	Bailes (2002); Janata & Paroo (2006); Weir et al. (2015)	A stimulus is played and then muted, the music returns after the muted period and participants make a judgement about it (e.g., was the music on time?)	-Participant responses required, correct responses serve as an indicator of the veridicality of the image	- Residual information from the perceived stimulus may be retained in echoic memory
Imagine then produce	Levitin & Cook (1996)	Participants imagine a piece in their head and then complete a production task, such as singing the piece	-Easy task for non-musicians to complete	-Performance in a singing task may be influenced by vocal constraints of the participant
Perform on an instrument without auditory feedback	Repp (2001)	Participants perform a piece without auditory feedback (e.g., on a muted keyboard)	-Close simulation of the actual conditions of musical performance	-Non-musicians cannot easily participate

* Denotes a neuroimaging study; see Section 1.2.4 for a description of the study.

1.2.3 Melodic Aspects of VMI

1.2.3.1 Pitch Imagery

As in early studies of temporal aspects of VMI, the first studies of VMI for pitch aimed to investigate whether pitch was represented within musical imagery. Farah and Smith (1983) addressed this question using a staircase design in which participants imagined a tone and then heard a second tone aloud. When the pitch of the tone played aloud was congruent to the pitch of the imagined tone, participants were able to detect tones at a much lower volume of presentation than when the two tones differed in pitch. This finding suggests that imagery for pitch can actually facilitate subsequent pitch perception. In later research by Okada and Matsuoka (1992), participants were asked to imagine a tone of 800 Hz and then detect which of five target tones was presented during concurrent white noise. In this case, the auditory imagery *interfered* with perception in terms of detecting the target tone when the target tone matched the imagined tone. Although this result appears at first to conflict the earlier findings of Farah and Smith, Okada and Matsuoka point out that they employed a different methodology (ascending method of limits) and focused on detection rather than discrimination. Hubbard (2010) notes that this discrepancy also parallels findings by Finke (1986) in the visual domain, whereby imagery improved performance on discrimination but not detection tasks. Thus, it appears that imagined pitches interact with perceived auditory stimuli in similar ways to the interactions of visual imagery with perceived visual stimuli.

1.2.3.2 Consistency of Pitch Imagery

The consistency of pitch representations within musical imagery across multiple trials has also been investigated. Halpern (1989) explored pitch consistency by asking participants to imagine the starting pitch to traditional, non-canonical songs such as "Twinkle, Twinkle Little Star".

Participants were then asked to sing or play on a keyboard (the keys of which were obstructed from view) the pitch they were imagining. Participants were quite consistent in producing the same or close to the same absolute starting pitch for a given song across multiple trials (mean *SD* = 1.28 semitones), suggesting that the absolute pitch level of an individual's auditory imagery for a specific song is quite stable over time. This study also examined productions of the starting pitch to the same song in different sessions separated by 48 hours. Similarly to Halpern's (1992) findings on long-term tempo consistency, pitch productions across the two sessions were highly stable for musicians, but much more variable for non-musicians.

1.2.3.3 Veridicality of Pitch Imagery

Levitin (1994) examined the veridicality of mental representations for pitch. Participants in this research sang two self-selected pop songs and the pitches of their productions were compared to the pitches of the original recordings of the songs. Overall, 40% of participants reproduced the correct starting pitch for a song on at least one trial, and 12% of participants produced the correct starting pitch on both trials. Levitin's study was subsequently replicated in a sample of 277 participants across six laboratories (Frieler et al., 2013). Overall, performance on the sung production task in the replication study was significantly above chance level, but the reported effect size was smaller than Levitin's. The data also suggested a wide range of individual differences in performance; however, no significant predictors (e.g., musical training, age, familiarity) were found to explain performance differences. A subsequent study of additional predictors of performance on this task by Jakubowski and Müllensiefen (2013) found that relative pitch memory abilities and emotional associations with the chosen songs significantly impacted on performance on the sung production task used by Levitin (1994) and Frieler et al. (2013).

One caveat of the studies of Levitin (1994) and Frieler et al. (2013), as well as the study by Halpern (1989) discussed above, is that singing might not always be an accurate reproduction of one's mental image (e.g., due to vocal range constraints or production difficulties). As discussed in Section 1.2.2.3, auditory feedback provided during singing could also alter or distort one's mental image, thereby influencing participants toward producing a tone somewhat different to the original imagined pitch. Additionally, no objective measure was taken in these studies to indicate that participants had actually formed a mental image before singing. However, taken together, these studies provide at least some first indication that pitch is preserved with high consistency and veridicality within musical imagery.

1.2.3.4 Comparisons of Pitch Imagery in Multiple Tasks

As discussed in Section 1.2.2.5, Janata and Paroo (2006) compared the acuity of pitch and temporal aspects of musical imagery. Participants' intonation judgements in a musical imagery condition were not significantly impaired as compared to intonation judgements made in a perceived music condition, suggesting that the ability to make fine-grained judgements (e.g., intonation) about musical pitches is preserved even when a section of the music is imagined without any perceptual input. However, participants did perform better in this task when they imagined three notes of the scale compared to when they were asked to imagine five notes, suggesting that they were aided by the additional sensory input when it was available.

Navarro Cebrian and Janata (2010) and Janata (2012) extended this work by comparing three types of imagery tasks that varied in their degree of sensory input. In a two-tone discrimination task (condition 1), a participant heard two pairs of notes in quick succession (600 ms apart) and indicated in which of the two pairs the two tones differed. In this task, an image of the first pair of tones needs to be maintained in sensory

memory for only a very brief period. In a cued-attention task (condition 2), participants heard two major scales on each trial and indicated which of the two scales contained a mistuned final note. This task was said to provide a moderate degree of sensory input, as an anticipatory image must be generated for the final tone of the scale, but the preceding sensory input (the other notes of the scale) provides a strong tonal context to aid image formation. In a third task (condition 3), participants heard the first four notes of a scale, imagined the next three, and then heard the final note and judged its intonation (similarly to the task used by Janata & Paroo, 2006). This context provided the least amount of sensory support. The main findings indicated that performance was most accurate in the condition with greatest sensory support (condition 1) and least accurate in the condition with the least sensory support (condition 3). An overall correlation between performance on these imagery tasks and musical training was also reported. This study provides a clear comparison of tasks used to measure the pitch dimension of musical imagery and suggests that differences in findings between previous studies may be at least partially explained by the degree of sensory input implicated in the imagery task. This work will serve to inform Study 3 of this thesis, which takes a similar approach to compare performance on different types of imagery tasks that vary in terms of sensory input, but with a focus on imagined *tempo,* rather than pitch.

1.2.3.5 Imagined Chords/Harmony

Hubbard and Stoeckig (1988) investigated the degree to which *simultaneous* pitches can be imagined. Their work compared imagery for single tones to imagery for three-note chords. Participants were presented a cue tone or chord and were asked to imagine the tone or chord a whole step (two semitones) higher than the cue. When a participant had formed an image, he/she pressed a button and a comparison tone or chord was presented aloud. The participant was asked to judge whether the comparison tone/chord was higher or lower

than the imagined tone/chord. It was found that image formation time was significantly longer in the chord condition than the tone condition, suggesting that more complex musical stimuli require more time and/or effort to generate. The authors also highlighted the fact that participants in previous music perceptual studies had performed more accurately in a same/different task when the target was the same rather than different, and that accuracy for different-unrelated[1] targets was greater than for different-related targets (e.g., Bharucha, 1984; Bharucha & Krumhansl, 1983; Cuddy, Cohen, & Miller, 1979). Hubbard and Stoeckig tested whether similar results could be revealed using their imagined tone/chord paradigm and were indeed able to replicate these results, providing evidence for the implementation of similar mechanisms in perception- and imagery-based musical comparison tasks.

1.2.3.6 Summary of Findings on Imagined Pitch

Previous research has provided a variety of evidence indicating that pitch is represented vividly within musical imagery and that many parallels exist between the perceptual experience of hearing a pitch and the mental experience of imagining a pitch (Farah & Smith, 1983; Hubbard & Stoeckig, 1988; Janata & Paroo, 2006). The various types of tasks that have been used to measure imagined pitch representations are summarized in Table 1.2. Evidence for highly consistent and veridical representations of pitch within imagery has been revealed (Frieler et al., 2013; Halpern, 1989; Levitin, 1994) and performance on imagined pitch tasks tends to be more accurate when a greater degree of sensory input is provided (Navarro Cebrian & Janata, 2010; Janata, 2012). As in the imagined tempo domain, positive effects of musical training have been found on performance in some imagined pitch tasks (Halpern, 1989).

---

[1] Different-related targets were defined in Hubbard and Stoeckig's study as a tone/chord two semitones above the imagined tone/chord and different-unrelated targets were defined as a tone/chord five semitones above the imagined tone/chord.

Table 1.2: *Tasks Previously Used to Measure Imagined Pitch*

Task	Studies using this task	Description	Advantages	Disadvantages
Imagined continuation	Bailes (2002); Janata (2012); Janata & Paroo (2006); Navarro Cebrian & Janata (2010); Weir, Williamson, & Müllensiefen (2015)	A stimulus is played and then muted, the music returns after the muted period and participants make a judgement about it (e.g., was the music in tune?)	-Participant responses required, correct responses serve as an indicator of the veridicality of the image	-Residual information from the perceived stimulus may be stored in echoic memory
Imagine then produce	Frieler et al. (2013); Halpern (1989); Levitin (1994)	Participants imagine a piece in their head and then complete a production task, such as singing the piece	-Easy task for non-musicians to complete	-Performance in a singing task may be influenced by vocal constraints of the participant
Judgement of a perceived stimulus based on imagined pitch	Farah & Smith (1983); Hubbard & Stoeckig (1988); Okada & Matsuoka (1992)	Participants imagine a pitch or chord and make a judgement about a subsequent perceived stimulus based on the image (e.g., is this pitch the same/different to the imagined pitch?)	-Participant responses required, correct responses serve as an indicator of the veridicality of the image	-Participants generally need to be exposed to a perceived stimulus before imagining-- residual information from the perceived stimulus may be stored in echoic memory
Two-tone discrimination task & cued-attention task	Navarro Cebrian & Janata (2010); Janata (2012)	Participants hear two stimuli in succession and make a comparison between them (an image of the first stimulus must be held in the mind for comparison to the second)	-Participant responses required, correct responses indicate that the first stimulus was successfully held as an image for comparison to the second	-Not traditionally used as an imagery task due to large degree of sensory input -Does not measure long-term pitch memory recall

1.2.4 The Neuroscience of VMI

Neuroimaging techniques such as PET, MRI, and EEG have provided a valuable window into the cognitive mechanisms underlying internal mental processes such as musical imagery. Neuroscientific research has revealed several parallels, as well as some inherent differences, between the brain regions recruited in perceiving and deliberately imagining music (see Zatorre & Halpern, 2005, for an overview). Similar activation patterns of the secondary auditory cortex have been found across several studies comparing perceived music to VMI (Kleber, Birbaumer, Veit, Trevorrow, & Lotze, 2007; Kraemer, Macrae, Green, & Kelley, 2005; Zatorre & Halpern, 2005). Primary auditory cortex activation has been found during VMI only under certain conditions, such as when participants imagined familiar songs with no lyrics (Kraemer et al., 2005). This finding suggests that the primary auditory cortex may play a less fundamental role in musical imagery than the key role it plays during music perception (Kleber et al., 2007). Bilateral frontal cortex activations have also been found during VMI tasks and may relate to the memory retrieval component inherent in imagery (Herholz, Halpern, & Zatorre, 2012; Zatorre, Halpern, & Bouffard, 2010; Zatorre, Halpern, Perry, Meyer, & Evans, 1996). Activation of the supplementary motor area (SMA) in multiple studies of VMI (Halpern et al., 2004; Herholz et al., 2012; Zatorre et al., 1996) indicates the potential involvement of a motor or sequencing component, perhaps related to the rehearsal or maintenance of the image within working memory.

The methods used to study the brain basis of VMI have predominantly emerged as adaptations of the behavioural methods for measuring various dimensions of imagery described in earlier sections of this chapter. For instance, two of the earliest studies of the neural correlates of musical imagery used a task based on the mental scanning paradigm of Halpern (1988a; as described in Section 1.2.2) to test patients with unilateral temporal lobe lesions (Zatorre & Halpern, 1993) and healthy

participants using PET technology (Zatorre et al., 1996). The development of elegant and effective behavioural paradigms is paramount to neuroscientific research on musical imagery. As musical imagery is a fairly intangible and private experience, paradigms are needed that can reliably ensure a musical image is being generated at a particular point in time in order to effectively capture this experience in an MRI scan or EEG data.

1.2.5 Functions of VMI

Musicians report that they use VMI for mental rehearsal while learning pieces of music and preparing for performance (Clark, Williamon, & Aksentijevic, 2011; Conolly, 2002; Haddon, 2007). Such mental rehearsal strategies often take a multimodal approach--combining auditory imagery with imagery of the movements required to perform a piece, visual imagery of a musical score, etc. Several studies suggest that deliberate mental practice of a piece is a useful strategy for learning and rehearsing music, though generally only when employed in conjunction with actual physical practice (Coffman, 1990; Ross, 1985). In music education, the importance of being able to generate an image of a musical sound before playing it is widely recognized and is often referred to via the term "audiation", as coined by the influential music educator, Edwin Gordon (1999, 2003).

Musicians also use VMI as an aid to action planning during performance (Keller, 2012). A correlation between sensorimotor synchronization and auditory imagery abilities has been found in musicians, suggesting that being able to anticipate the sound of a musical event before it occurs might aid in the ability to synchronize with it (Pecenka & Keller, 2009). Anticipatory auditory imagery generated during musical performances has also been found to facilitate the accurate timing of movements (Keller & Koch, 2006) and can aid in controlling the speed and force of

movements (Keller et al., 2010; Keller & Koch, 2008), thus helping to optimize the precision and efficiency of motor aspects of a performance.

Finally, musical imagery may serve wider functions in terms of expectation generation during music listening, as described by Janata (2001). When one listens to a piece of music, expectations are often generated about the next note to be heard in a sequence. An image of an expected note may be generated just before the note is actually played. If the perceived note matches the image, expectations are fulfilled; if not, expectations have been thwarted. The role of fulfilment and violation of expectations as key sources of our enjoyment and interest in music has been discussed in the music psychological and music theoretical discourse for many years (e.g., Meyer, 1956; Huron, 2006).

1.3 Involuntary Musical Imagery

1.3.1 Introduction

Involuntary musical imagery (INMI) can be defined as "a fragment of music that comes to the mind involuntarily and repeats at least once, on a loop, without conscious effort" (Williamson & Jakubowski, 2014). This definition highlights two primary features of the INMI experience: 1) it is recalled via associative and unplanned retrieval mechanisms and 2) it is involuntarily repetitive in nature. These two characteristics serve to distinguish INMI from other related cognitions such as voluntary musical imagery, which is imagined music that is strategically retrieved, as described in earlier sections of this chapter, musical *mind pops* (Kvavilashvili & Anthony, 2012), which comprise brief, single spontaneous appearances of a tune in the mind without repetition, and musical hallucinations (Griffiths, 2000), which are mental representations of musical sounds that are misattributed as originating from the external environment.

INMI has been referred to by various terms, perhaps most commonly by the term *earworm*, which originates from the German term for the phenomenon, *Ohrwurm*. INMI has also been referred to as *stuck song syndrome* (Levitin, 2006), *brainworms* or *sticky music* (Sacks, 2007), *Musical Imagery Repetition* (*MIR*; Bennett, 2002) and *intrusive songs* (Hyman et al., 2013) (see Williams, 2015, for an extensive overview of the various terms used to refer to the INMI experience). Throughout this thesis the term *involuntary musical imagery* (*INMI*; Liikkanen, 2008) will be consistently used, as this term represents a fairly comprehensive description of the phenomenon. The term *earworm* will be also be used specifically in the context of describing instructions given to participants in the experiments conducted for the thesis; this term was chosen for use in communications to participants, as it was a more familiar, colloquial expression.

Although the English term *earworm* entered the common vocabulary fairly recently (it first appeared in the *Oxford English Dictionary* in 2011), the INMI phenomenon has been alluded to in literature for over 100 years, suggesting it is not a specifically modern experience (Williamson & Jakubowski, 2014). For example, Mark Twain writes of an author who is plagued by a musical jingle that is stuck in his head day and night until he passes it on to a friend in the short story "Punch, Brothers, Punch!" In Arthur C. Clarke's "The Ultimate Melody", a scientist strives to develop the perfect melody and, in doing so, creates a tune that becomes so in sync with his brain waves that he goes into a catatonic state. Although these historical references demonstrate that having a tune stuck in one's head is not an exclusively modern experience, Oliver Sacks (2007) suggests in his popular science book *Musicophilia* that INMI may be more widespread and frequent in modern culture, due to the omnipresence of music as a result of recording technologies and portable music players. Despite these various literary, historical, and popular science references

to the phenomenon, INMI has only become a topic of psychological investigation in the past 10 to 15 years. This psychological literature will be reviewed in detail in the following sections.

1.3.2 Methods for Studying INMI

A main difference between INMI and VMI is that INMI is initially activated in a spontaneous and non-deliberate fashion. As such, INMI is less amenable to study within a laboratory setting and new methods have been devised to attempt to capture this experience for the purpose of empirical investigation. That being said, some researchers have successfully induced INMI within a laboratory setting, either by exposing participants to actual music (Floridou, Williamson, & Stewart, 2014; Hyman et al., 2013) or using cued-recall tasks in which participants are asked to recall the lyrics from familiar songs, thereby priming them for an INMI experience (Floridou, Williamson & Müllensiefen, 2012; Liikkanen, 2012b). However, induction rates in the laboratory can be low, with many participants reporting few or no INMI. Induction rates often depend on multiple factors such as concurrent task demands, song familiarity, and recency of exposure to a song (Floridou, Williamson, & Stewart, 2014; Hyman et al., 2013; Liikkanen, 2012b).

Questionnaire-based research has also been common, in which participants provide overall ratings of different aspects of the INMI experience (Beaman & Williams, 2010; Floridou, Williamson, Stewart, & Müllensiefen, 2015; Halpern & Bartlett, 2011; Liikkanen, 2012a). This method can be useful particularly in terms of collecting data from a large number of participants about more general aspects of the INMI phenomenon, such as the frequency with which it occurs and the overall attitudes and reactions to the experience expressed within the population. One disadvantage of questionnaires is that they rely on retrospective reports from participants and, as such, can be subject to recollection bias. Participants asked to report on their INMI experiences via questionnaires

may also not be able to describe the experience as vividly as when they are actually experiencing an INMI episode.

A number of researchers have employed more naturalistic methods to study INMI, namely Experience Sampling Methods (ESM) and diary studies (Bailes, 2006, 2007a, 2015; Beaman & Williams, 2010; Beaty et al., 2013; Byron & Fowles, 2013; Floridou & Müllensiefen, 2015; Halpern & Bartlett, 2011). ESM is a technique that was first developed by Larson and Csikszentmihalyi (1983) in which participants are contacted (by mobile phone, pager, text message, etc.) at random time points each day during the course of their daily activities. When contacted, they are asked to answer some questions of interest to the researcher, which might take the form of psychometric questionnaires, open questions, etc. One advantage of ESM is that participants are not aware of when they will next be contacted. Thus, they are likely be fairly immersed in their daily activities rather than thinking about the experiment, thereby minimizing demand characteristics or response biases. In the case of INMI research, ESM can be useful in capturing the spontaneous experience of inner music as it is occurring, and has been employed successfully in previous studies (Bailes, 2006, 2007a, 2015; Beaty, et al., 2013; Byron & Fowles, 2013; Floridou & Müllensiefen, 2015). However, one disadvantage of this method is that, as participants are contacted by the researcher a fixed amount of times per day, it is possible that many INMI episodes could be missed, as participants might not happen to be experiencing INMI at the specific times they are contacted.

Other diary-based methods have also been employed, in which participants record information about the experience each time they notice an INMI episode (Beaman & Williams, 2010; Halpern & Bartlett, 2011). This method has the potential to minimize the amount of missing data in the study as compared to ESM, as participants are asked to record all INMI episodes that they experience throughout the duration of the study. However, this method relies on the participant being able to remember to record his/her INMI on a regular basis over a period of days

or weeks without prompting from the experimenter. As such, it is important for researchers to choose an appropriate length of time for the experiment to run and to conduct the study during a time period in which participants are free to be fully engaged in the study.

The combination of these various techniques—laboratory studies, questionnaires, ESM, and diary methods—has led to a wealth of initial findings regarding the experience of spontaneous inner music, which will be described in the subsequent sections of this chapter. A summary of the various research methods used to study the INMI experience is presented in Table 1.3.

Table 1.3: *Methods Previously Used to Study INMI*

Method	Studies using this method	Description	Advantages	Disadvantages
Questionnaire	Beaman & Williams (2010, 2013); Beaty et al. (2013); Floridou et al. (2015); Halpern & Bartlett (2011); Hyman et al. (2013); Müllensiefen et al. (2014); Liikkanen (2008, 2012a); Williamson et al. (2012, 2014)	Participants fill in a questionnaire with fixed response choices ad/or open questions about the INMI experience	-Can reach a wide range of participants (e.g., via online survey platforms) -Can include questions on a diverse array of INMI features and evaluations	-Retrospective method, subject to memory biases and forgetting -Participants provide overall evaluations of the INMI experience, rather than on a single-episode basis
Diary	Beaman & Williams (2010); Halpern & Bartlett (2011)	Participants record INMI occurrences in a diary and answer questions related to the experience each time an INMI episode occurs	-Participants respond to questions about INMI as an episode is occurring, thus minimizing retrospective biases -Potential to collect many data points from a single participant	-Self-regulated by the participant, i.e., the experimenter doesn't generally remind the participant to fill in the diary during the study

Experience Sampling Method (ESM)	Bailes (2006, 2007a, 2015); Beaty et al. (2013); Byron & Fowles (2013); Floridou & Müllensiefen (2015)	Participants record INMI occurrences in a diary and answer questions related to the experience when probed by the experimenter (e.g., via text message, phone call, etc.)	-Experimenter probes serve as a reminder to fill in the diary -Participants are generally probed at random and, as such, won't be expecting a probe (helps to preserve the spontaneous nature of INMI) -Participants respond to questions about INMI as an episode is occurring, thus minimizing retrospective biases	-Some or many INMI episodes may not be caught by the probes -Participants who are late to respond to probes or forget to respond to probes may be problematic
Laboratory-based INMI triggering	Beaman, Powell, & Rapley (2015); Floridou et al. (2012); Floridou et al. (2014); Hyman et al. (2013); Liikkanen (2012b)	The experimenter attempts to trigger INMI episodes of a particular song in a laboratory setting (e.g., by playing the music loud, using cued-recall techniques, etc.)	-High degree of experimental control over the situational antecedents of INMI and concurrent tasks during INMI	-Not all participants will necessarily experience INMI -Prior familiarity and recency of exposure to a song can influence the likelihood that a song will become INMI

1.3.3 Features and Phenomenology of INMI

One notable feature of INMI is its ubiquity. Liikkanen (2012a) conducted an Internet survey of over 12,000 participants and found that 89.2% of respondents experienced INMI at least once per week, while 33% of the survey respondents reported that they experienced INMI daily. Liikkanen et al. (in press) collated over 80,000 reports of INMI experiences on Twitter and found that the data represented users from 173 geographic regions, suggesting that INMI is a highly global phenomenon. In a survey of 205 university students, Kvavilashvili and Mandler (2004) found that INMI was the most frequently experienced type of involuntary semantic memory, as compared to other involuntary semantic memories such as words, names, phrases, or images. Bailes (2006) conducted an ESM study in which musician participants reported that they were currently hearing music on 44% of their completed ESM booklet forms and that they were currently imagining music on 32% of the forms. This suggests that, at least for musicians, imagining music is an activity that is nearly as frequent in daily life as actual music listening. It should be noted, however, that this particular study did not discriminate between INMI and VMI and, as such, the percentages of time spent experiencing each of these types of imagery cannot be discerned.

Another key feature of INMI is that it often comprises a repeated fragment of a song, rather than a mental replay of an entire piece of music. This feature was noted by Brown (2006), in a case study of his own INMI experiences, and by Liikkanen (2012a) in his survey of over 12,000 respondents. In a survey by Hyman et al. (2013), over 70% of respondents indicated that their current or most recent INMI experience comprised one or two lines of a song. The reason behind this repetitive looping of the same fragment of musical material within INMI is still not well understood. It may be related to limitations on the amount of musical material that can be rehearsed within working memory, or to the fact that

a certain section of a song might be better remembered than other sections. This idea that certain song sections might be remembered better than others is supported by the finding that the chorus of a song tends to be more often reported as INMI than other song sections (Beaman & Williams, 2010; Williamson & Müllensiefen, 2012). It is likely that the chorus of a song would be highly memorable, as this section is often repeated verbatim multiple times during the course of a song. Related research on song "hooks" has also revealed that certain sections within a single song can be more salient in terms of long-term memory recall than other sections (Burgoyne, Bountouridis, Van Balen, & Honing, 2013).

The duration of INMI episodes can vary quite widely. In two diary studies, the average reported durations of INMI episodes were 27.25 minutes (Beaman & Williams, 2010) and 36 minutes (Halpern & Bartlett, 2011). Some participants reported INMI episodes as short as 2 minutes, while others reported episodes lasting an entire day (Halpern & Bartlett, 2011). One challenge of attempting to measure INMI duration is that defining an objective start and end point of an episode can be difficult for participants, as INMI often tends to fade in and out of consciousness depending on the concurrent task and amount of attention paid to the experience. However, the results of these studies provide at least some insight into how INMI duration can vary between episodes and individual participants. The average length of INMI episodes has also been found to relate to certain personality traits and individual differences in mental control (see Section 1.3.7 for more detail on this point).

Regarding the types of songs that are experienced as INMI, songs that are familiar and liked by the experiencer are much more likely to become INMI than unfamiliar or disliked songs (Halpern & Bartlett, 2011; Hyman et al., 2013). Reports of entirely novel music as INMI are rare but not unheard of (Liikkanen, 2008). Songs from a wide variety of genres have

been reported as INMI, including pop, rock, folk, jazz, classical, country, gospel, holiday music, children's songs, TV jingles, and film scores (Beaman & Williams, 2010; Halpern & Bartlett, 2011). Previous studies have also revealed little overlap in INMI tunes between participants, such that the vast majority of songs reported as INMI were named by only one participant in questionnaire and diary studies (Beaman & Williams, 2010; Williamson & Müllensiefen, 2012). Thus, INMI appears to be a very personal and idiosyncratic experience, which likely reflects the diverse listening habits present in today's society. However, it is not inconceivable that there are certain inherent musical features of these at first seemingly diverse set of INMI tunes that increase their likelihood of being spontaneously recalled and repeated within the mind. Study 5 of this thesis aims to address this issue in further detail by investigating the influence of both extra- and intra-musical features of a tune on the likelihood that it will become INMI.

Finally, multiple researchers have observed commonalties between the INMI experience and the perceptual experience of music. Brown (2006) kept a detailed report of his own constant INMI, which he referred to as a "perpetual music track". He noted that his INMI tended to be a quite faithful reproduction of the perceptual experience of music in terms of pitch, tempo, and timbre. Williamson and Jilka (2013) conducted intensive interviews regarding the INMI experience using six participants from varying musical backgrounds. These participants reported that their INMI often represented a close reproduction of the original song in terms of the instrumentation and vocals. However, the participants also reported that the complexity and vividness of the INMI experience could change over time. The evidence reported by both Brown (2006) and Williamson and Jilka (2013) that INMI can be a highly faithful representation of a perceptual experience parallels findings on VMI research, as VMI also tends to represent several aspects of the original music (e.g., pitch, tempo, timbre) with high fidelity (Crowder, 1989; Halpern 1988b, 1989; Levitin, 1994; Levitin & Cook, 1996). One limitation of this evidence on

the INMI experience is that researchers have relied primarily on self-report data from participants. To date, no research has systematically probed the musical content of the INMI experience, for instance by objectively comparing the pitch or tempo of each INMI episode to the original version of a song. Study 4 of this thesis aims to address this question in more detail, specifically in regard to the representation of tempo within INMI.

1.3.4 Situational Antecedents of INMI

Williamson et al. (2012) explored the environmental and endogenous cues that could trigger an INMI experience. The most common INMI trigger was recent and/or repeated exposure to music. Associations with people, sounds, words, or other perceptual triggers from one's environment and certain emotional states were other common triggers. The finding that a wide variety of cues can elicit an INMI episode provides some evidence as to why the experience of INMI may be so common in everyday life, as well as why such a diverse array of different songs can become INMI.

It has been found that INMI appears frequently under periods of low attentional demand (Williamson et al., 2012), during such activities as walking, doing chores, and driving (Hyman et al., 2013). Hyman et al. (2013) also reported a higher incidence of INMI during completion of puzzles requiring high cognitive load than during completion of less demanding puzzles. However, this finding was not replicated by Floridou et al. (2014), who reported more frequent INMI in participants engaging in low cognitive load activities but less frequent INMI during high cognitive load activities. It may be that in the work of Hyman et al. (2013), the high cognitive load activities were too challenging for participants such that they stopped attempting to complete them; thus, the activities could have actually become low in cognitive load.

Byron and Fowles (2013) also reported roles of familiarity and recency of exposure in increasing the likelihood that a tune would become INMI. Participants in their ESM study who had heard a previously unfamiliar song aloud six times were more likely to experience it as INMI than participants who had heard it twice. Additionally, participants were more likely to report experiencing a song as INMI when they had recently heard the song aloud than when they had not heard it recently. Thus, research has shown that a variety of factors, including recent exposure to a song, environmental and memory triggers, cognitive load, and familiarity, can be involved in the generation of an INMI experience.

1.3.5 Emotional Responses to INMI

On average, INMI episodes are more often rated as positive or neutral than negative in valence (Beaman & Williams, 2010; Halpern & Bartlett, 2011; Hyman et al., 2013). Beaman and Williams (2010) found that 67% of INMI episodes in a diary study were rated as pleasant or neutral in valence, while 89% of Halpern and Bartlett's (2011) questionnaire respondents rated their INMI as typically pleasant or neutral. These findings are in contrast to the commonly held conception that having an earworm "burrowed" in your mind is primarily an annoying and unwanted experience. This tendency to assume that INMI is a more annoying experience than is actually found in research may be due to a recollection bias to more vividly recall those occasional INMI episodes that are particularly troublesome (Beaman & Williams, 2010).

Despite the finding that INMI is a generally pleasant experience, respondents in Liikkanen's survey (2012a) who were queried about involuntary cognition in general rated INMI as more disturbing than involuntary visual, olfactory, and other semantic memories. This suggests that there are occasions in which INMI becomes distracting or

unpleasant, and that this may happen more often for INMI than other types of involuntary cognition. Some researchers have proposed a comparison between INMI and the phenomenon of intrusive thoughts (Hyman et al., 2013, 2015). An intrusive thought has been defined as "any distinct, identifiable cognitive event that is unwanted, unintended, and recurrent. It interrupts the flow of thought, interferes in task performance, is associated with negative affect, and is difficult to control" (Clark & Rhyno, 2005, p. 4). Although it is quite clear from the aforementioned research that INMI is not necessarily unwanted or associated with negative affect, it is a recurrent, unintended type of thought that can be difficult to control. Whether a song experienced as INMI is disliked and whether an episode might interfere with concurrent tasks are two factors that have been found to contribute to a feeling of intrusiveness during an INMI experience (Hyman et al., 2015). The repetitive nature of the INMI experience (the looping of short sections of music in the mind) might also contribute to the intrusiveness of the experience. This is supported by evidence that people who display traits related to repetitive thought patterns, such as neuroticism and sub-clinical obsessive-compulsive traits, tend to experience both longer and more disturbing INMI episodes (Floridou et al., 2012; Müllensiefen et al., 2014a; see Section 1.3.7).

1.3.6 Behavioural Responses to INMI

The literature cited above provides evidence that although INMI episodes are often appraised as positive or neutral in valence, there are also occasions in which INMI can be distracting, irritating, or unpleasant. When INMI episodes do become troublesome or worrying, many people engage in coping behaviours in attempts to eradicate the unwanted INMI. Williamson et al. (2014) revealed a wide range of INMI coping strategies in a qualitative study of responses to an open-ended survey question. Passive INMI coping strategies, e.g., letting the music fade away on its own, were the most commonly reported behavioural responses to INMI in

this study. However, various active coping strategies were also employed (see also Beaman & Williams, 2010, for additional reports of similar active coping strategies). The most commonly reported active strategies in Williamson et al.'s (2014) work included engaging with the INMI tune (by listening to the tune aloud, manipulating it in one's head, etc.) and distraction with a different song or verbal material, such as talk radio or having a conversation. Some participants even reported specific distractor songs that they believed could dispel INMI, e.g., "*play the lic from Layla by Eric Clapton in my head. It gets rid of whatever song I had there before but it doesn't actually stick in there*" (Williamson et al., 2014, Appendix S2). Overall, participants in the study reported that they found both active and passive strategies useful in expelling INMI.

A recent laboratory-based study tested one possible strategy for preventing the appearance of unwanted INMI. Beaman, Powell, and Rapley (2015) found that the number of times an INMI tune came to mind was reduced in a condition in which participants were asked to chew gum, as compared to various control conditions. Chewing gum has been used in previous research to interfere with sub-vocalization processes (Kozlov, Hughes, & Jones, 2012). Thus, the findings of Beaman et al. (2015) indicate a relationship between articulatory motor programming and the initiation of INMI episodes and suggest chewing gum as a potential INMI prevention strategy. These findings bear certain parallels to the work of Williamson et al. (2014) reviewed above, as verbal strategies such as singing a tune or having a conversation featured highly in list of INMI coping strategies, suggesting that making use of the articulators might also be effective in terminating unwanted INMI episodes once they have begun. However, further work is needed to test the INMI coping strategies of Williamson and colleagues in a more controlled setting to further evaluate their efficacy.

1.3.7 Individual Differences Related to the Occurrence of INMI

Researchers have also investigated the influence of demographic, personality, and musical background variables on INMI propensity and INMI evaluations. Liikkanen (2008) found a relationship between musical training and INMI incidence rates, such that participants with more musical training reported more frequent INMI. However, subsequent studies have indicated that musical engagement and the importance ascribed to music may be more key to predicting increased INMI incidence rates than formal musical training (Beaman & Williams, 2010; Müllensiefen et al., 2014a; Williamson & Müllensiefen, 2012). Liikkanen (2008) also reported a higher prevalence of INMI in women than men, but this finding was not upheld in other subsequent questionnaire studies (Beaman & Williams, 2010; Hyman et al., 2013).

Several links between INMI and personality traits have been reported. Positive relationships have been found between two personality constructs—openness to experience and neuroticism, as measured by the NEO Five Factor Inventory—and INMI frequency and the length of INMI episodes (Beaty et al., 2013; Floridou et al., 2012). Neuroticism has also been found to relate to several negative aspects of the INMI experience, including ratings of unpleasantness, interference, difficulty in dismissing INMI, and the feeling that a song is "stuck" in the mind (Beaty et al., 2013; Floridou et al, 2012). These findings suggest that, although a person who scores highly in openness to experience and a person who scores highly in neuroticism may both experience a greater frequency of INMI in their everyday lives, the person with neurotic traits will be more likely to appraise this as a negative experience. INMI propensity has also been positively associated with transliminality, a psychological construct that measures how penetrable people are to inner thoughts rising to consciousness (Wammes & Barušs, 2009).

Beaman and Williams (2013) investigated individual differences in thought suppression, as measured by the White Bear Suppression Inventory (WBSI), and psychosis proneness (schizotypy), as measured by the Schizotypal Personality Questionnaire (SPQ), and their potential relationships to INMI frequency and INMI reactions. Participants who reported greater difficulty in thought suppression and higher degrees of schizotypy exhibited INMI that were longer, more disruptive, and experienced more difficulty in dismissing INMI. Schizotypy was also positively related to worrying about INMI, but the thought suppression measure was not significantly related to INMI-related worry. Interestingly, neither measure of mental control was related to INMI frequency. This suggests that mental control-related traits might not play a direct role in the initiation of INMI, but do play a role in the appraisal and reactions to the experience once it has begun.

Finally, sub-clinical obsessive-compulsive (OC) traits have been found to be positively related to INMI frequency and disturbance (Müllensiefen et al., 2014a). Higher degrees of INMI disturbance were also associated with longer INMI episodes in this research. This is in line with previous literature suggesting that more effortful attempts to stop unwanted cognitions often result in a paradoxical increase in these types of thoughts (Wegner, Schneider, Carter, & White, 1987). Clinicians have also reported cases of "musical obsessions", which are essentially highly persistent and highly disturbing INMI, that are often co-morbid with obsessive-compulsive disorder (OCD) (Taylor et al., 2014).

1.3.8 Functions of INMI

The possible functions of INMI are still debated. Evidence for a recency effect—in which the most recently heard piece of music tends to become INMI more often than other music (Liikkanen, 2012b; Hyman et al., 2013)—suggests a link between INMI and mental rehearsal of recent material, which may aid in the consolidation and maintenance of musical

memories. However, empirical studies to date have not investigated whether, for instance, participants who experience more frequent INMI for a specific piece of music learn that piece faster or more accurately than those who experience less INMI.

Mood regulation is another possible function of INMI, although existing evidence for this hypothesis is also sparse. Williamson et al. (2012) found affective state to be a commonly reported trigger for INMI in qualitative research, indicating that certain moods might trigger certain types of INMI. Bailes (2012) reported that participants in an ESM study whose scores were at or higher than the midpoint on a bipolar scale from "alert" to "drowsy" reported less INMI, suggesting that a certain degree of alertness is associated with experiencing INMI. To date, however, no research has looked into whether certain mood constructs (e.g., happy, sad, aroused) relate specifically to certain features of INMI (e.g., mode, tempo, lyrics). Study 4 of this thesis begins to explore this highly under-investigated area of study, in regard to the relationships between specific musical features of the INMI experience and concurrent mood state.

1.3.9 The Neuroscience of INMI

To date, only one study has specifically investigated the brain basis of INMI. Farrugia, Jakubowski, Cusack, and Stewart (2015) used MRI to examine the brain structural correlates of INMI propensity and evaluations, as measured by a self-report questionnaire known as the Involuntary Musical Imagery Scale (IMIS; Floridou et al., 2015). INMI propensity was related to cortical thickness in the right Heschl's gyrus (HG) and right inferior frontal gyrus (IFG). Both of these areas are also implicated in auditory perception and VMI (Griffiths & Warren, 2002; Herholz et al., 2012; Zatorre et al., 1996). Affective evaluations of INMI, namely the extent to which participants wished to suppress INMI or considered them helpful, were related to grey matter volume in the right temporopolar (TP) and parahippocampal (PHC) cortices respectively. The

TP has previously been found to play a role in affective processing (Royet et al., 2000) and the PHC has been found to play a role in both voluntary and involuntary memory retrieval (Hall et al., 2014). Taken together, the findings from this study suggest that individual differences in INMI propensity and evaluations are directly related to brain areas implicated in auditory processing, affective evaluation, and memory retrieval. Although investigations of the neural correlates of INMI are still in their infancy, it is likely that the development of innovative behavioural paradigms will soon allow researchers to "capture an earworm" in the MRI scanner.

1.3.10 Summary of INMI Research

Research on the INMI experience has a shorter history than research on VMI, yet a recent surge of psychological interest in this topic has already revealed a number of robust findings. INMI is a common and widespread experience (Liikkanen, 2012a; Liikkanen et al., in press). A wide variety of musical genres can be represented within INMI (Beaman & Williams, 2010; Halpern & Bartlett, 2011), and INMI tunes are generally familiar to and liked by the experiencer (Halpern & Bartlett, 2011; Hyman et al., 2013). INMI can be triggered by recent exposure to music, but also by memory associations, affective states, and a variety of other internal and external cues (Williamson et al., 2012). Research also suggests that INMI occurs more frequently during low attentions states and tasks involving a low cognitive load than cognitively demanding tasks (Floridou et al., 2014; Hyman et al., 2013; Williamson et al., 2012). INMI tends to be rated more often as a positive or neutral than negative experience, despite the bias often present in common discourse that an "earworm" is an annoying experience that should be dispelled immediately (Beaman & Williams, 2010; Halpern & Bartlett, 2011; Hyman et al., 2013). When an INMI episode *is* evaluated as unpleasant, coping behaviours include both passive and active strategies, such as letting the INMI tune fade out on its own or playing a different tune aloud (Beaman & Williams, 2010; Williamson et al., 2014). Finally, a number of studies have investigated individual differences related to the INMI experience, and have found

relationships between INMI frequency and/or evaluations and musical training or engagement, personality traits (neuroticism, openness to experience), thought suppression, and obsessive-compulsive traits (Beaman & Williams, 2013; Beaty et al., 2013; Floridou et al., 2012; Müllensiefen, et al., 2014a; Wammes & Barušs, 2009).

1.4 Overview of the Thesis

The subsequent chapters of this thesis aim to empirically address several gaps and unanswered questions raised by the literature reviewed above. The first three studies reported in the thesis investigate the influence of several psychological factors on the representation of the tempo of familiar music within VMI. This research is crucial in providing a more thorough understanding as to why temporal elements of music have been found to be susceptible to distortion within VMI (Bailes, 2002; Janata & Paroo, 2006; Weir et al., 2015) and why some participants perform more accurately in imagined tempo tasks than others (Halpern, 1992; Janata & Paroo, 2006). The second two studies of the thesis extend techniques used previously in laboratory-based studies of VMI and musical memory to objectively examine specific features of the musical content of INMI. The findings of this research will help to further our understanding of how the musical content of INMI relates to the generation and functions of an INMI experience.

Chapter 2 presents the results of two related studies. The overarching research question addressed by these two studies is whether psychological factors that are known to affect time perception might also affect the tempo at which music is imagined. The first study examines the effect of physiological arousal on tempo of VMI and the second study examines the relationship between age and VMI tempo. Hypotheses are posited within the context of internal clock models of time perception and previous work on the effects of arousal and age on time-related tasks (Block, Zakay, & Hancock, 1998; Boltz, 1994; Gibbon, Church, & Meck,

1984; McAuley, Jones, Holub, Johnston, & Miller, 2006; Zakay & Block, 1996). The results of these two studies provide evidence as to whether factors known to influence time perception can also influence mental representations for complex temporal stimuli such as music that are drawn directly from long-term memory. The results of these studies are also used to inform the research presented in Chapter 4, which examines factors that affect temporal aspects of *involuntary* musical imagery.

The research reported in Chapter 3 represents both a methodological inquiry to compare different ways of measuring the tempo of VMI and an investigation into the veridicality of tempo representations within musical imagery. The results of this work provide insights into how methods previously employed in separate studies to measure the tempo of VMI (e.g., Halpern, 1988b, 1992) compare when administered to the same participants using the same stimuli across three conditions. The study also provides several novel insights into individual differences that affect the veridicality of tempo recall within VMI.

Chapter 4 examines temporal features of INMI using a novel method to track the tempo of individual INMI episodes during participants' daily lives. This method allows for an investigation into the precision with which involuntary musical memories are recalled, as well as an examination of how endogenous processes such as concurrent affective state relate to INMI tempo. The findings presented in Chapters 3 and 4 taken together also allow for some first comparisons between the temporal veridicality of voluntary and involuntary musical imagery.

Finally, Chapter 5 examines the melodic features of INMI using a computational modelling approach. This approach has been used in previous laboratory-based research on the memorability of tunes (Müllensiefen & Halpern, 2014), and represents a useful tool for investigating whether specific features of melodies contribute to the ease of mental replay, such that certain tunes become more easily "stuck" in the mind than others. A pre-existing survey database is used to compile a

dataset of frequently reported INMI tunes. These tunes are then compared in terms of their musical features to a control dataset of tunes that were 1) never reported as INMI and 2) matched to the INMI dataset in terms of their popularity and style. The results of this research reveal several musical features that predict the likelihood that a tune will become INMI.

Finally, the outcomes of the five studies are discussed in detail in Chapter 6, particularly in terms of the parallels and divergences between the two types of musical imagery and how the findings of the present research can help to inform theoretical accounts of musical memory and music cognition research in general.

Chapter 2. Investigating Psychological Factors that Influence Imagined Tempo: Effects of Arousal and Age

Abstract

The work presented in this chapter investigates whether two factors known to affect human time perception can also systematically influence tempo judgements for familiar, non-canonical songs recalled from long-term memory. Study 1 revealed a significant positive effect of a physiologically arousing task—jogging on the spot—on tempo choices for 14 familiar songs relative to a non-arousing control task (anagram completion). Tempo choices in both music perception and musical imagery conditions were similarly affected by the arousal manipulation. Study 2 investigated the influence of age on tempo judgements for familiar music in a cross-sectional design. No significant effect of age was found in a music perception condition, but an effect of age was found in a musical imagery condition in the opposite direction to that hypothesized. Results are discussed in relation to theories of time perception and implications for music performance.

2.1 General Background and Aims of the Research

Several authors have reported that tempo judgements for perceived and imagined music rely on a common mental representation (Halpern, 1988b; Weber & Brown, 1986) and that deliberate mental recall of tempo information can be both highly veridical and highly consistent (Halpern, 1992; Levitin & Cook, 1996). However, other studies have reported that tempo information is represented with less precision within musical imagery in comparison to pitch information (Bailes, 2002; Janata & Paroo, 2006; Weir et al., 2015). Previous research also indicates that some participants perform more accurately than others in imagining temporal aspects of a musical stimulus (Halpern, 1992; Janata & Paroo, 2006). These particular studies found evidence that participants with more

musical training were more consistent and precise in their mental reproductions of musical tempo within imagery. This could be due to the fact that musicians often use online musical imagery during practice and performances, in order to create an idealized mental representation before producing the music aloud (Keller & Appel, 2010; Keller, 2012). As such, this type of task is highly familiar and useful to musicians as compared to non-musicians. However, beyond the established relationship between musical training and performance on imagined tempo tasks, little additional evidence exists as to why imagined tempo representations are reported as highly precise in some studies and fairly malleable in others.

The mental representation of musical tempo requires one to imagine the dynamic unfolding of a sequence of events over a certain time course. This is dissimilar to several other aspects of musical imagery that have been studied in previous research, such as image generation for single pitches, timbres, or loudness levels (Crowder, 1989; Farah & Smith, 1983; Halpern et al., 2004; Intons-Peterson, 1980), which are essentially static in nature. Thus, one potentially informative approach for the present work is to investigate the influence of factors known to influence human time perception in general on mental representations of musical tempo. Previous psychological research has revealed a diverse array of factors that can influence time-related judgements, including arousal, age, certain clinical disorders, and psychoactive drugs (Boltz, 1994; Pastor, Artieda, Jahanshahi, & Obeso, 1992; Salthouse, 1996; Wittmann, Leland, Churan, & Paulus, 2007). The research in this chapter will focus specifically on two of these factors—physiological arousal, which is manipulated via exercise in Study 1, and age, which is examined in a cross-sectional design in Study 2—in order to discern whether such factors can systematically affect mental representations of the tempo of familiar music.

Study 1: The Effect of Physiological Arousal on Chosen Tempo for Perceived and Imagined Melodies

2.2 Introduction

2.2.1 The Effect of Arousal on Time Perception

A relationship between arousal and time perception has been posited at least since the introduction of Treisman's (1963) internal clock model of interval timing, which includes an arousal-sensitive pacemaker. Since then, several behavioural studies have confirmed that arousal can significantly affect human time-related judgements. These studies have manipulated arousal in various ways, including via stress (Boltz, 1994), physical activity (Schwarz, Winkler, & Sedlmeier, 2013), fear (Grommet et al., 2011), and fatigue (Miró, Cano, Espinosa-Fernández, & Buela-Casal, 2003). Across these studies, increased levels of arousal led to retrospective overestimations of the duration of visual or auditory stimuli (Grommet et al., 2011; Wearden, Pilkington, & Carter, 1999) and underestimations in time production tasks (also called prospective time judgement tasks; Miró et al., 2003; Ozel, Larue, & Dosseville, 2004). That is, in tasks that required a judgement of the duration of a stimulus after it occurred (retrospectively), participants tended to overestimate its duration, while in tasks in which one was asked to produce a certain duration of time (e.g., respond after 10 seconds has passed), participants tended to respond too early.

Findings within this area of research have often been explained in the context of internal clock models of interval timing. Treisman proposed a model of interval timing in 1963 that included a pacemaker mechanism that generates periodic pulses, which are stored in a counter. The pulses that are accumulated in the counter during a particular time period are compared to a store of reference durations in order to make a time-related response. Treisman's ideas were expanded by Gibbon, Church,

and Meck (1984) into a now widely-cited model, known as scalar expectancy theory (SET). Several further adaptations of these models also added an attention component, in which an attentional gate opens more fully between the pacemaker and the counter/accumulator as attention is allocated to a timing task (e.g., Zakay & Block, 1996). Thus, more pulses accumulate when time is seen as important and relevant, while some pulses may be lost if attentional resources are allocated elsewhere.

As mentioned above, previous research has revealed that increased arousal leads to overestimations in retrospective duration judgements and underestimations in time production tasks. Although these results may at first appear contradictory, both sets of results can be explained in terms of internal clock models. It has been posited that the rate of pulse emission within the pacemaker increases with heightened arousal (e.g., Boltz, 1994; Treisman, 1963; Wearden et al., 1999). An increase in the speed of the pacemaker thus leads to the collection of more pulses in a shorter duration of time within the accumulator. This causes retrospective time judgements to be overestimated, as more pulses accumulate during a specific time period than would under non-aroused conditions. In time production tasks (e.g., if a participant is asked to respond when 10 seconds have passed) arousal can lead to underestimations of time, as the accumulator's store will reach a particular number of pulses at a faster rate than during non-aroused states.

2.2.2 Previous Research on Arousal and Musical Tempo

Some evidence has accrued to suggest that arousal can also influence judgements of temporal aspects of musical and music-related stimuli. Iwanaga (1995) asked participants to adjust the speed of the song "It's a Small World" to their favourite tempo and found that participants' preferred tempi were close to speed of their heart rates. However, the use of only one song stimulus, rather than a range of songs with different original tempi, imposes some limitations regarding the generalizability of

these findings. North and Hargreaves (2000) found that participants preferred and listened longer to a faster tempo piece of music during an exercise condition as compared to a relaxation condition, suggesting that fast tempo music is deemed a more suitable fit to highly arousing situations than low arousal situations. Additionally, Boltz (1994) reported a significant increase in post-manipulation spontaneous tapping speed in participants assigned to a stress-inducing condition, and a significant decrease in tapping speed in a relaxation group, thereby suggesting that both increases and decreases in arousal can temporarily alter one's spontaneous motor tempo.

2.2.3 Overview of the Research

The primary aim of the present research was to investigate whether mental representations of the tempo of familiar songs recalled from long-term memory could be systematically altered via an increase in physiological arousal. This study diverges from the work of North and Hargreaves (2000), who asked participants to provide liking and appropriateness ratings of a musical stimulus within an arousing context, as participants in the present research were asked to indicate the tempo that felt internally "correct" for familiar songs both before and after an arousal manipulation. As previous research on the effect of arousal on time perception has frequently required duration judgements for experimenter-constructed, single events (e.g., Grommet et al., 2011; Schwarz, et al., 2013), the present study also offered a new approach by employing the types of complex temporal stimuli that are regularly encountered in everyday life: familiar songs. The task used in the study is fairly ecological, with parallels to the familiar situation of choosing the speed for a well-known song, such as "Happy Birthday" or "Jingle Bells". Finally, several of the previous studies reviewed above involved training phases for learning the duration of the experimental stimuli (e.g., Boltz, 1994), which can introduce confounds due to individual differences in the ability to learn and retain previously unfamiliar stimuli in memory. The

present work eliminated the need for a training phase by utilizing stimuli that were already highly familiar to participants.

In the present study, participants were asked to choose the tempo that "sounded right" for familiar, non-canonical songs such as "Happy Birthday". This task was completed in both perceived and imagined music conditions; findings of similarities between performance in these two conditions would provide further support for existing claims regarding the implication of shared mental representations in perceived and imagined music tasks (e.g., Halpern, 1988a, 1988b; Weber & Brown, 1986). Participants were assigned to one of two groups. In one group, physiological arousal was induced via aerobic exercise (jogging on the spot). A second group, which served as a control group, completed a mentally—rather than physically—demanding task (anagrams) that was not designed to increase physiological arousal. Participants' choices of tempi for 14 familiar songs were measured before and after exercise or anagrams in both the perceived and imagined music conditions. To investigate the potential influence of individual differences in musical training and self-reported auditory imagery abilities on chosen tempi in these tasks, questionnaires were administered to assess these factors. The main hypotheses of the study were as follows:

- Hypothesis 1: An increase in physiological arousal would increase the exercise group's chosen tempi for familiar songs relative to the anagram group, for whom no significant post-manipulation change in chosen tempi was predicted.

- Hypothesis 2: Similar chosen tempo increases will occur in perceived and imagined music conditions following exercise.

If the results of the present study were found to support these hypotheses, this would provide robust evidence for the influence of arousal on the speed of the internal clock, by demonstrating that even temporal representations of a complex and highly familiar stimulus such as music can be altered via changes in physiological arousal.

2.3 Method

2.3.1 Design

A 2x2 mixed design was employed. A between-subjects factor of manipulation (exercise or anagrams) aimed to increase physiological arousal via increase in heart rate in the exercise group and to maintain a constant state of physiological arousal in the anagram group. A within-subjects factor of task type (music perception or musical imagery) was employed in order to compare perceived and imagined musical tempo choices across all participants. The dependent variable of interest was the degree of change between pre- and post-manipulation tempo ratings in both the perception and imagery tasks.

2.3.2 Participants

A total of 39 participants (15 male) took part, ranging in age from 19 to 34 years ($M = 23$, $SD = 4.1$). The age range was intentionally constrained to only 15 years, as age is another a factor that has been shown to affect time-related judgements in previous research (McAuley et al., 2006; Vanneste, Pouthas, & Wearden, 2001; see Study 2). Twenty of the 39 participants were assigned to the exercise group and 19 were assigned to the anagram group.

Before signing up for the experiment, participants were asked to confirm that they were familiar with all 14 songs used as stimuli and that they were able to participate in 8 minutes of moderate aerobic exercise. Participants assigned to the anagram group were required to confirm their ability to complete 8 minutes of exercise as well as those in the exercise group, in order to prevent the introduction of systematic differences between the two groups related to physical fitness level.

Assignment of participants to the exercise and anagram groups was conducted by the experimenter to ensure that the two group did not differ

significantly in age, $t(37) = -1.70$, $p = .10$, or musical training, as measured by the Gold-MSI (see Section 2.3.4 for a description of this questionnaire), $t(37) = 0.042$, $p = .97$.

All participants received either course credit or payment of £10 for their participation.

2.3.3 Ethics Statement

The research was approved by the Ethics Committee of Goldsmiths, University of London. All participants gave written informed consent before participating.

2.3.4 Materials

Musical stimuli for both the perceived and imagined music conditions consisted of 14 well-known songs (see Table 2.1), based on those used in the work of Halpern (1988b). The selection criteria for the song stimuli were that they should be highly familiar to participants living in the UK and that they should not exist in a single canonical, recorded version. That is, the songs used in the study, such as "Happy Birthday" and "Jingle Bells", exist in many different versions and do not have a definitive recording that is commonly associated with these songs. Non-canonical songs were chosen as stimuli for the present study in order to allow for comparisons to the work of Halpern (1988b), who investigated perceived and imagined tempo recall for a similar set of non-canonical songs.

Once the 14 songs had been selected for inclusion in the stimulus set, the first musical phrase of each song was extracted for use and these excerpts were rendered in MIDI piano timbre. The tempo at which each song was initially produced was determined by averaging the chosen tempi of four independent judges in a pilot study (each judge selected the tempo that sounded "right" to him or her for each particular stimulus).

Four paper booklets were constructed for presenting the song titles and lyrics for each musical stimulus. The 14 songs were presented in four different random orders in each of these booklets. The order in which participants received each of the four booklets in the four tasks that they completed during the study was counterbalanced. On each page of the booklet, the participant saw the title of a familiar song and the lyrics to the section of the song that they would either hear aloud or be asked to imagine. Syllables in the lyrics corresponding to beats in the music were marked in bold and underlined, in order to aid participants in finding the beat and to ensure that all participants were feeling the beat at the same metrical level. Any beat that occurred between syllables of the lyrics was marked as "____". An example of the booklet contents is provided in Figure 2.1.

Table 2.1: *Songs Used as Stimuli in Perception and Imagery Tasks*

Song	Recorded tempo (bpm)	Fast start tempo for perception task (bpm)	Slow start tempo for perception task (bpm)
Deck the Hall	69	120.75	34.5
Happy Birthday	128	224	64
Hark the Herald Angels Sing	56	98	28
Jingle Bells	100	175	50
London Bridge	81.5	142.625	40.75
Old MacDonald Had a Farm	75	131.25	37.5
Row, Row, Row Your Boat	104	182	52
Rudolph the Red-Nosed Reindeer	75	131.25	37.5
Santa Claus is Coming to Town	132	231	66
Silent Night	112	196	56
Three Blind Mice	112	196	56
Twinkle, Twinkle Little Star	75	131.25	37.5
When the Saints Go Marching In	112	196	56
White Christmas	75	131.25	37.5

Note: Fast start tempi are 175% of the speed of the original, recorded tempo and slow start tempi are 50% of the original, recorded tempo. The click track for the imagery task was started at a tempo of either 48 bpm or 204 bpm. For any individual participant, a given song was always presented at either a slow start speed or a fast start speed for all 4 blocks. This was done to account for possible response biases toward the initial presented tempo of the song or click track, which could potentially minimize effects owing to the experimental manipulation.

LONDON BRIDGE

London **Bridge** is **fall**ing **down**, **fall**ing **down**, **fall**ing **down**

Figure 2.1. Example song title and lyrics from the booklets provided to participants in each task, with beat structure marked in bold and underlined.

In order for participants to adjust the speed of the stimuli in the study in real time, a Max/MSP application was utilized. This application employed the time stretching technology known as ZTX (formerly DIRAC) (http://www.zynaptiq.com/ztx/), which is widely used in professional audio applications and the music industry to apply tempo changes without changing pitch or loudness. The application allowed a participant to increase or decrease the speed of a stimulus as it was being heard via a Griffin Powermate assignable USB controller. The Powermate is a circular dial that can make continuous adjustments in very small increments but that has no external calibration marks. This controller thus reduces the possibility of remembering the position at which the last trial was set, which helps to ensure independence of observations from trial to trial. Participants turned the Powermate to adjust the speed of the actual music in the perception condition and to adjust the speed of a click track to the beat of their imagined music in the imagery condition.

For the between-subjects manipulation (exercise or anagrams), heart rate (HR) measurements were taken as an indicator of physiological arousal. An Ultrasport Pulsewatch Run50 heart rate monitor with chest strap was used to measure participants' heart rates in both the exercise and anagram groups. Heart rate measurements were taken at the beginning and end of each jogging or anagram period. For the anagram group, eight

lists of anagrams were compiled; four of these comprised seven scrambled names of countries and the other four comprised seven scrambled animal names. These lists were piloted to ensure participants would not complete them in less than the 60 s allotted in the study for each anagram task.

Finally, individual differences in music-related abilities were examined. The Musical Training dimension of the Goldsmiths Musical Sophistication Index (Gold-MSI; Müllensiefen, Gingras, Musil, & Stewart, 2014; see Appendix 1) was administered to measure formal musical training. This scale includes questions relating to prior music lessons and practice, music theory training, and performance experience. The Bucknell Auditory Imagery Scale (BAIS; Halpern, 2015; see Appendix 3) was administered to measure self-reported auditory imagery abilities. The BAIS measures participants' auditory imagery abilities for musical, speech, and environmental sounds in terms of how vividly they are able to voluntarily imagine these sounds (Vividness subscale) and how easily they are able to control and change these imagined sounds (Control subscale).

2.3.5 Procedure

Each participant in the study completed four blocks of 14 trials each (see Figure 2.2). All participants completed the perception tasks in the first two blocks and then proceeded to the imagery tasks in the third and fourth blocks.

BLOCK 1	BLOCK 2	BLOCK 3	BLOCK 4
14 trials perception task	-60 s jogging OR anagrams -3 trials perception task -60 s jogging OR anagrams -4 trials perception task -60 s jogging OR anagrams -3 trials perception task -60 s jogging OR anagrams -4 trials perception task	14 trials imagery task	-60 s jogging OR anagrams -3 trials imagery task -60 s jogging OR anagrams -4 trials imagery task -60 s jogging OR anagrams -3 trials imagery task -60 s jogging OR anagrams -4 trials imagery task

Figure 2.2. Procedural layout for the tempo tasks.

For each trial of the perception task, participants saw the printed name and first line of lyrics to one of the 14 songs (see Figure 2.1). They heard the section of music corresponding to the lyrics that were presented through Sennheiser HD 202 headphones and were asked to "please adjust the knob until the speed of the music sounds right to you". They were told to turn the dial to the right if they wanted to speed the music up and to the left if they wanted to slow it down. Two example trials were provided before the start of this block to ensure participants understood the task. The start tempo for each trial was set to either 50% or 175% of the speed of the original recorded sound file, in order to ensure that participants would make tempo adjustments for all or most of the trials (see Table 2.1 for a list of recorded and start tempi).

In the second block, participants completed four 60 s periods of the between-subjects manipulation (either exercise or anagrams), each of which was interspersed with the perception task (see Figure 2.2). The manipulation was administered in this interspersed format in an effort to deter heart rate from returning too quickly to baseline in the exercise group throughout the 14 tempo adjustment trials. During the manipulation

periods, the exercise participants jogged on the spot, whereas the anagram participants completed as many anagrams from one of the eight lists of anagrams as they were able during the 60 s period. [2] The experimenter watched the heart rate monitor during each jogging period to ensure the exercise group participants were within the range of 50 to 70% of their maximum heart rate (maximum HR was calculated as 220-age). This range was based on procedures from previous behavioural studies utilizing aerobic exercise in a similar fashion (Murray & Russoniello, 2012; Pontifex & Hillman, 2007). The experimenter asked participants to increase or decrease their jogging speed if needed, within their ability. All participants were able to exercise within this target range for all 60 s periods. Apart from the 60 s periods of the manipulation and the 14 songs being presented in a different random order, the perception task was administered in the second block in the exact same fashion with the same instructions as in the first block.

Following the second block, participants filled in the Gold-MSI and BAIS questionnaires. This phase ensured that the exercise participants' heart rates returned to baseline before beginning the third block. There was no significant difference between exercise participants' baseline heart rates prior to commencing the experiment and their heart rates following the questionnaire period, t (19) = 1.36, p = .19.

In the third block, participants completed an imagery task. This task was analogous to the perception task from Block 1, with the only difference being that participants were asked to indicate the tempi at which the same set of 14 songs sounded correct while *imagining* (rather than hearing) them. Participants heard a click track through the Sennheiser HD 202 headphones and were asked to align the clicks to the beat of the imagined music. They were told to turn the dial to the right if they wanted to speed the click track up and to the left if they wanted to slow it down.

---

[2] The eight lists of anagrams were presented in a randomized order to each participant.

Two example trials were provided before the start of this block to ensure participants understood the task.

The fourth block combined the manipulation with the imagery task in the same interspersed format as employed for the perception task in Block 2. The experimenter again watched the heart rate monitor to ensure exercise group participants were exercising at 50 to 70% of their maximum heart rate and the anagram participants completed different sets of anagrams in each of the four 60 s manipulation periods. All exercise group participants were again able to exercise within the target heart range for the 60 s periods.

Following Block 4, participants were debriefed as to the purposes of the study. Each testing session lasted approximately 75 minutes.

2.4 Results

2.4.1 Group Analyses

As detailed in the Method section, participants in the exercise and anagram groups were deliberately matched on age and Gold-MSI Musical Training scores. In addition, no significant differences were found between the two groups in terms of self-reported auditory imagery abilities, as measured by the BAIS Vividness, $t(37) = -0.007$, $p = .99$, and BAIS Control scales, $t(37) = -1.40$, $p = .17$.

2.4.2 Arousal Manipulation Check

A manipulation check was conducted to ensure that the exercise group participants showed a significant increase in physiological arousal following the manipulation in comparison to the anagram group. The mean post-manipulation heart rate for the exercise group was 132.72 beats per minute (bpm) in Block 2 and 130.18 bpm in Block 4; for the anagram group these values were 76.98 and 78.98 bpm respectively. A

significant manipulation by heart rate interaction was found between pre- and post-manipulation heart rates in both the perception and imagery tasks (perception: $F(1,37) = 251.03$, $p < .001$, partial $\eta^2 = .84$, imagery: $F(1,37) = 311.88$, $p < .001$, partial $\eta^2 = .89$). This interaction indicates that the exercise group showed a significant increase in heart rate following jogging relative to the post-anagram task heart rate levels of the anagram group.

2.4.3 Tempo Task Analyses

Ratios were calculated to express any change in participants' chosen tempi after versus before the manipulation of exercise or anagrams. This was done by dividing the post-manipulation chosen tempo by the pre-manipulation chosen tempo for the same song for each participant. The ratios calculated for each song were then averaged over all 14 songs to provide a measure of mean post-manipulation tempo change for each participant in each task (perception and imagery). The ratios averaged across all participants for the perception and imagery tasks for each group are presented in Figure 2.3. A ratio of greater than 1 indicates an increase in chosen tempo following the manipulation, while a ratio of 1 indicates no change between pre- and post-manipulation chosen tempo. The mean ratios for the exercise group were greater than 1 in both the perception and imagery tasks (perception: 1.03, 95% CI [1.00, 1.06]; imagery: 1.06, 95% CI [1.04, 1.09]), indicating an overall increase in chosen tempi following the exercise manipulation. The 95% confidence intervals for the mean ratios of the anagram group straddled 1 in both tasks (perception: 1.00, 95% CI [0.99, 1.02], imagery: 1.03, 95% CI [0.99, 1.07]), indicating that chosen tempi ratios did not change significantly after the anagram manipulation. These tempo change ratios will be used as the dependent variable of interest in the following analyses and will be referred to as "degree of tempo change".

A 2x2 mixed ANOVA revealed a significant effect of manipulation type (exercise vs. anagrams) on degree of tempo change, $F(1, 37) = 5.43$, $p =$

.03, partial η^2 = .13. The directionality of this result indicates that the degree of post-manipulation tempo change was significantly higher for the exercise group than the anagram group. No significant interaction was found between manipulation type and task type (perception vs. imagery), $F(1, 37)$ = 0.07, p = .79, indicating that the performance of the two groups in comparison to one another did not significantly differ between the perception and imagery tasks.

If the above analysis is performed again with age included as covariate of no interest (to control for any potentially confounding effects of age), the effect of manipulation type (exercise vs. anagrams) on degree of tempo change still holds, although the effect size decreases slightly, $F(1, 37)$ = 4.03, p = .05, partial η^2 = .10.
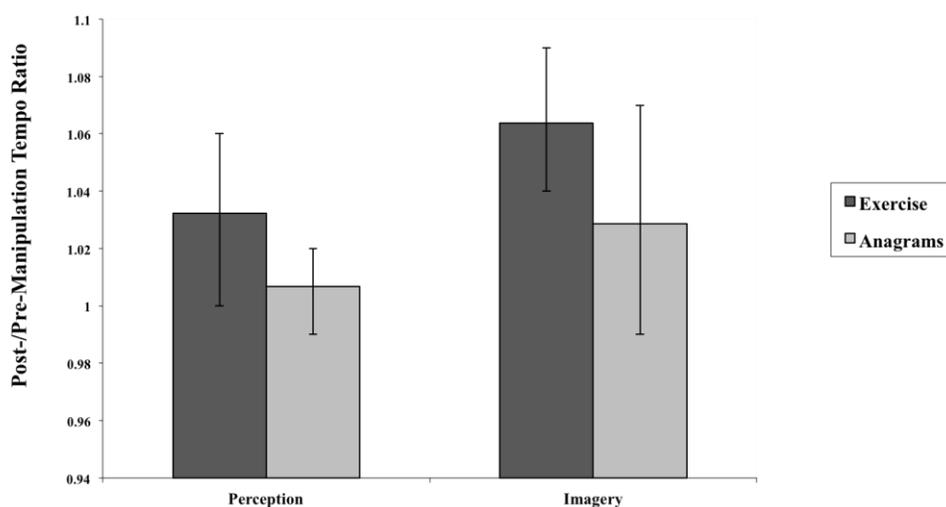
Figure 2.3. Mean ratios of post- to pre-manipulation chosen tempi for perception and imagery tasks (bars represent 95% confidence intervals).

2.4.4 Individual Differences

Correlations between the degree of tempo change ratios and several other experimental measures (age, Gold-MSI scores, BAIS scores, ratio

of post-exercise HR to baseline HR[3]) were calculated for each group separately. Significant correlations were found between some of the measures of music-related ability (exercise group: BAIS Control and BAIS Vividness, $r(18)= .80$, $p < .001$; anagram group: BAIS Control and BAIS Vividness, $r(17)= .58$, $p = .01$, Gold-MSI and BAIS Vividness, $r(17)= .57$, $p = .01$, Gold-MSI and BAIS Control, $r(17)= .54$, $p = .02$), but no significant correlations were found in relation to the tempo change variables.[4]

Finally, multiple regression analyses were conducted separately for the two groups (exercise and anagram groups) with age, Gold-MSI scores, BAIS scores, and degree of HR change as independent variables. No significant effects of these factors were found on chosen tempi (all ps > .05). Table 2.2 presents the results of these regression analyses; for brevity only results of the exercise group are presented here, since this was the main experimental manipulation of interest to the present research (anagrams served as a control task).

---

[3] This variable was included as a measure of degree of HR change, as some participants exerted more effort than others during the exercise task.

[4] An initial significant positive correlation was found between age and the imagery task tempo change ratio in the exercise group, $r = .64$, $p = .002$. However, this correlation was driven by one outlier in the distribution; when this participant is removed the correlation becomes non-significant, $r = - .35$, $p = .14$.

Table 2.2: *Multiple Regression Analyses for the Exercise Group*

1) Perception Task:

	Unstandardized B	S.E.	*t*-value	*p*-value
Constant	1.04	0.17	6.32	< .001
HR Change	0.05	0.09	0.49	.64
Age	-0.004	0.008	-0.48	.64
Gold-MSI	-0.001	0.001	-0.718	.49
BAIS Vividness	-0.02	0.03	-0.80	.44
BAIS Control	0.03	0.03	1.03	.33

2) Imagery Task:

	Unstandardized B	S.E.	*t*-value	*p*-value
Constant	1.08	0.13	8.21	< .001
HR Change	-0.001	0.08	-0.01	.99
Age	0.001	0.006	0.10	.93
Gold-MSI	< 0.001	0.001	-0.12	.91
BAIS Vividness	-0.03	0.02	-1.34	.21
BAIS Control	0.03	0.02	1.31	.22

2.5 Discussion

The present study tested two main hypotheses:

- Hypothesis 1: An increase in physiological arousal would increase the exercise group's chosen tempi for familiar songs relative to the anagram group, for whom no significant post-manipulation change in chosen tempi was predicted.

- Hypothesis 2: Similar chosen tempo increases will occur in perceived and imagined music conditions following exercise.

The results of the present study support Hypothesis 1 by revealing a significant effect of manipulation type on degree of post-manipulation change in chosen tempo, such that the exercise group showed significant increases in post-manipulation tempo choices relative to the anagram group. These findings provide evidence that the fairly stable mental representations for musical tempo that have been reported by several previous researchers (Halpern, 1988b, 1992; Levitin & Cook, 1996) can be systematically altered via a change in physiological arousal.

Support for Hypothesis 2 was also found. The absence of an interaction effect between manipulation type and task type indicates that the two groups performed similarly in relation to one another across the two tasks (perception and imagery). This result suggests that physiological arousal can affect one's mental representations of tempo not only when one is imagining a song, but even when the actual song is playing, which provides strong perceptual cues. This finding also provides support for previous research suggesting strong parallels between the mental processes involved in music perception and deliberate musical imagery (e.g., Halpern, 1988b; Zatorre & Halpern, 2005).

Previous research has indicated that arousal can affect time judgements for single-duration, experimenter-constructed stimuli (e.g., Grommet et

al., 2011; Schwarz et al., 2013), appropriateness ratings for music of different tempi (North & Hargreaves, 2000), and spontaneous tapping speed (Boltz, 1994). The present results extend this area of research by revealing that physiological arousal can also influence judgements of the tempo that sounds "right" for highly familiar musical melodies that have long-term representations in memory.

2.5.1 Relation to Internal Clock Models

The results of the present study can be interpreted in terms of the principles put forth in internal clock models of time perception (Gibbon et al., 1984; Treisman, 1963). The tempo choice tasks employed in this study require participants to make temporal judgements about a stimulus as it is occurring (or as it is being imagined) in real time; thus, this task can be considered a prospective time judgement task (rather than a retrospective time judgement task). The arousal manipulation employed in the present work serves to increase pacemaker speed in the exercise group, whereas pacemaker speed should remain relatively constant in the anagram group. Faster tempo choices are thereby made by the exercise group in order to compensate for the increased speed of pacemaker pulse emission. For example, consider a fictitious participant who allows 20 pulses of his/her pacemaker to accumulate during the course of the first phrase of "London Bridge" under normal, non-aroused conditions. This participant will need to set "London Bridge" to a faster than usual tempo during the exercise condition in order for it to occur over the course of 20 pulses, to compensate for the fact that 20 pulses will accumulate in a shorter time period than usual during an aroused state.

2.5.2 Real-World Implications

The findings of this research have various implications for music performance. Performing music is a mentally and physically demanding task that can be accompanied by a great deal of anxiety, which may lead to substantial increases in heart rate and subjective arousal prior to and

during a performance (LeBlanc, Jin, Obert, & Siivola, 1997). The present results suggest that even a moderate amount of physiological arousal (4 minutes of aerobic exercise) can increase chosen tempi for highly familiar music. Additionally, no significant relationship was found between musical training and the degree of post-manipulation tempo change in the exercise group, suggesting that the tempo choices of participants with more formal training in music were similarly affected by an increase in arousal to those of participants with little or no training. It should be noted, however, that participants in the present sample were not specifically selected on the basis of musical training, although a small handful were trained to a professional or semi-professional level. A future study involving a group comparison between professional musicians and non-musicians might provide further insight into how musicians cope with the increased arousal levels that often accompany a performance in order to maintain a stable performance tempo.

2.5.3 Limitations and Future Extensions of the Research

One limitation of the present study is the possibility that some participants may have inferred the experimental hypothesis, although none spontaneously reported this. However, it seems unlikely that demand characteristics, if present, could explain the present results due to the following reasons: 1) songs were always presented in a different order in different blocks, 2) the Powermate dial has no markings that could be used to remember tempo settings across blocks, and 3) the time delay between blocks makes it very difficult to retain prior tempo settings in memory.

The present study found no influence of self-reported auditory imagery abilities (BAIS Vividness and Control scores) on degree of post-manipulation tempo change in either the exercise or anagram group. This result suggests that participants who self-reported as being more able to vividly imagine sounds or more able to control their auditory imagery were not more immune to the impact of arousal on chosen tempi. In future

research in this area, a behavioural task could be employed to quantitatively measure auditory imagery abilities as a potentially more sensitive measure than a self-report scale, in order to further elucidate any differences in task performance due to pre-existing auditory imagery abilities.

Additionally, the degree of change in heart rate experienced by the exercise group participants had no significant influence on the amount of post-manipulation tempo change following the exercise periods. That is, participants who showed greater increases in heart rate did not systematically show greater increases in post-manipulation chosen tempi. However, the range for this heart rate change variable is fairly constrained, as most participants exercised at a moderate level and did not exert themselves beyond the minimum requirements of the study. A between-groups design in which multiple groups exercised at different levels of exertion might further elucidate any potential effects of this variable in the future. Finally, future research could aim to manipulate arousal in the opposite direction to the present study, e.g., decrease arousal through relaxation exercises. Such a study could provide further insights into the extent to which musical tempo choices can be manipulated by arousal and could also be useful in informing musical performance anxiety management strategies in order to achieve optimal tempo choices within performance situations.

2.5.4 Conclusion

In conclusion, the present study demonstrates that physiological arousal can influence temporal judgements for even complex and highly familiar stimuli such as well-known music. An effect of arousal on tempo choices for familiar music was present in both perceived and imagined music conditions. These results provide a rationale for examining the influence of other psychological factors known to influence time perception (e.g., age) and group differences (e.g., musicians vs. non-musicians) to provide

a wider range of evidence as to how the speed of the internal clock can influence temporal judgements within the music domain.

Study 2: Investigating the Relationship Between Age and Chosen Tempo for Perceived and Imagined Melodies

2.6 Introduction

Study 1 of this chapter represents a contribution to the literature on the effects of arousal on time-related judgements, particularly in relation to chosen tempo for familiar music. Study 2 will take a similar approach in order to investigate whether another factor known to influence time perception can systematically influence the tempo at which a piece of music sounds "right".

Several previous studies have revealed effects of age on time-related processes (e.g., Block et al., 1998; McAuley et al., 2006). The direction of these age effects is somewhat mixed, which is likely due to the variety of different time estimation paradigms used in previous literature (see Block et al., 1998 for a review and commentary on this point). Several researchers have reported that retrospective time judgements tend to be underestimated and prospective judgements tend to be overestimated more significantly in older adults than in younger adults (Craik & Hay, 1991; Perbal, Droit-Volet, Isingrini, & Pouthas, 2002) and that spontaneous tapping speed slows with age (Vanneste et al., 2001). These particular findings have often been explained in terms of a slowing of the internal pacemaker implicated in interval timing models (Craik & Hay, 1991; Droit-Volet & Wearden, 2002). However, several researchers have commented that a slowing of the pacemaker, in and of itself, is not sufficient for explaining all age-related findings in the timing literature (e.g., Block et al., 1998). A decline in working memory capacity

(Baudouin, Vanneste, Isingrini, & Pouthas, 2006) and slowing of processing speed (Perbal et al., 2002; Salthouse, 1996) with age have more recently been cited as additional contributing factors to age-related differences in time estimation. The importance placed on these additional factors may help to explain the fact that some previous research has revealed effects in the opposite direction to the references cited above, such that prospective *underestimations* of time increase with age (Block et al., 1998; Espinosa-Fernández, Miró, Cano, & Buela-Casal, 2003). The present work attempted to extend this literature and to clarify some of these mixed findings by utilizing ecological stimuli with which participants were previously familiar and could recall from long-term memory (familiar songs), in order to minimize the constraints imposed by needing to learn a stimulus's duration during the experimental session or to hold a novel stimulus within working memory.

2.6.1 Effects of Age on Music-Related Timing

Several researchers have found evidence for effects of age on spontaneous tapping rates and the perception of music-related stimuli. Vanneste et al. (2001) compared a group of young (ages 20 to 30) and older (ages 60 to 76) adults. The average spontaneous tapping rate was slower overall in older adults across five sessions, but no group difference was seen in terms of relative variability (measured as the coefficient of variation of tapping). The authors concluded that this finding supports the idea that the pacemaker of the internal clock becomes slower with age but not more variable. Drake, Jones, and Baruch (2000) also reported a slowing of spontaneous tapping rates with age in a study of children (ages 4 to 10 years) and adults (ages 21 to 58 years). Additionally, these researchers found a significant slowing of synchronization tapping rates with age for isochronous, rhythmic, and musical stimuli, indicating the direction of attention to higher (slower) metrical levels of temporal stimuli as age increases.

McAuley et al. (2006) conducted a large-scale study of 305 participants from ages 4 to 95 years. Measures were taken of participants' spontaneous motor tempo (SMT)[5], fastest and slowest tapping rates, and preferred perceptual tempo (PPT). The PPT task involved rating isochronous monotone sequences presented at different rates as "too slow", "too fast", or "just right". The authors reported that SMT decreased with age in a cubic relationship. This relationship suggests that SMT slows with age during childhood and late adulthood, but remains relatively constant in middle age. The authors also reported a significant correlation between participants' SMT and PPT, suggesting that preferred tempo measures are highly similar regardless of whether preferred tempo is measured by production or perception tasks.

2.6.2 Overview of the Research

The above findings suggest an age-related slowing of self-paced tapping, a preference for slower perceived isochronous stimuli, and a shift in the focus of attention toward higher metrical levels with increasing age. However, previous research has not explored the extent to which tempo representations of actual pieces of music with long-term representations in memory might also be affected by ageing. In the present study (Study 2), participants were presented familiar, non-canonical songs and asked to adjust the songs to the speed at which the music sounded "right" to them.

Study 2 tested adult participants across a wide age range (19 to 85 years). As in Study 1, participants were asked to set the speed of 14 familiar tunes to the speed that sounded right to them in two conditions—a perceived and an imagined music condition. The procedure was highly similar to that of Study 1, but with age as the variable of interest, rather than physiological arousal. Also as in Study 1, participants turned a dial to either adjust the actual music to the speed that sounded "right"

---

[5] SMT is essentially an analogous measure to the spontaneous tapping rates described in the aforementioned studies.

(perception condition) or to adjust the speed of a click track to the beat of an imagined song (imagery condition). As most previous studies of age-related effects on preferred tempo have used tapping tasks, it was of interest in the present study to use a task that did not rely on motor constraints of the participants, in order to confirm and extend the findings of McAuley et al. (2006) on the PPT task. To the author's knowledge, the present study is also the first to investigate age-related effects on tempo representations within musical *imagery*. Thus, the primary contributions of the present study were: 1) the investigation of the influence of age on tempo judgements for familiar songs recalled from long-term memory, 2) the comparison of these tempo judgements in perceived and imagined music conditions, and 3) the use of a task that aimed to minimize the reliance on fine motor movements in collecting participant responses.[6]

As the findings of McAuley et al. (2006) and Drake et al. (2000) suggest an age-related slowing of preferred tempo across several types of musical tempo-related tasks, it was hypothesized in the present study that the tempo at which familiar songs sounded "right" might also decrease with age. Similar patterns of results were also expected across both musical perception and imagery tasks, given the parallels between the cognitive mechanisms implemented in each (e.g., Zatorre & Halpern, 2005). In sum, the hypotheses of Study 2 were as follows:

- Hypothesis 1: Average chosen tempo for familiar melodies will decrease as age increases.

- Hypothesis 2: Similar chosen tempo decreases will be seen in perceived and imagined music conditions with increasing age.

---

[6] The task of turning the dial to adjust the speed of the music or click track does of course require movement, but participants were allowed to move the dial back and forth as much as needed when deciding on a tempo. The movement that participants made, in and of itself, was not the response being measured, as is the case with tapping.

2.7 Method

2.7.1 Design

The design of the present study was highly similar to Study 1, but with age as the predictor variable of interest. A cross-sectional design was employed, in which a sample of participants across a wide range of ages was tested in both a perceived music block and an imagined music block. The dependent variables of interest were the mean tempo choices for 14 familiar songs in the perceived and imagined music tasks.

2.7.2 Participants

Participants were an opportunity sample of visitors to the Science Museum, London. Testing in the Science Museum was made possible by an 11-week residency awarded to the author as part of the museum's "Live Science" series.[7] As a result of this unique testing environment, any museum visitors were welcomed and encouraged to participate in the experiment. However, many of these participants did not finish the entire experiment or were not familiar with all of the songs used as stimuli, due to the wide range of backgrounds and countries of origin within this sample. In total, 189 participants completed at least some part of the experiment, but it was decided to include only those participants who 1) completed the full experiment and 2) reported that they knew at least 12 of the 14 songs used as stimuli. This resulted in a total sample size of 92 participants (54 females, 38 males). These participants ranged in age from 19 to 85 years ($M = 41.4$, $SD = 20.8$).

2.7.3 Ethics Statement

The research was approved by the Ethics Committee of Goldsmiths, University of London and was also approved by the Science Museum,

---

[7] http://www.sciencemuseum.org.uk/about_us/new_research_folder/livescience.aspx

London. All participants gave written informed consent before participating.

2.7.4 Materials

The study used the same sections of the same 14 non-canonical songs as stimuli as Study 1 (see Table 2.1). Four different random orders of these songs were generated for both the perception and imagery conditions and the presentation of each of these random orders was counterbalanced across participants. On each trial of the study, the name of one of the 14 songs and the lyrics corresponding to the section from the MIDI recording were presented on a computer screen. The text of the song lyrics was marked in bold and underlined whenever a word/syllable corresponded with a beat in the music, in order to help participants to discern the beat of the music, as in Study 1 (see Figure 2.1).

An adapted version of the Max/MSP application used in Study 1 was employed. The song title and lyrics for each trial were presented on a computer screen within this application (instead of in a paper booklet as in Study 1) for ease of usage within the Science Museum testing environment. The same tempo shift algorithm (ZTX/DIRAC) was employed to make tempo changes in real time without changes in loudness or pitch and the Powemate dial was again used by participants to control the speed of the music or click track. Participants wore Sennheiser RS 160 wireless headphones during both the perception and imagery tasks.

As in Study 1, it was ensured that if a stimulus was presented at a fast start tempo in the perception condition, the click track was also started at a fast tempo for that same song in the imagery condition and vice versa. This was done so that any bias toward the start tempo of the stimulus (music or click track) would not confound interpretation of the potential effect of interest.

2.7.5 Procedure

At the beginning of the testing session each prospective participant was provided with a list of the 14 songs used as stimuli in the study. He/she was asked to look through the list and indicate to the experimenter whether any of the songs on the list were unfamiliar to him/her. The experimenter recorded the names of the unfamiliar songs for each participant accordingly. As detailed in Section 2.7.2 above, any museum visitor who wanted to try the experiment was allowed to participate, but only those who knew at least 12 of the 14 songs were included in the present analysis.

Next, the experimenter recorded demographic information concerning each participant's age and gender. The participant then completed two blocks of tempo tasks. A perceived music block was completed first, followed by an imagined music block, as in Study 1. Instructions were given at the start of each of the two blocks, in which participants were asked to adjust the music to the speed that sounded right to them (perception condition) or to imagine each song at the speed that sounded right to them and adjust the click track to the beat of this imagined music (imagery condition). They were instructed to turn the dial to the right if they wanted to speed up the music/click track and to the left if they wanted to slow it down. Two example trials were provided before each block to ensure participants were comfortable with the task. The testing session lasted approximately 20 min. All participants received a verbal and written debrief following the testing session.

2.8 Results

2.8.1 Descriptive Statistics

The following results are reported using the dataset of 92 participants who were familiar with at least 12 of the 14 songs used as stimuli. Trials for which a participant reported that they did not know a song were

excluded from all analyses. This resulted in the exclusion of 25 trials total from each task (1.9% of trials), leaving 1263 total usable trials in each task.

A correlation between each participant's perception and imagery task tempo choices (in bpm) was calculated. The overall correlation across all trials from all participants was highly significant, r = .49, p < 001. The correlation between perception and imagery task tempo choices by participant is visualized in Figure 2.4.
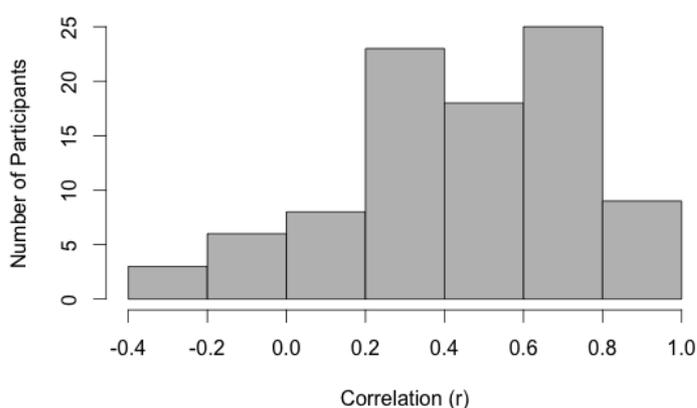
Figure 2.4. Distribution of correlations between perception and imagery task tempo choices for the 92 participants.

Figure 2.4 suggests that although many participants displayed at least moderate correlations in performance on the two tasks (68% of participants had correlations above r = .30), some participants showed very little correlation in performance between the two tasks. These data were examined further and it was found that some trials of the imagery task strongly suggested that participants were doubling or halving their tempo choices in comparison to the perception task (setting the click track to a different metrical level than the one indicated by the beat markings provided to them for the lyrics, as in Figure 2.1). As such, it was decided that trials where the ratio between imagery and perception task

tempo choices was 1.9 or greater or 0.6 or less would be excluded, following the exclusion criteria used by Halpern (1988b) in previous work comparing perceived and imagined tempo. This resulted in an exclusion of 237 trials (18.8% of trials), leaving 1026 usable trials. Following the exclusion of these trials, the overall correlation between perception and imagery task performance increased substantially in magnitude, $r = .67$ ($p < .001$). No effect of age on the number of usable trials remaining was found in a regression analysis, indicating that this issue of doubling or halving metrical levels was not biased towards any particular age group in the sample, $F(1, 90) = 0.76$, $p = .38$.

The subsequent inferential statistics will be calculated using this reduced dataset of 1026 trials. The same analyses have also been conducted on the full dataset, and the results followed the same pattern in all instances.

2.8.2 Investigating the Relationship Between Age and Chosen Tempo

Before conducting analyses related to the predictor variable of interest (age), the distribution of this variable was examined (see Figure 2.5). As the age variable is highly positively skewed it was determined that, in addition to performing the subsequent analyses regarding the relationship of age to chosen tempo on the full dataset, these analyses would also be conducted using two randomly selected subsets of 20 participants under the age of 40 years (see Section 2.8.3 for these additional analyses). The two dependent variables of interest to the present study were the average final tempo settings in bpm for all usable trials completed by each participant in the perception and imagery tasks. These variables will hereafter be referred to as "mean perception tempo choice" and "mean imagery tempo choice".
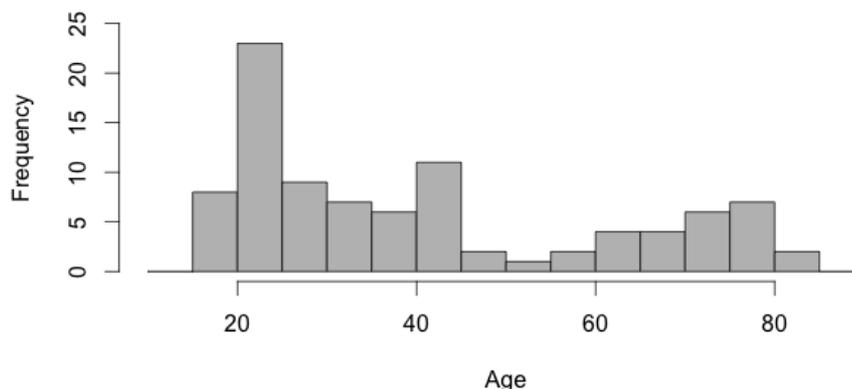
Figure 2.5. Age distribution for the 92 participants.

Non-parametric correlations (Kendall's tau) were calculated between age and mean perception tempo choice and age and mean imagery tempo choice, due to the non-normal distribution of the age variable. The correlation between age and mean perception tempo choice was non-significant, $r_T = .06$, $p = .42$, as was the correlation between age and mean imagery tempo choice, $r_T = .12$, $p = .09$.

Regression analyses were then performed. A rank transformation was applied to the age variable, following the procedures of Conover and Iman (1981), in order to account for its non-normal distribution. As in the above correlations, no significant linear relationships were found between age and mean perception tempo choice (see Table 2.3) or age and mean imagery tempo choice (see Table 2.4). However, evidence of significant quadratic relationships between age and both dependent variables emerged (see Tables 2.3 and 2.4). These relationships are visualized in Figures 2.6 and 2.7. Both graphs show a gradual decrease in tempo choices in middle age followed by a sharper increase in tempo choices in older age.

One can see in the data from the perception task presented in Figure 2.6 that two outliers are present in the distribution toward the older end of the age range. If these two outliers are removed the above pattern of results still holds, with no significant linear effect of age, $F(1, 88) = 0.01$, $p = .93$, but a significant quadratic effect of age, $F(2, 87) = 3.31$, $p = .04$, $R^2 = 0.07$, on mean perception tempo choice.

Table 2.3: *Regression of Mean Perception Tempo Choice on Age*

Equation	R^2	Constant	b1	b2	*F*-statistic	*p*-value
Linear	0.01	110.51	0.06	NA	1.15	.29
Quadratic	0.12	121.24	-0.62	0.01	5.84	.004*

Note: * signifies a significant predictor at the level of $p < .05$.

Table 2.4: *Regression of Mean Imagery Tempo Choice on Age*

Equation	R^2	Constant	b1	b2	*F*-statistic	*p*-value
Linear	0.04	110.66	0.14	NA	3.42	.07
Quadratic	0.15	125.71	-0.82	0.01	8.11	.001*

Note: * signifies a significant predictor at the level of $p < .05$.
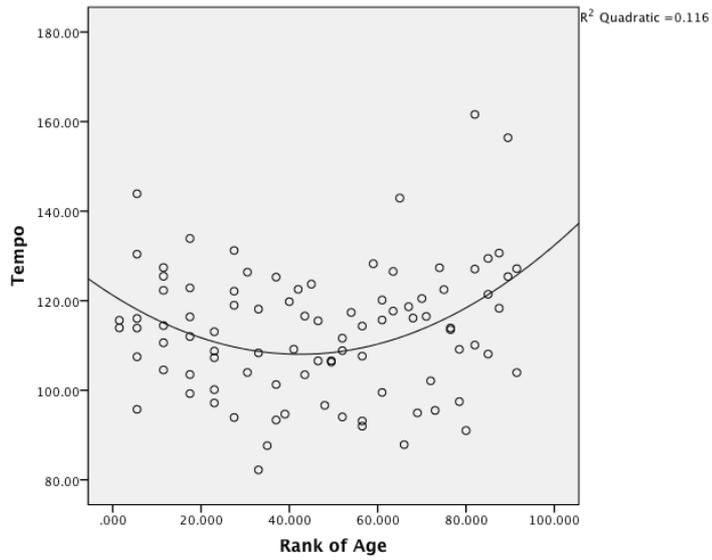
Figure 2.6. Plot of quadratic relationship between age and mean perception tempo choice.
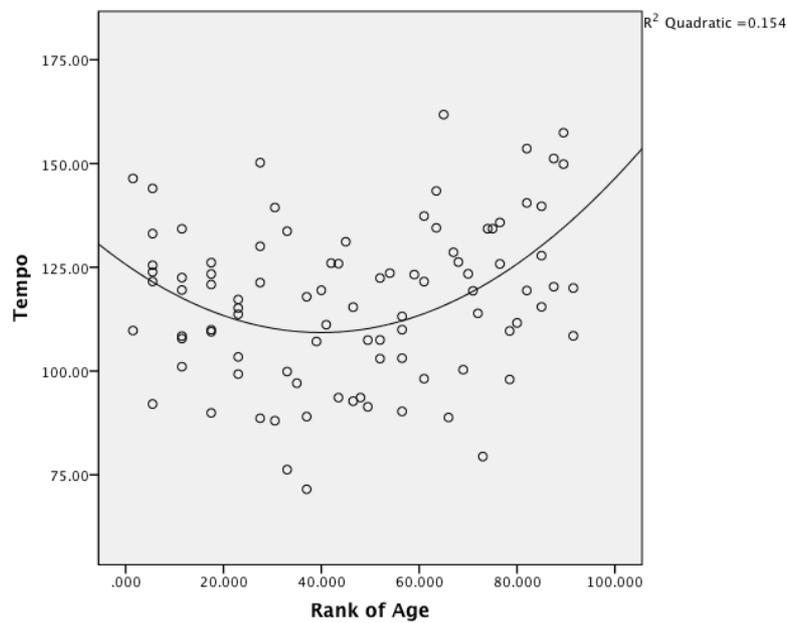
Figure 2.7. Plot of quadratic relationship between age and mean imagery tempo choice.

2.8.3 Regression Analyses on a Subgroup of Participants

A second set of regressions was run with two different random samples of 20 participants under the age of 40 years. This reduced the overall sample size to 62 participants. This analysis was done because the full dataset comprised a much larger proportion of participants in this young age group than in the older end of the age range. As such, it was of interest to test whether the relationships reported above still held when the age variable was less biased toward the younger end of the distribution.

2.8.3.1 Regression Analyses on a Subgroup of Participants: Random Sample 1

A random sample of 20 participants under the age of 40 years was selected. A square root transformation was then applied to further reduce the skewed distribution of the age variable (skewness of Age = 0.18, skewness of SqrtAge = -0.03).

A significant positive linear relationship was found between age and mean perception tempo choice in a regression analysis, $F(1, 60) = 5.15$, $p = .03$, with evidence of a significant quadratic relationship also present, $F(2, 59) = 5.00$, $p = .02$. However, when two outliers are removed from this distribution (as in Section 2.8.2), both these linear and quadratic relationships become non-significant ($p = .15$ and $p = .18$ respectively). A significant positive linear relationship was found between age and mean imagery tempo choice, $F(1, 60) = 12.46$, $p = .001$, with evidence of a quadratic relationship also present, $F(2, 59) = 6.44$, $p = .003$. No outliers were present in the distributions for the imagery task.

2.8.3.2 Regression Analyses on a Subgroup of Participants: Random Sample 2

Next, a second random sample of 20 participants under the age of 40 years was selected. A square root transformation was again applied to reduce skewness of the age distribution (skewness of Age = 0.13, skewness of SqrtAge = -0.11).

In this analysis, no linear effect of age was found on mean perception tempo choice, $F(1, 60) = 2.19$, $p = .14$, but a significant quadratic relationship was found, $F(2, 59) = 3.50$, $p = .04$. However, as in the first analysis of a subset of 62 participants above, these relationships are both non-significant when two outliers are removed from the analysis ($p = .54$ for the linear model, $p = .19$ for the quadratic model). As in the analysis of the first subgroup of 62 participants, evidence for both linear ($F(1, 60) = 7.39$, $p = .01$) and quadratic effects ($F(2, 59) = 4.74$, $p = .01$) of age on mean imagery tempo choice was also found.

2.8.4 Summary of Results

To summarize, all three analyses reported above found evidence of a quadratic effect of age on mean imagery tempo choice (as in Figure 2.7). A small quadratic effect of age on mean perception tempo choice was found in some analyses, but this effect appears to be driven primarily by two extreme values in the data distribution. When these cases are removed, the effect size is reduced and the effect becomes non-significant in two of the three analyses reported above.

2.9 Discussion

The hypotheses of Study 2 were as follows:
- Hypothesis 1: Average chosen tempo for familiar melodies will decrease with an increase in age.

- Hypothesis 2: Similar chosen tempo decreases will be seen in perceived and imagined music conditions with increasing age.

The results of Study 2 do not support either of the *a priori* hypotheses. One significant result was found that was present across all analyses—a quadratic effect of age on average chosen tempo for familiar melodies in a musical imagery task. This effect revealed a decrease in average imagined tempo for familiar songs during early adulthood and an increase in average imagined tempo in older adulthood. These results will be interpreted in the context of existing literature below.

2.9.1 Relationships to Previous Literature

McAuley et al. (2006) found evidence of a cubic relationship between age and spontaneous motor tempo (SMT), as measured by a spontaneous tapping task. This relationship revealed a decrease in SMT in childhood, a fairly stable SMT in mid-adulthood, and a further decrease in SMT in older adulthood. The results of the imagery task in the present work are similar to this finding in some respects, in terms of the fact that a decrease in chosen tempo was observed in early adulthood and then chosen tempo judgements became fairly stable in mid-adulthood (see Figure 2.7). However, these results diverge from McAuley et al.'s (2006) findings when one considers the older adult age range. In the present study, an *increase* in chosen tempo was found in older adulthood, following the fairly stable tempo judgements made during mid-adulthood.

One potential explanation for this finding is that several previous studies have demonstrated that working memory capacity often declines in older adulthood (Baudouin et al., 2006; Salthouse & Babcock, 1991; Salthouse, 1991). Additionally, several researchers have provided evidence that working memory is implicated during musical imagery in order to retain a mental representation of music within the "mind's ear" (Baddeley & Andrade, 2000; Kalakoski, 2001; Smith, Wilson, & Reisberg, 1995). As

such, an age-related decline in working memory capacity might affect performance on the musical imagery task utilized in the present study, in particular impairing the ability to imagine familiar music at slow tempi over the course of a longer time period. Additionally, as attentional resources and the ability to divide attention may also decline with age (McDowd & Craik, 1988; Ponds, Brouwer, & Van Wolffelaar, 1988), older adults who participated in the present study may have been more distracted by the sound of the click track during the imagery task than younger adults, which could have caused difficulties in adjusting this perceived stimulus to a mental representation of a song. The imagery task may be especially difficult when the click track begins playing at a fast tempo, as a greater number of events occurring each second could create more interference with one's mental representation of the music. This may help to explain the finding that older adult participants tended to select faster tempi more often than younger participants in the imagery task.

The perception task utilized in the present study relies less on working memory and attentional resources than the imagery task, as the actual music is presented aloud in the absence of any external distractions from a click track. This may provide some explanation for the divergence of the perception task results from the imagery task results, particularly in regard to the older age range. The absence of any clear effect of age on perception task tempo choices is also notable. This finding suggests that judgements for the tempo at which familiar music sounds "right" may be invariant to age-related changes in time perception when tested using a task in which perceptual cues for the actual music are provided.

However, the above accounts are all post hoc explanations for unexpected findings, and further research is necessary to fully understand the mechanisms underlying these effects (or lack thereof, in the case of the perception task). The results of the present work provide evidence that a slowing of the pacemaker of the internal clock implicated within models of interval timing cannot explain all age-related timing effects, in line with previous suggestions by Block et al. (1998). Future

work including additional measures of working memory, processing speed, attentional capacities, and other types of musical imagery tasks would help to shed further light on this area of study.

2.9.2 Limitations and Future Directions

A variety of additional individual differences (imagery abilities, musical training, stimulus familiarity) were not accounted for, which may have driven the results of the present work in a different direction than hypothesized. One of the limitations imposed by the Science Museum, London testing environment was that all experiments run in the museum should last a maximum of 15 min. As such, very little demographic and background information on the participants could be collected. Future work should account for participant-level differences in musical, imagery, and cognitive abilities to provide a more comprehensive account of how these factors relate to differences in tempo choices for familiar music. A longitudinal study might also help to clarify the present findings. Comparisons within the same person over the course of time would allow for more control over individual differences and would also allow for the investigation of potential cohort effects.

2.9.3 Conclusion

The results of Study 2 do not support the hypothesis that chosen tempo for familiar songs decreases with age. No robust effect of age was found in a music perception condition and a quadratic effect of age was found in a musical imagery condition, indicating a decrease in chosen tempo in early adulthood and an increase in chosen tempo in older adulthood. Further work is needed using additional tasks to measure temporal aspects of musical imagery representations and related cognitive abilities, such as working memory span and attentional capacities, in order to fully understand the potential effects of age on time-related judgements for mental representations of familiar music recalled from long-term memory.

Chapter 3. Measuring Imagined Tempo for Music: A Comparison of Methods and Individual Differences

Abstract

The accuracy of tempo recall for familiar pop music was compared across three tasks—an imagery task that required movement in sync with the beat of the music, an imagery task that did not involve movement to the beat of the music, and a perception task—in a repeated-measures design. Tempo recall was most veridical in the task utilizing a perceived stimulus, with all musical cues present. However, it was also found that tempo recall in the imagery task that required moving to the beat (tapping) was significantly more veridical than tempo recall in the imagery task that did not involve moving to the beat (adjusting a click track to the beat of imagined music). Veridicality of tempo recall in the imagery tasks was positively predicted by musical experience, while performance on the perception task was positively influenced by stimulus familiarity. An effect of the original, recorded tempo of the song stimulus on task performance was also found in both perception and imagery tasks. These results are discussed in relation to previous literature on auditory-motor interactions and musical expertise.

3.1 Introduction

The first two studies within this thesis examined the influence of two specific factors—arousal and age—on perceived and imagined tempo judgements for familiar, non-canonical music. The research outlined in the present chapter extends this line of investigation to examine the influence of several additional factors on perceived and imagined tempo judgements, specifically in terms of how these factors affect the veridicality of tempo recall for a canonical version of a familiar song. The results of the present study, in combination with the results of Studies 1 and 2, will provide a detailed account of the potential influence of a wide

variety of participant-, task-, and stimulus-related factors on both the stability and accuracy of tempo representations within voluntary musical imagery.

Study 3: An Investigation of Factors that Influence the Temporal Veridicality of Voluntary Musical Imagery

3.1.1 Memory for Musical Tempo

The stability and precision with which the tempo of familiar music can be reproduced and recognized is an area of great interest to musicians and researchers of musical memory. As detailed in Chapter 1, musicians often exhibit remarkably stable timing profiles when performing the same piece of music months or even years apart (Clynes & Walker, 1982, 1986; Collier & Collier, 1994). Furthermore, this ability to consistently recall the tempo of familiar music across an extended time period does not appear to be limited only to trained musicians, who have often spent many hours purposefully honing their timing skills. For instance, Bergeson and Trehub (2002) found that a sample of mothers who were unselected for musical training varied in tempo by an average of 3.1% when singing the same songs to their infants at two time points one week apart. Conversely, spoken utterances from these same mothers measured at the same time points showed a 20% average variation in timing. Cross-cultural evidence for this high consistency of memory for musical tempo has also been found in comparisons of performances of aboriginal songs across a period of 34 years, which deviated on average in tempo by only 2 to 3% (Bailes & Barwick, 2011). Levitin and Cook's (1996) findings that 72% of sung productions of familiar songs by non-musicians were within 8% of the tempo of the original recording provide evidence for a high degree of temporal veridicality when recalling well-known songs from long-term memory. In addition to these findings indicating a high degree of precision in tempo production tasks, research

on temporal discrimination by Friberg and Sundberg (1995) has revealed that adult participants display an average just noticeable difference (JND) of only 2.5% for discriminating tempo differences between two isochronous tone sequences and also provided evidence that this JND is not affected by musical training. Finally, it appears that some degree of musical temporal precision is present from very early in life. Trainor, Wu, and Tsang (2004) found that after a one-week exposure phase to novel melodies, 6-month-old infants were able to discriminate these melodies when played at the original tempo from versions played 25% faster or slower.

3.1.2 Memory for Imagined Tempo

The precision of recall for musical tempo specifically within voluntary musical imagery has been investigated by Halpern (1992), who found that the consistency with which the same individual imagined the tempo of a familiar song was very high within the same session, although only musically trained participants maintained this high consistency of imagined tempo recall when testing sessions were separated by 2 to 5 days. These results are somewhat in contrast to the aforementioned findings of highly consistent long-term tempo recall in non-musicians reported by authors such as Bergeson and Trehub (2002). One source of this discrepancy might lie in the fact that Bergeson and Trehub used a sung recall paradigm, in which participants received both auditory and motor feedback during the course of their singing. In contrast, Halpern's task involved tapping to the beat of imagined music and thereby only provided minimal motor feedback and little other sensory feedback that could be used to refine tempo judgements in real time. Also of relevance to the present discussion is the work of Janata and Paroo (2006), who found timing judgements to be impaired in an imagery task in comparison to a perception task, and Janata (2012), who reported a general increase in performance accuracy with an increase in the amount of sensory support provided across three imagined *pitch* tasks. Both of these studies

provide additional support for the idea that increased sensory feedback can aid performance in imagery tasks.

The present research will take a somewhat similar approach to Janata's (2012) pitch imagery research within the imagined *tempo* domain, by comparing performance of the same participants across three tempo recall tasks that vary in terms of the auditory as well as motor feedback afforded by each task. If one considers the aforementioned studies of imagined musical tempo (Halpern, 1992; Janata & Paroo, 2006) along with various other methods used in previous research for measuring the tempo of musical imagery (see Table 1.1 in Chapter 1 for a overview), it becomes apparent that a great variety of different tasks have been used to quantify imagined tempo representations. It is, as of yet, unclear whether some of these tasks might constitute a more precise measure of imagined tempo than others, whether some might pose more difficulties to participants in terms of the degree of sensory feedback incorporated in the task and the memory demands imposed on participants, and whether the way in which task performance is assessed (e.g., correct/incorrect judgement, assessment on a continuous scale) has an influence on the results reported in previous literature. The present work aimed to address some of these important questions by comparing performance in two imagery tasks and one perception task using the same participants and same stimuli for each.

In addition, the veridicality of tempo recall within imagery, in terms of how accurately the tempo of a piece of music is recalled in comparison to the original version of the piece, has not previously been investigated in a paradigm in which the music is recalled purely as imagery (in the absence an auditory component such as singing the song aloud or making a tempo judgement about a presented recording, as in e.g., Levitin & Cook, 1996). As such, the present study aimed to provide some first evidence to address this under-investigated question using two types of imagined tempo tasks that do not involve any perceptual exposure to the song stimulus. This particular research question, in addition to being

of interest to musical imagery researchers, is also highly relevant to musical performers, who regularly make use of musical imagery representations when performing (Keller, 2012). These anticipatory images are often generated either in the absence of perceptual input (e.g., at the beginning of a piece or following a rest) or in concurrence with conflicting perceptual input (e.g., imagining an upcoming note while playing another note) and musicians require these images to be as precise and vivid as possible for optimal performance. As such, research on the veridicality of tempo recall for familiar music in conditions that vary in terms of auditory and motor feedback can be potentially informative to musical performers.

3.1.3 Imagined Tempo and Motor Engagement

Evidence from several sources suggests that probing musical imagery representations using a task that engages the motor system might improve accuracy of tempo judgements over non-motor probing methods. For example, Clynes and Walker (1982) employed a repeated-measures design in which they asked musicians to perform, imagine, and conduct while imagining the same piece of music. The authors reported that the total duration of the conducted version of the piece was generally closer to the duration of a musician's actual performance than the duration of the purely imagined version. These results suggest a role of motor engagement in refining temporal aspects of imagery representations, at least in highly trained musicians. Manning and Schutz (2013) found that tapping during silent intervals improved performance on subsequent timing tasks (i.e., judging whether a final probe note is on time or not) relative to a task in which participants relied solely on auditory imagery without an overt motor component. This led the authors to suggest that moving in time with the beat (even an imagined beat) can actually improve the perception of timing. Additionally, correlational research has revealed a positive relationship between auditory imagery and sensorimotor synchronization abilities in musicians, such that individuals who were able to form pitch images with greater acuity were also more

precise in their performance on synchronization tasks involving tapping to the beat of both a stable metronome and a dynamic pacing signal (Pecenka & Keller, 2009). Finally, anecdotal reports from participants in previous research have suggested that tapping a finger along to imagined music supported more accurate imagery timing during the course of various musical imagery tasks (Schaefer, Vlek, & Desain, 2011; Schaefer, 2014). The present study aimed to extend these findings by comparing the veridicality of tempo recall between an imagery task in which participants moved in time to the beat of the imagined music and an imagery task in which they did not move to the beat.

3.1.4 Participant- and Song-Level Factors

Another potential source of the discrepancies in imagined tempo task performance reported in previous literature is the possible influence of participant-level factors, such as expertise, or song-level factors, such as the range of song tempi used in the stimulus set. The present study aimed to measure several such factors in order to provide a more comprehensive account of the influence of participant- and song-level variables on performance in each of the three tempo recall tasks used in the present study.

Some previous research has reported that musicians perform better in various aspects of musical imagery tasks than non-musicians (Aleman, Nieuwenstein, Böcker, & de Haan, 2000; Bishop, Bailes, & Dean, 2014). Of particular relevance to the present research are studies that have found positive effects of musical training on performance in tasks that required participants to tap to the beat of imagined music (Halpern, 1992) or to tap out the rhythm of imagined music (Bishop et al., 2013). Additionally, Janata (2012) found a significant positive relationship between musical training and performance on his two imagined pitch tasks that involved the *least* degree of sensory support, indicating that training can aid performance on pitch discrimination tasks that are more reliant on imagery abilities than perceptual input. In light of these previous

findings, musical training was included as a predictor variable of interest in the present research. In this study, formal musical training was assessed using the Musical Training dimension of the Goldsmiths Musical Sophistication Index (Gold-MSI; Müllensiefen et al., 2014b). This self-report tool confers certain advantages over methods used to quantify musical training in previous research, as it provides a composite score that includes measures of several facets of the formal musical training experience, such as years of lessons, time spent practicing, and years of music theory training.

Aside from musical training-related effects, minimal research has been conducted to examine the potential influence of other variables related either to participants' previous experience or to the properties of a stimulus itself on performance in imagined tempo tasks. To extend this line of research, several additional potentially relevant participant- and song-level variables were examined. The Active Engagement dimension of the Gold-MSI was administered, which captures information such as how often a participant attends concerts and actively listens to music. The inclusion of this measure represents a new line of investigation in the musical imagery literature, as it is presently unknown whether only formal training on an instrument is related to one's imagined tempo abilities, or whether "informal training", in terms of regular listening and engagement with music in a non-performative context, can also enhance these abilities. Auditory imagery abilities are also known to vary across the population (Aleman et al., 2000; Highben & Palmer, 2004). Therefore, it was predicted that participants who reported more vivid auditory imagery experiences in general or who were better able to control their auditory imagery would perform more accurately on the imagined tempo tasks in the present study. Self-reported vividness and control of auditory imagery were assessed using the Bucknell Auditory Imagery Scale (BAIS; Halpern, 2015). Finally, previous research indicates that music listeners tend to exhibit a "preferred" tempo for perceived music of around 100 to 120 beats per minute (bpm), such that metrical levels around this tempo are more salient to participants in listening experiments than other metrical

interpretations (McAuley et al., 2006; McKinney & Moelants, 2006; Van Noorden & Moelants, 1999). When asked to tap at a "comfortable rate" in tasks used to measure spontaneous motor tempo, adult participants also tend to produce a tapping rate around 100 to 120 bpm (McAuley et al., 2006; Moelants, 2002). In light of these findings, the effect of the original, recorded tempo of each of the 12 songs used as stimuli in the present study on performance in the imagined tempo tasks was examined. This allowed for an investigation as to whether any sort of "preferred tempo" could also be found for imagined music, such that participants performed more accurately in recalling stimuli from certain recorded tempo ranges than others.

3.1.5 Overview of the Research

To summarize, the present study compared tempo recall for the same 12 songs in three conditions: two imagery conditions and one perception condition. The two imagined tempo tasks employed in the present study were purposely constructed to be similar to those used in previous literature, to enable comparisons to this literature. For the imagery task that did *not* require movement in time with the beat of the imagined music, a similar task was used to that of Halpern (1988b), who asked participants to set the speed of a metronome to an imagined song. The task in the present study used an isochronous click track, the speed of which could be changed in real time in very small increments, as in the imagery tasks constructed for Studies 1 and 2 of this thesis. For the imagery task in which participants moved in time with the music, participants tapped along to the beat of the music as they imagined a song, as in Halpern (1992). Although these two types of imagery tasks have been used to probe imagined tempo in various ways within previous literature, the novelty of the present study lies in the comparison of the veridicality of tempo recall for pieces of music from within long-term memory using both types of imagery tasks, as well as a perceived music task, within the same participants.

In summary, the main aims of the present study were to assess the veridicality of tempo recall for familiar pop songs in two imagery tasks and one perception task, and to examine the impact of individual differences on performance on these tasks. Hypotheses were made based on previous literature:

- Hypothesis 1: An imagery task in which participants moved in time to the beat of the imagined music (by tapping) would result in more veridical tempo recall than an imagery task in which participants did not move in time with the beat (click track adjustment).

- Hypothesis 2: Overall, tempo recall would be most veridical in a perception condition, where a high degree of auditory feedback (lyrics and full instrumentation) was present.

Performance on these three tempo tasks was also expected to be influenced by participant-level factors—such as musical training, musical engagement, and auditory imagery abilities—and song-level factors—such as the original tempo of the song and prior familiarity of the song to the participant. The results of the research were anticipated to provide novel evidence regarding the veridicality of tempo recall within musical imagery and the influence of task demands and individual differences on performance on different types of tempo recall tasks.

3.2 Method

3.2.1 Design

All participants completed three tasks that used the same 12 pop songs in a repeated-measures design. In all three tasks, participants were instructed to produce or adjust a stimulus to the original, recorded tempo of a familiar pop song. Two musical imagery tasks, hereafter referred to as the Imagery (motor) task and Imagery (non-motor) task, and one musical perception task, hereafter referred to as the Perception task, were employed. Participants also completed questionnaires (detailed below) in order to investigate individual differences in musical training and

engagement, auditory imagery vividness and control, and familiarity that might relate to task performance.

3.2.2 Participants

Participants were 25 university students (5 male), ages 22 to 36 years (M = 26.72, SD = 3.48). Prior to the experiment, all prospective participants were sent an email with the title and artist of each of the 12 pop songs used as stimuli in the experiment. Only prospective participants who confirmed that they were familiar with all 12 songs, such that they could recall each song in their head, were selected to participate in the experiment. All participants received a payment of £10 for their time at the end of the testing session.

3.2.3 Ethics Statement

The study protocol was approved by the Ethics Committee of Goldsmiths, University of London, UK. Written informed consent was obtained from all participants.

3.2.4 Materials

3.2.4.1 Musical Stimuli

In order to select the pop songs that would be used as stimuli in the present work, a pilot study was conducted with 20 participants. None of these 20 volunteers participated in the main study reported in this chapter, but these volunteers were sampled from the same population as the main study participants (undergraduate and postgraduate students at Goldsmiths, University of London). A list of the song titles and performers of 145 pop songs was distributed to the 20 volunteers, which was created in previous research to represent a wide range of chart-topping pop songs from within the past 50 years (Jakubowski & Müllensiefen, 2013). A check box was provided next to the name of each of the 145 songs. Participants were asked to check the box next to a song if they could

replay at least some part of that song in their head. Using this pilot data, 12 songs that were rated as being familiar to at least 75% of these participants were chosen as stimuli for the present experiment. These 12 songs were also chosen to represent a wide range of original, recorded tempi (see Table 3.1 for original tempi).

For the purposes of the main experiment, a section from each of the 12 pop songs of 16 beats in length was selected. In pilot work, these sections were deemed to be highly recognizable parts of the songs. These excerpts were generally from the chorus, except in cases where the chorus had very repetitive lyrics that might be confusing as to what section of the song was being probed (e.g., "Let it Be") or in which the chorus began with long note durations that might obscure a sense of beat (e.g., "My Heart Will Go On"). The selected sections are presented in Table 3.2.

The tempo of each of these 16-beat musical excerpts was calculated directly from the original recording of the song, by using the sound editing software Audacity to locate the onset time of the first note and offset time of the last note of the musical excerpt. The duration (in seconds) between these two points was divided by 16 beats in order to calculate the duration of one beat. Dividing 60 by the duration of one beat then gives the tempo of the excerpt in beats per minute (bpm). The tempo calculations for each musical excerpt are listed in Table 3.1. These tempo calculations were then compared to the tempo data for each song listed on two online databases—songbpm.com and Geerdes MIDI music database. This was done to ensure that the metrical level at which the experimenter determined each song's tempo was in line with other existing measurements of the song tempo. This was indeed the case for all 12 songs.

For the Perception task, excerpts of the original recordings corresponding to these 16-beat sections were extracted using Audacity. Fade-ins and fade-out of 0.25 s were applied to each musical excerpt. For the two

Imagery tasks, participants did not hear any music aloud and were presented only with the song title, artist, and lyrics of the song excerpt.[8] The lyrics that coincided with the beat of the music were marked in bold and underlined to assist the participants in finding the beat and to ensure that all participants were defining the beat at the same metrical level. Any beats in the music that occurred between lyrics were marked as " ___ ". Some examples of this beat marking technique are provided in Figure 3.1.

---

[8] Participants were also presented the song title, artist, and lyrics in the same format during the Perception task to ensure that the three tasks were as similar in design as possible.

Table 3.1: *Songs Used as Stimuli with Recorded Tempo of Each Section Used*

Song	Artist	Recorded Tempo (bpm)
Baby One More Time	Britney Spears	93.0
Billie Jean	Michael Jackson	117.1
Every Breath You Take	The Police	117.5
Hotel California	The Eagles	73.6
Imagine	John Lennon	76.0
Last Christmas	Wham	104.8
Let it Be	The Beatles	75.7
Like a Virgin	Madonna	120.2
My Heart Will Go On	Celine Dion	49.8
Stairway to Heaven	Led Zeppelin	72.5
Thriller	Michael Jackson	116.3
Wonderwall	Oasis	86.6

Table 3.2: *Sections of the 12 Pop Songs Used as Stimuli in All Tasks*

Song	Lyrics to Section Used as Stimulus
Baby One More Time	When I'm not with you I lose my mind, Give me a sign. Hit me baby one more time.
Billie Jean	Billie Jean is not my lover. She's just a girl who claims that I am the one.
Every Breath You Take	Every breath you take, Every move you make
Hotel California	Welcome to the Hotel California Such a lovely place, such a lovely place, such a lovely face
Imagine	Imagine there's no heaven. It's easy if you try.
Last Christmas	Last Christmas I gave you my heart But the very next day you gave it away.
Let it Be	When I find myself in times of trouble Mother Mary comes to me, Speaking words of wisdom, let it be.
Like a Virgin	Like a virgin, Touched for the very first time
My Heart Will Go On	Every night in my dreams, I see you, I feel you That is how I know you go on.
Stairway to Heaven	There's a lady who's sure all that glitters is gold, and she's buying a stairway to Heaven.
Thriller	Cause this is thriller, thriller night And no one's gonna save you from the beast about to strike.
Wonderwall	Today is gonna be the day that they're gonna throw it back to you. By now you should've somehow realised what you gotta do.

Figure 3.1. Two example screenshots of the song title, artist, and lyrics with beat marking technique as they were presented to participants in Study 3.

3.2.4.2 Tasks

To administer the three tasks in the experiment, applications were developed in Max 6.1. The applications used for the Imagery (non-motor) and Perception task were adaptations of the applications used for the imagery and perception tasks in Studies 1 and 2 of this thesis (see Chapter 2). As in Studies 1 and 2, ZTX (formerly DIRAC) time stretching technology was used to apply tempo changes to an isochronous click track (in the Imagery (non-motor) task) and the musical recordings (in the Perception task) without changing pitch or loudness. Also as in Studies 1 and 2, a Griffin Powermate assignable USB dial was used to adjust the speed of the click track or music. Sennheiser HD 202 headphones were worn by the participants during both the Imagery (non-motor) and Perception tasks.

In the Imagery (motor) task, participants tapped to the beat of an imagined song on the touchpad of a Mac laptop computer. A series of tap onset times was recorded by detecting a change in the status of the

touchpad using the output of the fingerpinger object for Max 6.1.[9] Inter-tap intervals (ITIs) were then calculated by subtracting consecutive tap onset times.

3.2.4.3 Questionnaires

To examine the relationship between individual differences in musical background and veridicality of tempo recall in the three tasks, the Musical Training and Active Engagement dimensions of the Goldsmiths Musical Sophistication Index (Gold-MSI; Müllensiefen, et al., 2014b) were administered to measure formal musical training (music lessons, instrumental practice, etc.) and engagement with music (concert attendance, listening habits, etc.) respectively (see Appendices 1 and 2 for full set of questions). The Bucknell Auditory Imagery Scale (BAIS; Halpern, 2015; see Appendix 3) was used to assess self-reported auditory imagery experiences in terms of vividness of imagery (Vividness subscale) and control over the sounds within one's imagery (Control subscale). Participants also rated their familiarity with each of the 12 songs used as stimuli in the experiment on a scale from 1 to 5, with 1 being "not very familiar at all" and 5 being "extremely familiar".

3.2.5 Procedure

Each participant completed three blocks of tempo recall/adjustment tasks. In all blocks, participants completed two practice trials of the task for that block using the well-known songs "Happy Birthday" and "Row, Row, Row Your Boat" to ensure they understood the procedure for each task. Participants were allowed to ask questions after completing the practice trials before proceeding to the task.

Blocks 1 and 2 for all participants were Imagery tasks (non-motor and motor). The order of presentation of the two Imagery tasks was

---

[9] Available at http://www.anyma.ch/2009/research/multitouch-external-for-maxmsp/

counterbalanced, such that 12 participants completed the Imagery (non-motor) task in Block 1 and 13 participants completed the Imagery (motor) task in Block 1.

In the Imagery (non-motor) task, participants were presented with the title, artist, and lyrics to the 16-beat excerpt of each of the 12 pop songs (as in Figure 3.1). On each of the 12 trials, they were asked to imagine the part of the song corresponding to the presented lyrics as closely as possible to the original version of the song by the artist named on the screen. Ten seconds of silence were included at the start of each trial, during which participants were to bring a musical image to mind before hearing the isochronous click track. When the click track began playing, participants were to set the speed of the click track to the beat of the song being imagined by turning the Griffin Powermate dial to the right to increase the speed and to the left to decrease the speed of the click track. The start speed of the click track was randomized and started at either 36 bpm or 180 bpm. These start speeds were chosen to lie well outside the range of the correct tempi for the pop songs so that the participants would make tempo adjustments on all or most trials. Participants were told that within each trial they could turn the dial back and forth as much as they wanted until they had decided on a tempo. Participants were also instructed not to move to the imagined music during this task (e.g., by foot tapping, head bobbing). Thus, although this task did require some small hand movements in order to adjust the click track to the beat of the music, participants were not moving in such a way that their body motions could become entrained to the beat of the imagined music. In contrast, the Imagery (motor) task required entrainment to the beat of the imagined music via tapping.

In the Imagery (motor) task, participants were presented with the same sections of the same 12 songs in a different random order than in the Imagery (non-motor) task (the presentation of the song stimuli was randomized across all blocks for all participants). In this task, participants were asked to imagine the song presented on the screen as closely as

possible to the original version by the artist named on the screen and to tap all the way through the part of the song that was presented to them (the full 16-beat excerpt). Participants were allowed to test out their tapping speed on the desk next to them before recording their response on the Mac laptop touchpad. This was done so that this task was analogous to the Imagery (non-motor) and Perception tasks, in which participants could turn the dial back and forth multiple times before making a final tempo decision.

The Perception task occurred in Block 3 for all participants. This ordering was employed so as not to prime any participants with the real, recorded music before either of the Imagery tasks. In the Perception task, participants heard the recording of the excerpt from each pop song aloud while being presented with the song title, artist, and lyrics on the screen. The music started playing at either a relatively fast or relatively slow tempo (60% or 150% of the original tempo), again to ensure that participants made tempo adjustments on all or most trials. The start tempi in this task corresponded to the start tempi of the Imagery (non-motor) task, e.g., if a participant received a slow start speed (36 bpm) for the click track in the Imagery (non-motor) task for the song "Let it Be," he/she also received a slow start tempo (60% of the original speed) for the actual music in the Perception task for "Let it Be." The reason for this was that pilot work for both the present study and Study 1 of this thesis showed a bias of participants' tempo judgements toward the start tempo of the stimulus. As such, it was decided to keep the start tempo of the stimuli (click track or music) similar across the blocks to minimize any potential effects of such a bias in the present study.

Following the three blocks of tempo recall/adjustment tasks, each participant filled out the Gold-MSI and BAIS questionnaires and rated their familiarity with the songs used as stimuli in the experiment. All participants were then debriefed as to the purposes of the study. The entire testing session lasted approximately one hour.

3.2.6 Analysis

3.2.6.1 Imagery (Non-Motor) Task

The data for the Imagery (non-motor) task were filtered in order to exclude trials that likely represented participants doubling or halving the tempo compared to the metrical level at which they were instructed to imagine the song. As in Study 2 of this thesis, a criterion borrowed from Halpern (1988b) was applied such that trials in which the ratio of a participant's chosen tempo to the original, recorded tempo for that song was 1.9 or greater or 0.6 or less were discarded. This resulted in discarding 20% of trials in the present study, leaving a total of 239 usable trials for the Imagery (non-motor) task. For each of the usable trials, an absolute deviation from the original, recorded tempo was calculated as a percentage (0% deviation = participant chose the exact original tempo) and a ratio of the chosen to original tempo was also calculated (ratio of 1 = participant chose the exact original tempo).

3.2.6.2 Imagery (Motor) Task

For the Imagery (motor) task, the data were analysed using a similar methodology to previous tapping research (Benoit et al., 2014). In the pre-processing steps, outliers and artefacts were removed from the data. Outliers were defined as inter-tap interval (ITI) values of more than three times the interquartile range from the median ITI; such values often represent missing taps. Artefacts were defined as ITIs of less than 100 ms in duration, which often represent two parts of the hand/finger touching the tapping surface in quick succession. Following the removal of outliers and artefacts, all trials with fewer than 8 total taps were then discarded, as these trials provided too few taps to produce reliable tempo estimates. This resulted in discarding 2% of the present data (6 trials).

For each of the remaining trials, a mean ITI value was calculated and converted to a bpm value for ease of comparison to the other tasks in the

study. Next, 28 trials (9% of the total data) were discarded on the basis of the ratio of the tapped tempo to the original, recorded tempo being 1.9 or greater or 0.6 or less, in order to exclude trials that strongly suggested participants were halving or doubling the tempo in an analogous fashion to the Imagery (non-motor) task data. Due to a technical problem with the data collection software, one trial of tapping for one participant was not recorded. [10] Following these processing steps, a total of 265 trials remained for subsequent analyses. For each of these usable trials, an absolute deviation from the original, recorded tempo was calculated as a percentage (0% deviation = participant tapped the exact original tempo) and a ratio of the tapped to original tempo was calculated (ratio of 1 = participant tapped the exact original tempo).

3.2.6.3 Perception Task

For each trial of the Perception task, an absolute deviation from the original tempo of the song was again calculated as a percentage and a ratio between the chosen tempo and original tempo was also calculated.

3.3 Results

3.3.1 Imagery (Non-Motor) Task

The mean absolute deviation from the original, recorded tempo across all trials of the Imagery (non-motor) task was 25.9% (SD = 17.3%). The mean ratio of the chosen to original tempo across all trials was 1.11 (SD = 0.25). The absolute deviations from the original tempo and ratios of chosen to original tempo, averaged across all trials for each of the 25 participants, are displayed in Figure 3.2.

---

[10] All other trials were completed in all tasks.
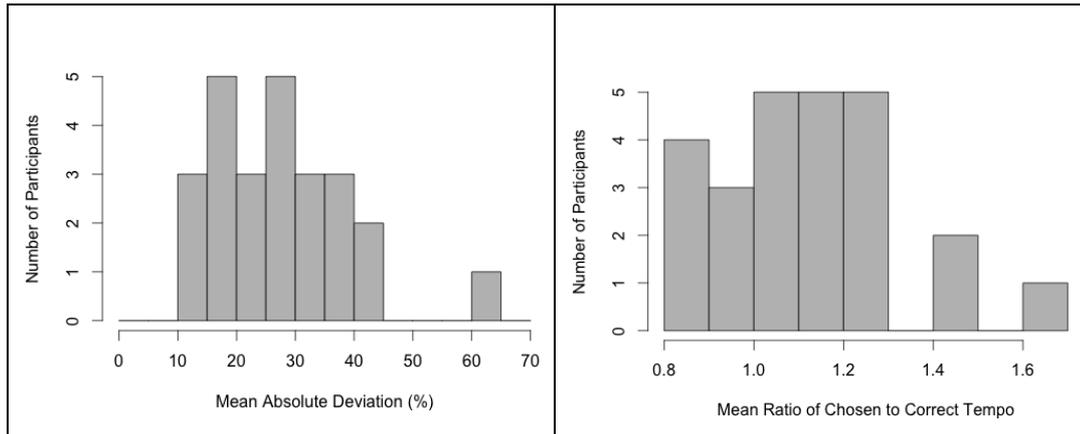
Figure 3.2. Distributions of mean absolute deviations from the original tempo (as percentages) and ratios of the chosen to original tempo for each participant (Imagery (non-motor) task).

Following the initial data filtering steps reported in the Analysis section, some participants had fewer usable trials than others. Figure 3.3 displays the number of usable trials by participant. These data suggest that some participants may have experienced more difficulty with the task of setting a click track to imagined music than others. Thus, a second exploratory analysis was conducted using only the 21 participants who had at least 8 usable trials for the Imagery (non-motor) task. The mean absolute deviation from the original tempo across these 21 participants was 24.1% (SD = 18.3%) and the mean ratio of chosen to original tempo was 1.08 (SD = 0.27). The most accurate of these 21 participants deviated from the original tempo by 10.3% on average and the least accurate participant deviated by 36.8% on average.
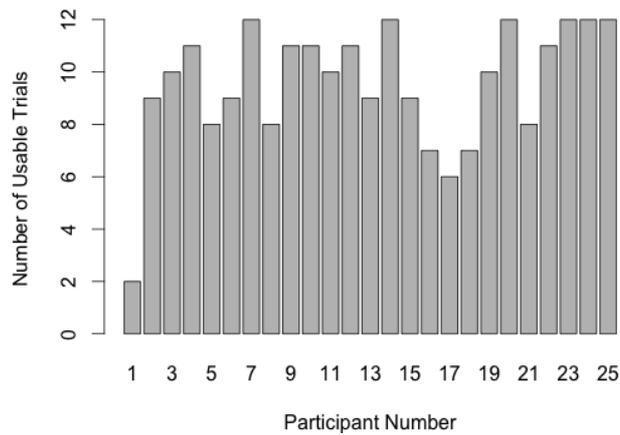
Figure 3.3. Number of usable trials for each participant (Imagery (non-motor) task).

3.3.2 Imagery (Motor) Task

In the Imagery (motor) task, the mean absolute deviation from the original tempo across all trials was 18.4% (*SD* = 13.4%). The mean ratio of the tapped to original tempo across all trials was 1.02 (*SD* = 0.20). Results of this task are visualized in Figure 3.4.
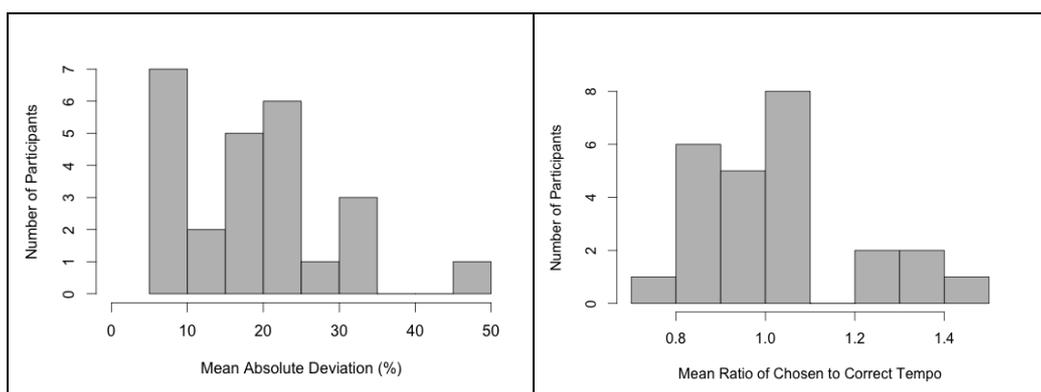
Figure 3.4. Distributions of mean absolute deviations from the original tempo (as percentages) and ratios of the chosen to original tempo for each participant (Imagery (motor) task).

Similarly to the Imagery (non-motor) task, some participants had more usable trials than others (see Figure 3.5), suggesting that certain participants may have experienced difficulty in tapping to the beat of imagined music. A second exploratory analysis was conducted using the 22 participants who had 8 or more usable trials. The mean absolute deviation from the original tempo for all trials across these 22 participants was 18.1% (*SD* = 13.8%) and the mean ratio was 1.03 (*SD* = 0.19). The highest performing of these participants deviated from the original tempo on average by 6.5% and the lowest performing participant by 45.5% on average.
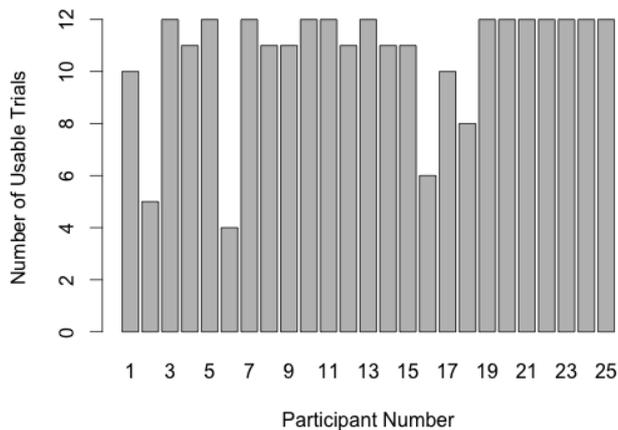
Figure 3.5. Number of usable trials for each participant (Imagery (motor) task).

3.3.3 Perception Task

On average, participants' chosen tempi deviated from the original tempi by 8.1% (*SD* = 6.3%) in the Perception task. The mean ratio of chosen to original tempo across all trials was 0.98 (*SD* = 0.10). The mean absolute deviation values for each participant ranged from 2.8% to 18.6%. Mean absolute deviations and ratios for each participant for the Perception task are plotted in Figure 3.6.
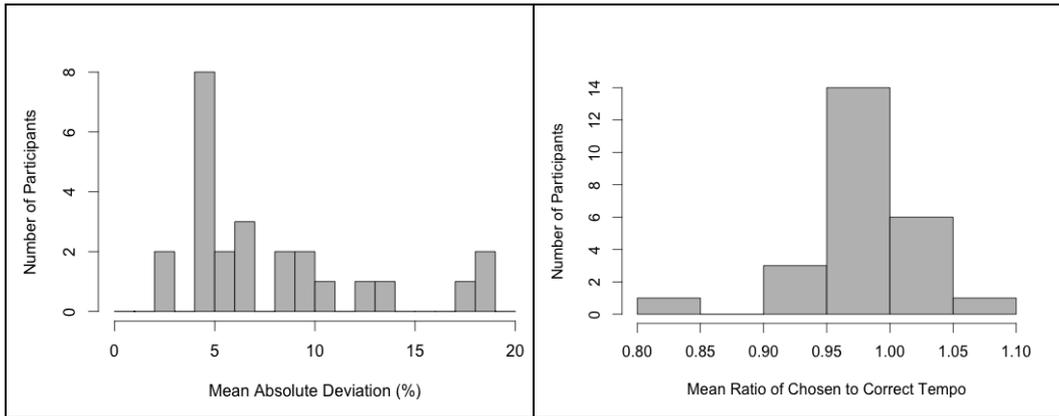
Figure 3.6. Distributions of mean absolute deviations from the original tempo (as percentages) and ratios of the chosen to original tempo for each participant (Perception task).

An inspection of the data using the boxplot() function in R revealed that there are two outliers present in the absolute deviation distribution for this task. If these two cases are excluded, the mean deviation from the original tempo for the sample of 23 remaining participants is 7.2% (SD = 5.8%) and the mean ratio of chosen to original tempo is 0.99 (SD = 0.09). Overall, 10 participants deviated on average by 5% or less from the original tempo of the pop songs in this task.

3.3.4 Comparisons Between Tasks

The overall mean absolute deviations from the original tempo across all usable trials for each task are shown in Table 3.3 and Figure 3.7. On average, performance was most accurate in the Perception task and least accurate in the Imagery (non-motor) task. The effect of task on performance (measured as absolute deviation from the original tempo) was investigated using a linear mixed effects model. A mixed effects model was used as it can take into account the differences in variance across the three tasks and "Participant" could be included as a random effect to take account of the multiple observations per participant and the different number of trials for each participant on each task. Overall, there was a significant difference in task performance between all three tasks (Perception vs. Imagery (motor): t(777) = 8.63, p < .001; Perception vs.

Imagery (non-motor): $t(777) = 14.41$, $p < .001$; Imagery (motor) vs. Imagery (non-motor): $t(777) = 5.86$, $p < .001$; all comparisons Bonferroni corrected).

Table 3.3: *Mean Absolute Deviations for Each Task (All 25 Participants)*

Task	# of Usable Trials	Mean Absolute Deviation (SD)
Perception	300	8.1% (6.3%)
Imagery (motor)	265	18.4% (13.4%)
Imagery (non-motor)	239	25.9% (17.3%)
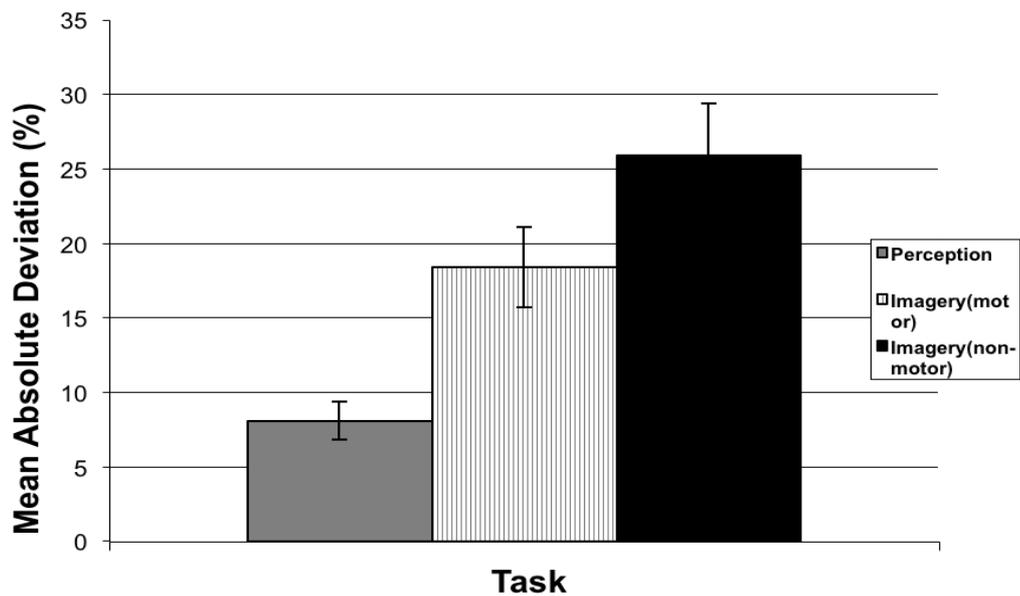
Figure 3.7. Mean absolute deviations for all three tasks (bars represent standard error of the mean).

Next, correlations in performance between the three tasks were examined (see Table 3.4). Performance on the tasks was quantified as the mean absolute deviation from the original tempo for each participant

for each task. Spearman correlations were calculated due to the presence of some outliers and positive skewness in the data. A significant positive correlation was found between performance on the two Imagery tasks, $\rho(23) = .58$, $p = .003$, and a significant positive correlation was found between performance on the Perception task and Imagery (non-motor) task, $\rho(23) = .40$, $p = .05$. Performance on the Perception and Imagery (motor) tasks was also positively correlated, but this correlation did not reach statistical significance, $\rho(23) = .32$, $p = .12$.

3.3.5 Participant- and Song-Level Factors

In order to investigate individual differences related to task performance, three linear mixed effects models were fitted. Separate models were constructed for each of the three tasks with musical training (GoldMSI-MT), active musical engagement (GoldMSI-AE), auditory imagery vividness (BAIS-V), auditory imagery control (BAIS-C), familiarity with each song (Familiarity), and the original, recorded tempo of each song (OrigTempo) as predictor variables. "Participant" was included as a random effect in the models to take account of the multiple trials recorded per participant on each task and the dependent variable of interest was the absolute deviation from the original tempo of a song.

After the full models were fitted with all six predictor variables included, all non-significant predictors were removed from each model. Each of the three models was then fitted again with only the previously significant predictors included, which improved the fit of the model substantially in all cases over the full model, based on the Bayesian Information Criterion (BIC) (see Table 3.4). In the Imagery (non-motor) task, only active musical engagement was a significant predictor of task performance, such that higher engagement scores predicted more accurate task performance (smaller absolute deviations from the original tempo). In the Imagery (motor) task, musical training and a faster original, recorded tempo predicted more accurate performance. Finally, songs that were more familiar to participants and had a faster original, recorded tempo

were recalled at a more accurate tempo in the Perception task than less familiar and slower songs. The reduced models for each of the three tasks are presented in Table 3.4.

Table 3.4: *Linear Mixed Effects Models for Performance on Each Tempo Task (Reduced Models)*

1. Imagery (non-motor) Task

	Estimate	S.E.	*t*-value	*p*-value
Intercept	0.546	0.097	5.62	< .0001
GoldMSI-AE	-0.007	0.002	-2.94	.007

Note: BIC of full model with all predictors = 2.27, BIC of reduced model = -58.89

2. Imagery (motor) Task

	Estimate	S.E.	*t*-value	*p*-value
Intercept	0.466	0.052	8.95	< .0001
GoldMSI-MT	-0.004	0.001	-3.36	.003
OrigTempo	-0.002	0.0004	-4.46	< .0001

Note: BIC of full model with all predictors = -128.69, BIC of reduced model = -177.47

3. Perception Task

	Estimate	S.E.	*t*-value	*p*-value
Intercept	0.215	0.027	7.87	< .0001
Familiarity	-0.016	0.005	-3.63	.0003
OrigTempo	-0.0007	0.0002	-3.96	.0001

Note: BIC of full model with all predictors = -592.67, BIC of reduced model = -647.88

3.4 Follow-up Study

As a final step in the analysis of the present results, a follow-up study was conducted to investigate whether the possibility of familiarity with other

versions of the 12 pop songs used as stimuli may have affected the veridicality of recall in the main study. This question is particularly relevant to the Imagery tasks, in which no music was presented and participants were required to rely solely on their memory of the musical content of each song. Although participants were clearly instructed during all tasks to imagine the version of the song that was performed by the artist whose name was presented on the screen (e.g., Britney Spears, Celine Dion), the possibility that participants were more familiar with a different version of the song and recalled this version instead cannot be entirely ruled out. In the modern era it is extremely difficult to control for participants' exposure to cover versions of songs; even if no commercially successful cover version of a song exists there is still the possibility that a participant has performed or heard a cover version of a song by a local band, sang a version of this song around a campfire with friends, etc. As such, a follow-up study was conducted to investigate the versions of the 12 pop songs used as stimuli in the present study that most readily come to mind.

3.4.1 Method

Two samples of participants were recruited for the follow-up study. The first was a subset of 16 of the 25 participants who had participated in the original study. The other group was a new, independent sample of 33 participants who were selected to be from the same age demographic as the original study participants. The mean age of the follow-up study participants was 29.6 years (range = 22-42, SD = 4.6) and 9 of these participants were male.

The same online task was administered to both samples of participants. On each trial of this task, participants saw the title of one of the 12 pop songs and the lyrics to a section of that song. The sections of the songs were exactly the same sections as had been selected for use in the main study of this chapter. Participants were asked to imagine the section of the song playing in their heads and to advance to the next page of the

task once they had imagined the song in their heads. On the next page of the task they were presented a recording of the section of the song (the same recording used in the Perception task of the main study). They were then asked to respond to the question "Was this the version of the song you thought of?" The response options to this question were "Yes", "No", and "I wasn't able to bring to mind any version of this song". If a participant answered "No" to the question, he/she was also encouraged to provide information on the version of the song that had been thought of, including the performer and YouTube links if possible. This task was completed for all 12 of the pop songs that had been used as stimuli in the main study. In total, the follow-up study took approximately 5 to 10 minutes to complete.

3.4.2 Results and Discussion

3.4.2.1 New Sample of Participants

As the new sample of 33 participants were recruited solely on the basis of being a similar age to the original study participants, not all of these new participants were familiar with every song used as stimuli. On 9.8% of trials a participant indicated that he/she was unfamiliar with the song in question. On 0.8% of trials a participant reported having accidently thought of a different song, for example "Tears in Heaven" in place of "Stairway to Heaven".

The main question of interest to this follow-up study was to investigate how often participants imagined a different version of one of the 12 pop songs. After excluding from the dataset the trials mentioned above in which participants didn't know a song or imagined the wrong song, the new sample of 33 participants imagined a different version of the song than the version used in the main study of this research in 4.5% of trials. This finding suggests that the versions of the songs used in the main study were the most commonly recalled versions. For four songs ("Wonderwall", "Thriller", "Billie Jean", and "Let It Be"), no other version

was recalled by any participant. For the other eight songs, a different version was imagined by 1 to 4 participants ($M = 2$, $SD = 1.20$). The song for which other versions were most often recalled was "Like a Virgin", for which 4 participants reported imagining a cover version.

It should be noted that the design of the follow-up study was such that an artist name was not presented to participants during the imagery task. This was done in order to directly probe which single version of a song was most readily accessed from memory. However, it is possible that for some of the 4.5% of trials in which participants imagined a different version of a song that a participant did actually know the version used in the main study, but that he/she was just able to recall a different version more readily when probed to think of the first version that came to his/her mind. For instance, one participant thought of a version of "Stairway to Heaven" that her choir had performed. This does not imply that this participant was completely unfamiliar with the Led Zeppelin version of this song; if provided the artist name "Led Zeppelin", as was done in the main study, she may have been able to vividly recall that version of the song as well as the choir version. Thus, the 4.5% of trials in which participants imagined a different song may actually be an *overestimate* of the degree to which the results of the main study were affected by participants imagining different versions of the stimuli.

3.4.2.2 Original Study Participants

A subset of 16 participants from the original study also completed the follow-up study. As the follow-up study was conducted approximately 12 months after the main study, it is likely that these participants did not remember exactly which versions of the 12 pop songs were used in the original study in which they participated. In total, these participants imagined a different version of one of the 12 pop songs on six trials (3.1% of trials). For six of the 12 pop songs no other version was imagined; for the other six songs only one participant imagined a different version of the song. The same comment as made above for the new sample also

applies here: participants more readily thought of a different version of a song on 3.1% of trials, but this does not necessarily mean that they did not know the version used in the original study.

Of the six trials for which a different version of the song in question was imagined, four of these were reported by the same participant. In light of this result, the main data for the study were reanalysed while excluding this particular participant. This subset of data followed the same pattern of results as the original analysis. The mean absolute deviations for the three tasks when excluding this participant are reported in Table 3.5. The differences in performance between the three tasks all remained significant (Perception vs. Imagery (motor): $t(742) = 8.55$, $p < .001$; Perception vs. Imagery (non-motor): $t(742) = 14.14$, $p < .001$; Imagery (motor) vs. Imagery (non-motor): $t(742) = 5.67$, $p < .001$; all comparisons Bonferroni corrected).

Table 3.5: *Mean Absolute Deviations for Each Task (Subset of 24 Participants)*

Task	# of Usable Trials	Mean Absolute Deviation (SD)
Perception	288	8.2% (6.4%)
Imagery (motor)	253	18.8% (13.5%)
Imagery (non-motor)	227	26.4% (17.7%)

3.5 Discussion

The present study aimed to investigate the veridicality of tempo recall for familiar pop songs in both imagery and perception conditions, as well as to compare two methods of probing imagined tempo for music. One

method of probing imagined tempo used a task in which participants moved in time to the beat of imagined music and one used a task with only minimal auditory feedback that did not involve moving to the beat of the music (adjusting the speed of a click track). The hypotheses for the study were as follows:

- Hypothesis 1: An imagery task in which participants moved in time to the beat of the imagined music (by tapping) would result in more veridical tempo recall than an imagery task in which participants did not move in time with the beat (click track adjustment).

- Hypothesis 2: Overall, tempo recall would be most veridical in a perception condition, where a high degree of auditory feedback (lyrics and full instrumentation) was present.

Overall, Hypotheses 1 and 2 were supported by the present research. Several (but not all) of the factors related to individual differences in participants' musical backgrounds and stimulus features were also significant predictors of performance on the tempo recall tasks. These findings will be summarized and interpreted in the following sections.

3.5.1 Task Comparisons

Overall, tempo recall was found to be significantly more veridical in the Perception task than in the Imagery tasks. This may be due to the fact that the Perception task provides a greater amount of auditory feedback to participants when making their tempo choices than the Imagery tasks. This account supports findings from Janata (2012) on the role of "sensory support" in facilitating performance in pitch imagery tasks by revealing a similar result in the tempo domain. The Perception task in the present study also imposes fewer memory demands than the Imagery tasks, as the majority of the musical content of the songs is presented to participants and only the tempo of the music must be recalled.

The findings from the Perception task are comparable to those of Levitin and Cook (1996) who, in their experiment, found that participants sang familiar pop songs within 8% of the correct tempo on 72% of trials. In the

present study, 65% of trials of the Perception task were within 8% of the correct tempo. The slightly lower percentage found in the present study could be due to the use of self-selected music in Levitin and Cook's work, which might facilitate even more accurate performance by increasing familiarity. This idea is supported by the fact that familiarity with a song was indeed a significant predictor of Perception task performance in the present work.

In terms of performance in the Imagery tasks, tempo choices in the Imagery (motor) task were significantly more veridical than tempo choices in the Imagery (non-motor) task. This suggests a refined accuracy for temporal aspects of musical imagery when the motor system is engaged. This finding extends previous research suggesting a relationship between auditory imagery and sensorimotor synchronization (Clynes & Walker, 1982; Manning & Schutz, 2013; Pecenka & Keller, 2009) by demonstrating that the tempi of familiar melodies are able to be reproduced more closely to the correct (veridical) tempo during an imagery task that requires moving along to the beat of the imagined music than one that does not.

Recent research has highlighted the role of close interactions between the auditory and motor systems in the production of music, as well as in purely perceiving a musical beat (Bangert et al., 2006; Grahn & Brett, 2007; Zatorre, Chen, & Penhune, 2007). The present study provides some first evidence that the auditory component within these auditory-motor interactions might be able to be replaced by *imagined* sound, such that the motor output (tapping) interacts with auditory imagery to improve accuracy of time-related judgements. This idea is further supported by the wealth of evidence regarding the similarities between perceiving and imagining music (Hubbard, 2010), particularly the fact that both processes rely on similar brain regions (Zatorre & Halpern, 2005). Future research should investigate in detail how the present findings fit within current models of auditory-motor coupling and sensorimotor feedback

mechanisms (e.g., Warren, Wise, & Warren, 2005; Wolpert, Ghahramani, & Jordan, 1995; Zatorre et al., 2007).

Performance on the Imagery (non-motor) task, in which participants adjusted the speed of a click track to the beat of imagined music, was the least accurate of the three tasks in the study. This task provided fairly minimal auditory feedback, in which participants heard only the sound of the click track and were required to compare the speed of this sound to the beat of their imagined music. In contrast to the Imagery (motor) task, this task also did not provide useful motor feedback; although the participants did move their hand to turn the dial in order to change the speed of the click track, these motions were not done in time with the beat of the music. Thus, the action of moving the dial may actually have created a source of motor interference with the beat of the music, rather than beneficial motor feedback as was present in the Imagery (motor) task. Another difficulty that may have been experienced in the Imagery (non-motor) task is that the sound of the click track at the wrong speed could have influenced or interfered with participants' imagery representations. However, this is somewhat negated by the fact that participants did perform very well overall in the Perception task, which is a very similar task in that the music is heard aloud at the incorrect tempo and could also therefore interfere with a participant's internal representation of the tempo of a song. Further research is needed to clarify the relatively poor performance on the Imagery (non-motor) task. However, as they stand, the present findings suggest that the Imagery (motor) task may be a more suitable method for probing imagined tempo for music than the Imagery (non-motor) task for use in future studies, as this task resulted in more accurate tempo productions overall.

3.5.2 Participant- and Song-Level Factors

The main findings of the present study suggest a wide range of average accuracy between participants for each of the three tempo recall tasks. For example, in the Imagery (motor) task, the highest performing

participant deviated on average from the original tempo by 6.5%, while the lowest performing participant deviated on average by 45.5%. As such, there is much scope in exploring potential individual differences related to performance on each of the three tasks. The present study investigated specifically the influence of music- and imagery-related abilities, song familiarity, and recorded tempo of the song stimuli.

Performance on the Perception task was positively influenced by familiarity with the pop songs. The absence of an effect of musical training or musical engagement suggests that even everyday music listeners who are not formally trained or avid concert goers can make highly accurate judgements about the speed of music in a task where all the musical cues are present-- as long as the music is familiar enough to them. These results might also help to account some previous findings, such as those of Levitin and Cook (1996) and Bergeson and Trehub (2002), in which non-musicians performed extremely accurately at reproducing the tempo of songs that were indeed highly familiar to them (all songs in these studies were self-selected) in tasks in which participants received auditory feedback from their own singing.

Although those with less musical experience tended to perform as well as those with more musical experience in a tempo recall task in which the actual music was heard aloud, it appears that musical background does begin to play a role in accuracy of tempo judgements in imagined music tasks. In the Imagery (motor) task, musical training was a significant positive predictor of task performance. One possible explanation for this result, supported by several behavioural and neuroscientific studies, concerns the enhanced integration between the auditory and motor systems in trained musicians as compared to non-musicians (Bangert et al., 2006; Zatorre, Chen, & Penhume, 2007). Musicians may have a distinct advantage in this type of task, as they quite regularly engage in scenarios (music practice and performance) in which their auditory and motor systems provide refined feedback to one another. This finding also serves to support previous research that has revealed an advantage of

musical training in general in sensorimotor synchronization tasks, e.g., tapping to perceived music (Chen, Penhune, & Zatorre, 2008), by extending this advantage to the imagined music domain. On the other hand, active musical engagement (e.g., concert attendance, regular listening) was a significant positive predictor of task performance in the Imagery (non-motor) task. This finding suggests that, although musical experience appears to enhance performance on this task, this experience does not necessarily need to be in the form of structured, formal lessons on an instrument.

A significant positive effect of the original, recorded tempo of the song stimulus was found on performance in both the Perception and Imagery (motor) tasks, such that songs with a faster original, recorded tempo were reproduced at a more veridical tempo than slower songs. The presence of this effect in the Perception task may be related to previous research suggesting a general preferred perceptual tempo in adults of around 100 to 120 bpm (McAuley et al., 2006; McKinney & Moelants, 2006; Van Noorden & Moelants, 1999), while the results in the Imagery (motor) task may relate to similar findings that the most comfortable spontaneous tapping rates in adults also lie around 100 to 120 bpm (McAuley et al., 2006; Moelants, 2002). The present findings extend this area of research by indicating that accuracy of recall for the tempo of familiar pop songs in both perception and imagery tasks is facilitated as the original song tempo approaches this 100 to 120 bpm range, since the fastest original tempo for a stimulus used in the present study was approximately 120 bpm (for the song "Like a Virgin"). Future work should be conducted to include an even wider range of pop song stimuli, including songs with original tempi both above and below 120 bpm, in order to investigate whether this linear relationship becomes a quadratic relationship when a wider range of original song tempi are included. If accuracy in the Perception and Imagery (motor) tasks was found to decrease for songs with original tempi greater than 120 bpm, this would provide further support for the idea of a facilitation effect of tempi around 100 to 120 bpm in these two tasks.

Finally, no significant effects of self-reported imagery abilities on performance on any of the three tasks were found in the present study, as measured by the two subscales of the BAIS. One reason for this lack of an effect might be that the BAIS measures quite general auditory imagery abilities for environmental, speech, and musical sounds. As such, ratings of the vividness or amount of control over one of these sounds might not be directly related to the more specific task of making refined tempo judgements for imagined music. In future research it might be useful to include more specific self-report measures that assess solely musical imagery abilities—in particular temporal aspects of the experience—as well as objective tests of imagery abilities to complement self-report measures.

3.5.3 Future Directions

Two Imagery tasks were employed in the present research, one of which provided motor feedback while tapping to the beat of imagined music and one of which provided auditory feedback as the click track was adjusted to the beat of the imagined music. A potential extension of the present research would be to include a third Imagery task, in which one's tapping actions were accompanied by auditory feedback, e.g., the click track sound could be played each time the touchpad is tapped. This would allow for an investigation as to whether the *combination* of auditory and motor feedback could further enhance the accuracy of tempo recall within imagined music.

It is apparent from this research that participants varied widely in their performance in terms of veridicality of tempo recall, in particular on the Imagery tasks. Additionally, those who were less musically inclined performed less accurately on the Imagery tasks than participants with more extensive musical engagement or training. These results suggest that there is a need to continue to develop new, more sensitive methods for measuring imagined tempo, particularly in non-musicians. It is not

entirely possible to discern from these results whether more musically experienced participants performed more accurately on the Imagery tasks because they were actually better and/or more experienced at imagining music or whether the Imagery tasks were just more difficult and abstract to non-musical participants, such that other methods might be more suitable for quantifying the internal musical experiences of these participants. Of particular note is the fact that both Imagery tasks required participants to extract the beat of the music, which can be a fairly complex mental process even when an actual perceived musical stimulus is presented (e.g., Repp, 1999). It would be fruitful in future research to test whether other methods, such as asking participants to tap out the rhythm of an imagined song, might be less reliant on musical training than the tasks used in the present research.

3.5.4 Conclusion

In summary, the present study suggests that while accuracy in judging the tempo of familiar music is greater in a perception condition with all musical cues present than in imagined music conditions, engagement of the motor system, musical background (both formal training and informal musical engagement), and the original tempo of the song stimulus itself can significantly affect performance in imagined tempo tasks. These findings have key implications to researchers of musical memory and imagery, as well as to musical performers, who use imagery for tempo as a means of planning and adapting their performances in real time and rely on these images to be as accurate as possible.

Chapter 4. Investigating Temporal Features of Involuntary Musical Imagery

Abstract

A novel method for measuring the tempo of involuntary musical imagery (INMI) within a naturalistic setting was developed, allowing for the investigation of research questions regarding 1) the precision of memory representations within INMI and 2) the interactions between INMI and concurrent affective state. Over the course of 4 days, INMI tempo was measured by asking participants to tap to the beat of their INMI with a wrist-worn accelerometer, while participants documented additional details of their INMI in a diary. Overall, the tempo of music within INMI was recalled from long-term memory in a highly veridical form. A significant positive relationship was also found between INMI tempo and subjective arousal. The results suggest several parallels between INMI and voluntary imagery, music perceptual processes, and other types of involuntary memories.

4.1 Introduction

The first three studies of this thesis have revealed several novel findings regarding the representation of musical tempo within voluntarily generated imagery. These findings include evidence for 1) an effect of arousal on chosen tempo for deliberately recalled familiar tunes, 2) fairly veridical recall of the tempo of familiar music within voluntary imagery, and 3) the use of methods involving motor engagement with imagery as a means of attaining more accurate imagined tempo judgements than non-motor methods. The next study in this thesis will extend this line of research by investigating whether such findings can be generalized to the realm of *involuntarily* recalled musical imagery.

Involuntary musical imagery (INMI), the experience of spontaneously recalled and repetitive music in one's head, appears to be a much more frequent experience during the course of daily life in the Western world than voluntarily imagined music (Beaty et al., 2013; Liikkanen, 2012a). Despite the commonality of the INMI experience, no previous quantitative research has been conducted to measure temporal aspects of INMI. In order to begin to address this highly under-investigated area of research, the first step in the present research was to transfer methods for measuring imagined tempo outside the laboratory in order to study temporal representations within everyday INMI experiences. INMI, as a form of spontaneous cognition, is ephemeral by nature and its occurrence can be difficult to anticipate in a laboratory setting. For this reason, a new method was developed in order to measure the tempo of INMI as it occurred during the course of participants' daily lives, with the aims of maximizing the number of INMI episodes that could be captured from each participant and increasing the ecological validity of the study. This study allowed for an investigation as to whether the findings from the first three studies of this thesis generalize to the INMI experience, particularly in regard to the veridicality and consistency of imagined tempo representations and the relationship between imagined tempo and arousal. Several additional hypotheses based on previous literature were also tested regarding the precision of recall for musical tempo within INMI.

Study 4: Tracking the Tempo of Involuntary Musical Imagery in Daily Life

4.1.1 Developing a New Method for INMI Research

A large amount of previous research on the INMI experience has been descriptive in nature, focusing on its content, phenomenology, and situational antecedents (e.g., Beaman & Williams, 2010; Halpern & Bartlett, 2010; Liikkanen, 2012a; Williamson et al., 2012; Williamson &

Jilka, 2013). These studies have often provided detailed reports regarding the subjective experience of INMI. For instance, participants in some studies have reported that INMI often represents a highly authentic replication of a familiar song in terms of pitch, tempo, rhythm, instrumentation, lyrical content, and timbre (Brown, 2006; Williamson & Jilka, 2013). However, the degree of precision with which the imagery replicates the original music, in terms of pitch, tempo, etc., has not previously been quantitatively measured.

In addition to these self-report-based studies, some experimental research on INMI has been conducted (Beaman et al., 2015; Floridou et al., 2014; Hyman et al., 2013; Liikkanen, 2012b). These studies focused primarily on methods of inducing INMI in the laboratory and whether INMI could be suppressed during different types of concurrent tasks, such as tasks that engaged the phonological loop and articulatory systems (Beaman et al., 2015; Floridou et al., 2014). However, no published studies reporting quantitative measures of the *musical content* of INMI are available at the time of writing.

The present study implemented a new method that quantitatively measured one musical feature of INMI—tempo—as it occurred during daily life. INMI tempo measurements were collected by asking participants to tap to the beat of their INMI while wearing a wrist-worn accelerometer. Additional details regarding each INMI episode were reported in a diary. A tapping-based method was chosen for collecting INMI tempo measurements in light of the results of Study 3 of this thesis, as the tapping method employed in Study 3 resulted in significantly more accurate reproductions of the tempo of familiar pop songs than the click track adjustment method. Additionally, the tapping method was quite suitable for adaption to an ecological setting through the use of a portable, wrist-worn accelerometer to record tapping data.

By obtaining tempo information for individual INMI episodes, the present study was able to gain detailed insights into several questions regarding

the representation of musical tempo within INMI in a more objective manner than has previously been possible. Specifically, the main research questions investigated in the present study concerned 1) the precision of memory representations within INMI and 2) the interactions between INMI and concurrent affective state.

4.1.2 Investigating the Precision of INMI Tempo Recall

As reviewed in Chapters 1 and 3, the precision of *deliberately recalled* musical memories, in terms of veridicality and consistency of recall, has previously been investigated in laboratory-based studies. Several studies have indicated that both musicians and non-musicians are often highly precise in reproducing the tempo of familiar music when singing or performing (Bailes & Barwick, 2011; Bergeson & Trehub, 2002; Clynes & Walker, 1982, 1986; Collier & Collier, 1994; Levitin & Cook, 1996). In Study 3 of this thesis, a mean deviation of 18.4% from the original, recorded tempo was found when participants deliberately imagined and tapped to the beat of familiar songs. This finding suggests somewhat veridical tempo recall within a purely imagined tempo task, although performance on this task was generally less veridical than in previous sung recall paradigms (e.g., Levitin & Cook, 1996). In another study of voluntary musical imagery for tempo, Halpern (1988b) reported a significant correlation between the tempo at which participants set familiar tunes in a perceived music condition and the tempo at which they imagined the same tunes in an imagery condition. However, she also found evidence for a "regression to the mean" for imagined tempo, such that relatively slow songs tended to be imagined faster than their preferred perceived tempo, and relatively fast songs tended to be imagined slower than their preferred perceived tempo.

The present study represents the first known quantitative investigation into the veridicality of *involuntarily recalled*, everyday occurrences of musical imagery. Tempo measurements for each INMI experience of a song that exists in a canonical version were compared to the tempo of the

original, recorded version of the tune. Veridical representations of musical tempo within INMI would suggest a parallel between the memory mechanisms implicated in deliberately recalled music and spontaneously recalled musical imagery occurring within a naturalistic setting. Such a finding would also provide links to other types of involuntary memory. For instance, involuntary *autobiographical* memories have often been found to be as, or even more, specific and detailed in comparison to voluntary autobiographical memories (Berntsen, 1998; Mace, 2006; Schlagman & Kvavilashvili, 2008). As INMI has been classified as a type of involuntary *semantic* memory (Kvavilashvili & Mandler, 2004; Liikkanen, 2012a), the present findings would provide some first evidence as to whether involuntarily retrieved semantic memories might also be recalled with a degree of specificity similar to their voluntarily recalled counterparts. However, if tempo information was not preserved with high fidelity during INMI, this could suggest that other elements, such as one's internal affective state or external environmental sounds and distractions, might significantly affect the stability of tempo within INMI. The present design also allowed for the investigation of two secondary questions regarding the temporal veridicality of INMI: 1) the influence of recent hearing of a song on veridicality of tempo recall, and 2) whether evidence for a regression to the mean could be found for INMI, similar to that reported for voluntary imagery by Halpern (1988b).

The temporal consistency between multiple INMI episodes of the same piece of music was also explored in the present research, in a similar manner to which Halpern (1992) investigated the temporal consistency of voluntarily imagined songs. This was achieved by comparing the tempi of tunes that were repeatedly experienced as INMI within the same participant over the data collection period. Work by Byron and Fowles (2013) has shown a quick exponential decay in the recurrence of the same tune as INMI, thus the number of instances of recurrent INMI tunes was expected to be few. Nevertheless, this exploratory analysis contributed to the overall investigation of the temporal stability of the INMI experience.

4.1.3 Investigating the Relationship Between INMI and Affective States

If spontaneously recalled musical memories are found to rely on similar memory mechanisms to deliberately retrieved music, might they also serve similar functions to purposeful music selection? One of the most common uses of music within Western society is for mood regulation (Juslin & Laukka, 2004; Saarikallio & Erkkila, 2007; Sloboda, O'Neill, & Ivaldi, 2001; Tarrant, North, & Hargreaves, 2000), and a handful of existing studies provide support for the idea that INMI might also serve as a mood regulatory mechanism in the absence of an external music source. For instance, participants in qualitative research have reported associations between their current mood and the type of INMI they experience (Williamson et al., 2012). Some participants in this study reported that the emotional tone of their INMI appeared to match their current mood state, while others reported specific INMI tunes that reoccurred whenever they were in a certain mood state. Diary-based methods have suggested that INMI is more frequent in more alert mood states (Bailes, 2012), such that self-reported drowsiness (on a scale of "alert" to "drowsy") was a negative predictor of whether participants were experiencing INMI. Research on voluntary musical imagery has also revealed parallels between participants' decoding of emotions in perceived and imagined music (Lucas et al., 2010), indicating that music can convey similar emotions whether it is imagined or heard aloud. Finally, research on involuntary autobiographical memory indicates that these types of memories are often more emotional than their deliberately recalled counterparts, suggesting that involuntary retrieval of memories might enhance their emotional qualities (Berntsen & Hall, 2004). However, no previous empirical evidence exists as to whether certain musical dimensions of the INMI experience might relate to specific mood constructs, in a similar fashion to the way in which different features, such as the tempo, musical mode, or texture of a piece of music, can elicit

different emotional responses during music listening (Husain, Thompson, & Schellenberg, 2002; Webster & Weir, 2005). As such, the second main aim of the present study was to use the newly developed method for measuring the tempo of INMI in order to investigate how musical features of the INMI experience might relate to one's concurrent affective state.

Hypotheses for this research question were based on previous findings regarding the relationships between features of perceived music and emotional responses. Several previous studies suggest a link between musical tempo and arousal. Listening to fast tempo music can increase subjective arousal (Husain et al., 2002) and fast tempo music is preferred in high arousal conditions, such as exercise (Edworthy & Waring, 2006; North & Hargreaves, 2000). Study 1 of this thesis has also revealed that physiological arousal can influence tempo choices when participants are asked to indicate the tempo that "sounds right" for familiar, non-canonical songs in both voluntary imagery and perceived music tasks (Jakubowski, Halpern, Grierson, & Stewart, 2015). As such, a positive relationship was predicted between subjective arousal and INMI tempo in the present study.

Emotional valence appears to be less clearly related to musical tempo, but has been related to other features of music such as mode, i.e., major vs. minor, such that the major mode is generally associated with more positive emotional valence and the minor mode with more negative valence (Gagnon & Peretz, 2003; Husain et al., 2002; Webster & Weir, 2005). In accordance with this previous research, the musical mode for each reported INMI song was also determined, with the prediction that major mode INMI would co-occur with more positive emotions than minor mode INMI.

4.1.4 Summary of Research Questions

In summary, the present study employed a novel method to collect INMI tempo data during participants' everyday activities over the course of 4

days. The acquired data were used to investigate two specific research questions. The first main question examined the precision of tempo recall within INMI, particularly in regard to the veridicality of recall in comparison to the original version of the song and the consistency of recalled tempo across multiple INMI episodes within the same participant. The second main question examined the relationships between musical features of INMI--specifically tempo and musical mode--and self-reported affective states--specifically in terms of emotional arousal and valence.

The four main hypotheses of the study were as follows:

- Hypothesis 1: The tempo of INMI for canonical songs will be recalled highly veridically from memory.

- Hypothesis 2: The tempo of INMI will be recalled highly consistently from memory (although the number of instances of recurring INMI episodes within the same participant is expected to be fairly low).

- Hypothesis 3: INMI tempo will be positively related to concurrent ratings of arousal.

- Hypothesis 4: Ratings of emotional valence will be more positive during INMI in the major mode than minor mode INMI.

The results of the study represent a first step towards an understanding of temporal aspects of INMI within daily life.

4.2 Method

4.2.1 Design

A naturalistic study was conducted in which participants 1) tapped to the beat of their INMI while wearing an accelerometer that recorded their movements and 2) recorded information about each INMI episode in a diary during their daily lives over a period of 4 days.

4.2.2 Participants

Participants were 17 volunteers (7 male), ages 20 to 34 years ($M = 24.59$, $SD = 4.20$). The age range of the sample was deliberately constrained in order to control for any potential age-related effects on INMI tempo (see Study 2 of this thesis). All participants were recruited on the basis that they reported experiencing earworms[11] several times a day and the sample comprised both musically trained and untrained participants. A screening task was administered to all participants who did not engage in professional musical activities, in order to exclude any prospective participants who exhibited difficulties in tapping to the beat of musical imagery. The screening task was the Imagery (motor) task from Study 3 of this thesis, which required tapping to the beat of familiar, voluntarily imagined songs in the laboratory. Any participant who was not an outlier on this task in terms of mean tapping variability was deemed eligible for inclusion in the present study. Outliers were defined as data points more than 1.5 times the interquartile range below the first quartile or above the third quartile. All participants received £40 compensation for their participation.

4.2.3 Ethics Statement

The study protocol was approved by the Ethics Committee of Goldsmiths, University of London, UK. Written informed consent was obtained from all participants.

---

[11] Although the term *involuntary musical imagery (INMI)* is used throughout this thesis as a more complete description of the phenomenon, the term *earworm* was used in all instructions given to participants, as this represents a more familiar and colloquial term. Hence, all descriptions of the instructions to participants here will use the term *earworm*.

4.2.4 Materials

4.2.4.1 Measuring INMI Tempo

To record INMI tempo, a GENEActiv wrist-worn accelerometer was employed (http://www.geneactiv.org/; see Figure 4.1). This device resembles a wristwatch and, as such, is a fairly non-invasive tool for measuring participants' movement data throughout the day (Rowlands et al., 2014; Zhang, Rowlands, Murray, & Hurst, 2012). Measurements for the present study were taken at the GENEActiv's maximum sampling rate of 100 Hz. Participants were asked to tap to the beat each time an earworm occurred while wearing the device.

Figure 4.1. GENEActiv accelerometer device used for INMI tempo measurements.

In order to validate the viability of the GENEActiv device for measuring tapping data, a pilot study was conducted in which self-paced tapping was simultaneously measured by tapping on a laptop touchpad (as in Study 3) while wearing the accelerometer. Three participants were instructed to tap at their most comfortable speed for 1 min, as well as their fastest and slowest possible speeds for 1 min, in order to obtain data across a wide tempo range. Timing of tap onsets were registered by the laptop touchpad using MAX 6.1[12]; tap onsets were also extracted from

---

[12] Tap onset times were recorded by detecting a change in the status of the touchpad using the output of the fingerpinger object for MAX (available at http://www.anyma.ch/2009/research/multitouch-external-for-maxmsp/).

the accelerometer data using the analysis procedure described below (see Section 4.2.6.2). Both tap onset time series were processed in order to calculate the mean tempo of the time series in bpm. All tempi calculated with the accelerometer data were within 1 bpm of the tempi measured using the touchpad, thus demonstrating the viability of the device for measuring tapping data.

4.2.4.2 Self-Report Diary Measures

A paper diary was given to participants for reporting on the occurrence and circumstances of each earworm experienced over the course of the 4-day period. Each page comprised 11 questions pertaining to a single earworm episode. Participants were asked to fill in this booklet as soon as possible once they had finished tapping to the beat of the earworm. Each page of the diary asked for the time and date of the episode, the time the diary questions were completed, the name, artist, and section of the earworm song, the last time the song was heard aloud, whether the episode occurred during a repetitive movement (e.g., walking), and information on internal or external events that might have triggered the earworm episode (see Appendix 4 for the full set of diary questions). The categories for the question on how the earworm episodes were triggered were based on categories derived in previous large-scale, qualitative research on INMI triggers by Williamson et al. (2012). The diary also comprised seven mood pairs used in previous musical imagery research by Bailes (2006, 2007, 2012) that were adapted from a study of music in everyday life by Sloboda et al. (2001). These seven mood pairs group into three factors: Positivity, Arousal, and Present Mindedness. Participants were asked to rate their mood on these seven scales in terms of the way they felt just before the earworm began.

4.2.5 Procedure

Participants were asked to choose a 4-day block of time during which they felt they could complete the study most effectively (e.g., they didn't

have other commitments that would greatly impede the recording of their earworms as they occurred). Each participant then met with the experimenter for approximately 15 min to receive the instructions and materials for the study. The experimenter provided a definition for the term "earworm", as follows: "An earworm is a short section of music that comes into your mind without effort (it is involuntary; without any intention to retrieve or recall the music) and then repeats by itself (immediately repeated at least once, on a loop, without you consciously trying to replay the music)". Participants were informed that they would be asked to keep track of their earworms over the next 4 days. Participants were instructed that when they experienced an earworm, they were to tap along to the beat of the music in their heads as accurately as possible while wearing the accelerometer device. They were asked to tap at least 20 times during each earworm episode, in order to provide enough data for analysis. Examples of familiar songs ("Jingle Bells" and "Row, Row, Row Your Boat") were provided in which the syllables of the lyrics corresponding to the beats in the music were marked in bold and underlined to ensure that participants understood what was meant by the beat of the music (see Figure 4.2). The experimenter then showed the participants how to wear the accelerometer and demonstrated the tapping method, in which they were asked to tap with their full forearm on their leg. The experimenter also showed the participants how to press the button on the accelerometer to serve as a marker of the end of each tapping episode. No button press was required at the start of the tapping episode so that participants could begin tapping as soon as they noticed an earworm, without impeding upon the spontaneous nature of the event. The experimenter asked each participant to test out both the tapping method and the button press in the lab. The experimenter then showed participants the paper diary and explained each question to ensure clarity.

Participants wore the accelerometer and carried the paper diary with them for a period of 4 days (96 hours). During this period, participants were asked to tap to the beat of their earworms whenever possible and to

fill out the diary as soon as possible after each tapping period. Participants returned the materials to the experimenter at their earliest convenience after 96 hours had elapsed and were debriefed as to the purposes of the experiment upon their return.

Row **row** **row** your **boat** **gent**-ly **down** the **stream** __

Figure 4.2. Example text used to explain the meaning of a "beat" to participants. Bold and underlined syllables correspond to beats in the music.

4.2.6 Analysis

4.2.6.1 Diary Data Analysis

Hand-written diary data for each participant was inputted into Microsoft Excel for further analysis in Excel and R. A total of 275 INMI episodes were reported in the diaries.

Scores on each of the seven mood pairs were grouped into the three factors (Positivity, Arousal, and Present Mindedness) designated by the original authors of the mood scale (Sloboda et al., 2001) and summed. Scores on reverse-scored items were recalculated before adding as necessary. The Positivity factor comprised the happy/sad and tense/relaxed mood pairs, Arousal comprised the alert/drowsy and energetic/tired pairs, and Present Mindedness comprised the interested/bored, involved/detached, and lonely/connected pairs. In the present study, the Arousal factor was used as a measure of subjective arousal during INMI experiences and the Positivity factor was used as a measure of self-rated emotional valence.

4.2.6.2 Tapping Data Analysis

To isolate individual tapping episodes, each participant's movement data was viewed within the Data Analysis feature of the GENEActiv software. Each episode was located using the time and date reported in the diary booklet, with the button press as a marker of the endpoint for the episode. The start of a tapping episode was detected by examining the 2 min preceding the button press and locating the onset of a sequence of successive acceleration peaks corresponding to repetitive tapping. Once an episode was isolated, it was extracted and saved for further analysis. No discernable corresponding tapping sequence was found in the accelerometer data for 10 of the episodes reported in the diaries (3.64% of the 275 reported episodes). This could be due to a variety of different reasons, such as the participant forgetting to tap, writing down the incorrect time in the diary, or not tapping a regular or discernable beat pattern.

Next, each INMI episode was analysed with a tap detection algorithm in MATLAB, using similar methods to previous tapping research (Benoit et al., 2014; see Figure 4.3). The magnitude of the acceleration vector was computed as the square root of the sum of the three squared acceleration signals (x, y, and z). The resulting signal was smoothed using three passes of a running average filter in order to remove high frequency noise, and local maxima detection[13] was performed on the smoothed signal. Detected maxima were considered to be tap onsets if their absolute height was higher than a threshold; this threshold was defined as a ratio in relation to the highest maximum for the current tap sequence. The default threshold was set to 0.4, but was adjusted manually for each episode due to the fact that tapping strength and patterns varied greatly between and even within participants.

---

[13] Both of these steps were implemented using the lmax function, available at http://www.mathworks.com/matlabcentral/fileexchange/3170-local-min-max-nearest-neighbour/content/lmax.m.

For each file, the resultant tap series was then processed using the following steps. The first 10 taps were excluded from analysis, in line with previous tapping literature (e.g., Benoit et al., 2014; Sowiński & Dalla Bella, 2013; Zelaznik & Rosenbaum, 2010), and all numerical measurements were calculated based on the remaining taps. If there were fewer than 10 taps remaining after excluding the first 10 taps, this was recorded as a missing value, as the tapping period was deemed too short to extract a reliable tempo estimate. Overall, 30 INMI episodes were excluded on this basis (10.91% of the total data).

Next, the time series of inter-tap intervals (ITI) was calculated by computing the difference between all successive tap onsets. This ITI series was then processed further to remove artefacts and outliers (similar to the procedure used in Benoit et al., 2014 and Study 3 of this thesis). Artefacts were defined as two taps that were registered in brief succession; these can originate from rebounds (e.g., two fingers or two parts of the wrist/hand hitting the tapping surface in brief succession) or signal glitches. In the present study, artefacts were defined as ITI values of less than 100 ms. Outliers were defined as ITI values greater than three times the interquartile range from the median value of the ITI series; such values generally indicate the presence of missing taps. Overall, the average percentage of outliers (outliers divided by total number of taps) across all usable tapping sequences was 2.6%, and 55.7% of the usable tapping sequences contained no outliers at all. Using the artefact- and outlier-free ITI series, an average ITI value, coefficient of variation (CV; a normalized measure of tapping variability defined as the standard deviation of the ITI series divided by the mean ITI value), and tempo in beats per minute (bpm) were outputted from each episode for further analysis.

Finally, all remaining tapping episodes were visually inspected in a graphical format in MATLAB, as in Figure 4.3. In this visual inspection stage, 6 more episodes (2.18% of the total data) were excluded on the basis of comprising a noisy signal without clearly discernable tapping

peaks and one episode (0.36% of the total data) was excluded on the basis of a participant apparently halving the tempo in the middle of a tapping episode. Following all exclusion steps, 228 INMI episodes remained with usable tempo data (82.91% of the total reported INMI episodes).
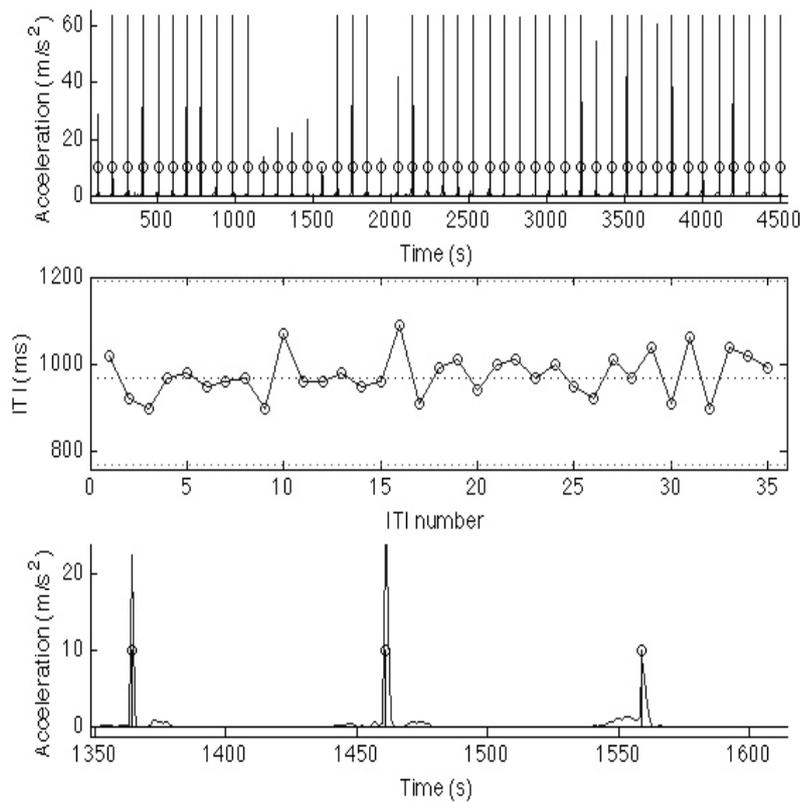
Figure 4.3. Graphical examples (from top to bottom) of 1) accelerometer movement data (minus the first 10 excluded taps; circles denote each local maximum), 2) series of corresponding inter-tap intervals, and 3) three individual taps from graph 1 (enlarged for clarity).

4.2.6.3 Mode Data Analysis

Two musicians who were naïve to the purposes of the study were recruited to independently code the musical mode (i.e., major or minor) of each song reported as INMI in the diaries. The coders followed a protocol resembling that of Schellenberg and von Scheve (2012), who also hand-

coded the mode of pop songs. The coders were required to find a recording of each INMI tune, listen specifically to the section of the song reported as INMI by the participant, and code the mode as major, minor, or ambiguous.[14] The independent ratings of the two coders were then collated and examined for any discrepancies by the present author. For 25 INMI episodes, the participant provided insufficient information to determine the mode of the INMI tune (e.g., the participant didn't know the title of the song or named an entire symphony). For the remaining 250 INMI episodes, the mode of 203 episodes (81.2% of the remaining data) was coded identically by the two coders. Episodes that were not coded identically were excluded from further mode-related analyses on the basis of being tonally ambiguous. Of the 203 episodes that were coded identically, 160 were in major keys and 43 were in minor keys.

4.3 Results

Descriptive statistics related to the music experienced as INMI and the circumstances surrounding the experience will be reported first to provide context and opportunities for comparison to previous literature. Results pertaining to the two main research questions of the paper—regarding the precision of tempo recall within INMI and the relationships between INMI and affective state—will then be reported.

---

[14] In the case of songs with no definitive, canonical version, such as classical music and Christmas songs, the coders were asked to find one standard recording, as it was not presumed that the mode of a piece would change between different recorded versions.

4.3.1 The INMI Experience: Descriptive Statistics

Of the 275 INMI episodes reported in the diaries, the number of episodes reported per participant during the full 4-day period ranged from 7 to 32 episodes. The number of INMI episodes reported during one day ranged from 0 to 10. The median number of episodes reported per participant was 16, or approximately 4 episodes per day (median episodes per day = 3.5).

As reported in Section 4.2.6.2, a discernable corresponding tapping sequence was found in the accelerometer data for 265 of the 275 reported episodes. As the co-occurrence of both a tapping sequence and a diary entry provide strong evidence that an INMI experience actually occurred at the time it was reported, the following descriptive statistics will be reported based on these 265 episodes.

In total, 182 unique songs were reported as INMI. For the vast majority of episodes, the song title and performer were reported, indicating that the songs were familiar to participants; however, for 11 episodes both the title and performer field were left blank. Two participants reported experiencing self-composed music as INMI for one and two episodes respectively. Other reported songs comprised a mix of pop, classical, rock, rap, jazz, folk, musical theatre, Christmas, TV, and children's music. Only one song ("Barbie Girl" by Aqua) was reported by two different participants; other repetitions of INMI songs occurred only within the same participant. A total of 42 songs were reported at least twice by the same participant, 15 songs were reported 3 or more times, and 2 songs were reported 6 times. Of the 42 songs occurring as INMI at least twice, 19 of these songs had recurrences on two separate days.

To investigate the frequency of INMI during different times of the day, the dataset was broken into bins of 3-hour periods. Less frequent INMI were reported in the early morning and late night hours. However, between the hours of 9:00 and 21:00, when all participants were most likely awake

and engaged in the study, the number of INMI episodes was quite stable across each 3-hour bin (see Figure 4.4).
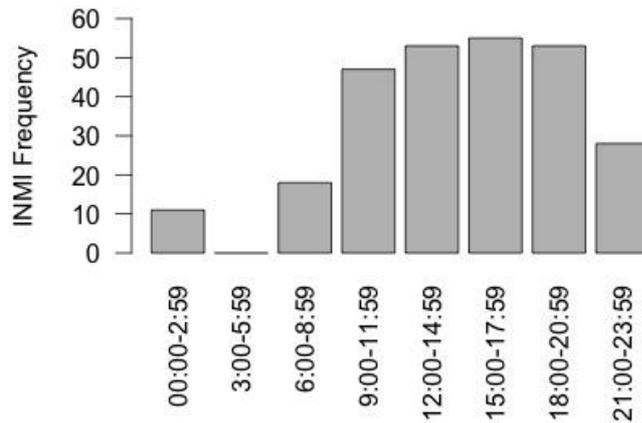
Figure 4.4. Frequency distribution of INMI at different times of day.

Participants also reported on how long it had been since they had heard the song experienced as INMI played aloud, e.g., on an iPod, radio, live performance, etc. For 16.2% of episodes, the song experienced as INMI had been heard less than 1 hour ago, and for 23.4% of episodes the song had been heard less than 3 hours ago. However, for 40.0% of episodes, participants reported that they had not heard a recording or performance of the song experienced as INMI in over one week. Figure 4.5 presents a summary of the results for this variable.
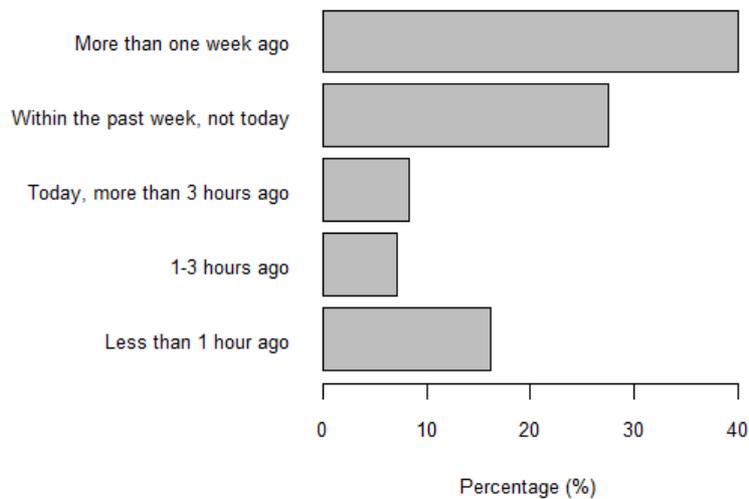
Figure 4.5. Distribution of how recently songs experienced as INMI had been heard aloud.

A total of 67 INMI episodes (25.3% of episodes) were reported to have occurred while a participant was engaged in a repetitive movement. An open question in the diary also probed the type of repetitive movement that the participant was engaged in. The majority of these reports of INMI during repetitive movements occurred while walking (57 episodes) and 3 episodes occurred while typing. The remaining repetitive movements, which comprised a single report each, were: brushing teeth, climbing, cutting vegetables, cycling, dying hair, washing dishes, and washing hair.

Finally, participants were asked to report any reasons they thought a song might have occurred as INMI. The distribution for this variable is visualized in Figure 4.6. Recent exposure to a song was named as a likely trigger for 40.4% of INMI episodes. Association with an environmental trigger, such as a person, word, or sound was attributed as a potential cause in 15.1% of episodes. For 27.6% of episodes, participants reported, "I have no idea why this tune came into my head," in the absence of any other trigger. Participants were also invited to record additional reasons an INMI episode might have occurred in

response to an open question (if they ticked the box "Other"). "Other" reasons were reported for 16.6% of episodes, however a substantial proportion of these appeared to be instances where participants were providing additional details that still fit within existing categories. For instance, one participant reported "memory of the TV show" as a trigger of an INMI episode, which could potentially be classified in the category "The song reminds me of something from my past". Another participant wrote "Aphex Twin is to release a new album on Sept. 22", which could be classified as "I was thinking about a future event and related it to this song". Additional recurring triggers reported in the "Other" category that were *not* captured by the existing categories relate to features of a melody (N = 3; e.g., "it's a nice melody") and mood states (N = 10; e.g., "maybe because I feel a bit tired, sitting at my desk makes me relax (it's a slow song)").
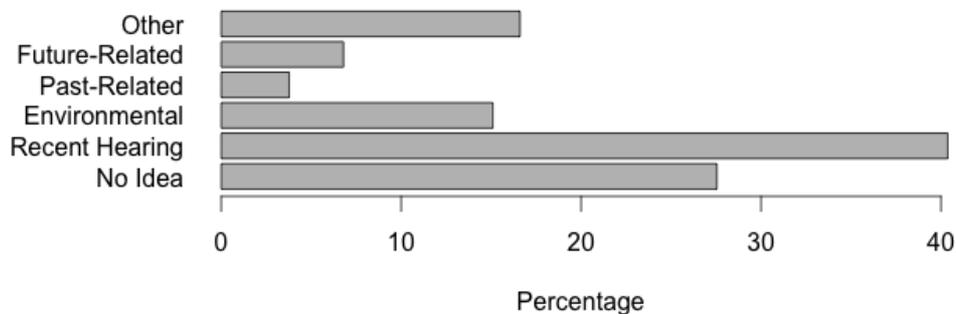
Figure 4.6. Percentages of INMI episodes for which each trigger was reported. (Note: As multiple triggers could be reported for each episode, these percentages total to over 100%).

4.3.2 The Tempo of INMI: Descriptive Statistics

Overall, 228 INMI episodes had usable tempo data (see Section 4.2.6.2 in *Method* section for data exclusion criteria). The number of total taps in each usable sequence ranged from 20 to 121 taps, and the duration of

the tapping sequences ranged from 7.6 to 92.4 s (M = 23.6, SD = 10.1). These episodes ranged in tempo from 42.0 bpm to 196.5 bpm. The mean tempo across all INMI episodes was 100.9 bpm (SD = 29.9; median = 98.5). The mean CV (coefficient of variation) of tapping across all 228 episodes was 0.06 (SD = 0.02; range = 0.02-0.13). The tempo distribution across all INMI episodes is shown in Figure 4.7.
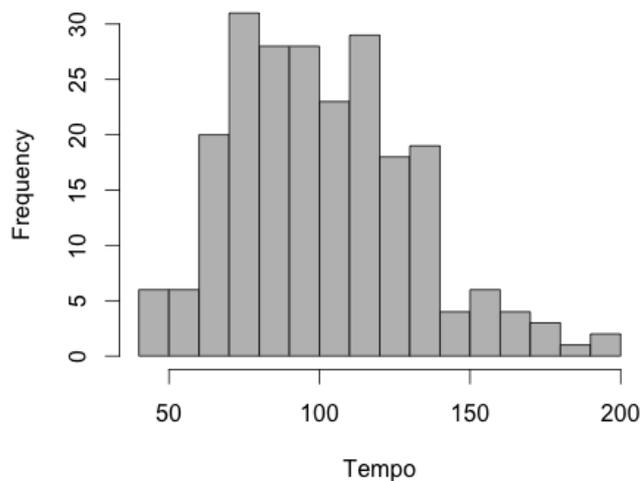
Figure 4.7. Tempo distribution of all 228 INMI episodes with usable tempo data.

The median tempi for INMI episodes at different times of the day are reported in Table 4.1. These data suggest a tendency to experience INMI that are somewhat slower between the waking hours from 6:00 to 8:59 than INMI that are experienced in the midday hours. The 8 INMI reported between 00:00 and 2:59 displayed the fastest median tempo of these time windows. However, it should be noted that the relatively low number of INMI episodes reported during the earliest and latest hours of the day makes it difficult to assess whether these differences in tempo at different time points are systematic or simply related to sample size differences.

Table 4.1: *Median INMI Tempo by Time of Day*

Hours	Total INMI with Usable Tempo Data	Median Tempo (bpm)
00:00-2:59	8	109.3
3:00-5:59	0	NA
6:00-8:59	17	80.1
9:00-11:59	44	93.1
12:00-14:59	46	108.22
15:00-17:59	43	100.2
18:00-20:59	44	98.8
21:00-23:59	26	88.72

4.3.3 Memory for INMI Tempo: Veridicality and Consistency

The next aim of the study was to investigate the veridicality of tempo recall within INMI, in comparison to the original version of each reported INMI tune. Of the 228 episodes with usable tempo data, 132 of these comprised INMI experiences of pieces of music that exist in a canonical version. Canonical songs are defined here as those that exist in one standard, recorded version. Examples of non-canonical songs include most Christmas songs and classical music, for which recordings might exist but no one "definitive version" is apparent from which to obtain tempo information. The original, recorded tempo for the section of the tune reported as INMI for each canonical song in the present dataset was obtained from two databases—songbpm.com and the Geerdes MIDI music database—and was also manually checked with a metronome to confirm the accuracy of the tempo information provided by the databases and to resolve any discrepancies between the two databases.

The tapped tempo of each of these 132 INMI episodes was then compared to the original, recorded tempo of the song by examining 1) ratios of the tapped to recorded tempo, 2) absolute deviations (as

percentages) of the tapped tempo from the recorded tempo, and 3) the correlation between tapped and recorded tempo across all episodes. A ratio of 1 or an absolute deviation of 0% for a particular episode would indicate that a participant tapped at the same tempo as the recorded version of the song. Figure 4.8 presents the original versus the tapped tempo for all 132 episodes comprising canonical songs.
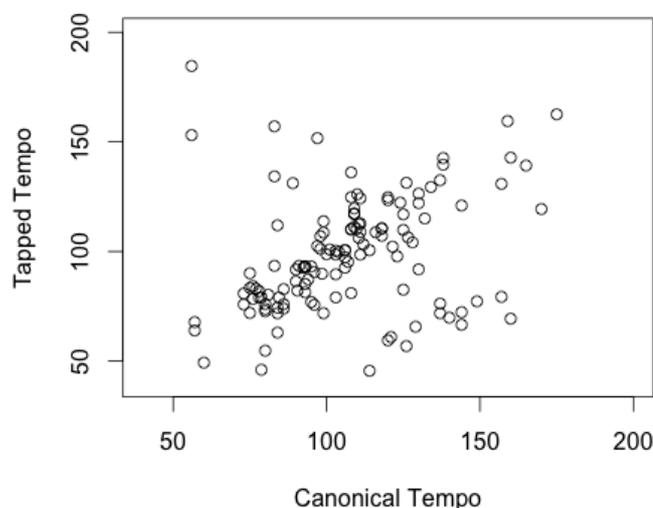
Figure 4.8. Original, recorded tempo for each of the 132 songs that exist in canonical versions plotted against the tempo each song was tapped at when experienced as INMI.

Some extreme ratios of the tapped to recorded tempi likely represented participants halving or doubling the tempo of a song, i.e., tapping at a different metrical subdivision than expected. In accordance with other previous research (Halpern, 1988b and Studies 2 and 3 of this thesis), episodes where the ratio of the tapped to recorded tempo was 1.9 or greater or 0.6 or less were omitted, given the likelihood of participants having doubled or halved the song tempo. This resulted in the exclusion of 17 episodes, leaving 115 episodes for further analysis.

The mean ratio of tapped to recorded tempo for these 115 episodes was 0.98 (*SD* = 0.15; median = 0.97) and the mean absolute deviation of the

tapped tempo from the original, recorded tempo was 10.8% (SD = 10.8%; median = 7.9%). A highly significant correlation was also found between the tapped and recorded tempi, r = 0.77, p < .001. Overall, 59.1% of songs were recalled within 10% of the original, recorded tempo and 77.4% of songs were recalled within 15% of the original tempo.

It is plausible that recent hearing of a melody might influence the above results, i.e., that the quite veridical reproduction of tempo within INMI might be explained solely by recent exposure. A Wilcoxon rank-sum test was employed (due to non-normal data distributions) to compare the absolute deviation from the recorded tempo for INMI tunes heard within the past week (N = 64) to the absolute deviation from the recorded tempo for tunes heard over one week ago (N = 51). The result of the test was non-significant, W = 1399, p = .19. A more stringent criterion was also applied to compare only tunes heard within the past day (N = 32) to tunes heard over one week ago (N = 51); this result was also non-significant, W = 754, p = .57.

A final question into the veridicality of INMI was related to previous research on voluntary musical imagery. Halpern (1988b) found a regression to the mean for the tempo of voluntarily imagined songs, i.e., slow-tempo songs tended to be recalled faster than the original tempo and fast-tempo songs tended to be recalled slower than the original tempo. The present data allowed for the first exploration into this question in INMI. An independent-samples t-test was performed to compare the original, recorded tempo between songs that were recalled slower than the recorded tempo within INMI (those songs with ratios of tapped tempo to recorded tempo of less than 1) and songs recalled faster than the recorded tempo (songs with ratios greater than 1). The mean recorded tempo for songs that were recalled slower than the original tempo was significantly faster than that of songs that were recalled faster than the original tempo, t(107) = 2.71, p = .01, thus suggesting regression to the mean within INMI.

Finally, the consistency of tempo recall for recurrences of the same tune as INMI was examined using the eight songs with usable tempo data that were experienced by the same participant at least three times. It should be noted that, unlike in the veridicality analyses, this sample of eight songs included both canonical and non-canonical songs. The mean tempo difference between the slowest and fastest version of a song experienced as INMI was 19.6% (SD = 14.0%; median = 14.6%). When comparing the slowest and fastest rendition of each song, two of these eight songs differed in tempo by over 40%. The remaining six songs differed in tempo by less than 20%, and five songs differed by less than 15% (see Table 4.2).

Table 4.2: *Consistency of Tempi for Songs Experienced as INMI at Least Three Times*

Song	Number of INMI Episodes	Slowest Tempo (bpm)	Fastest Tempo (bpm)	Difference Between Slowest and Fastest Version (%)
A Sky Full of Stars	3	82.5	116.9	41.7
Miss You	3	98.5	112.7	14.4
Ponta de Areia	3	63.0	74.3	18.0
Spirited Away One Summer's Day	3	78.6	82.6	5.1
One	4	81.4	93.2	14.4
You're So Vain	5	111.2	119.6	7.5
For Unto Us a Child is Born	6	74.6	105.0	40.6
Non Voglio Cantare	6	117.4	134.8	14.8

4.3.4 Musical Features of INMI and Affective States

The second main research question of this study aimed to investigate the relationship between participants' affective states, as recorded in the diary booklets, and specific musical features of their concurrent INMI. Specifically, 1) a positive relationship was predicted between the Arousal dimension of the mood scale and the tempo of INMI and 2) the Positivity dimension of participants' mood ratings was predicted to be higher during INMI in the major versus the minor mode, in line with previous findings from research on perceived music (e.g., Gagnon & Peretz, 2003; Husain et al., 2002; Webster & Weir, 2005).

The correlations between INMI tempo, INMI mode, Arousal, and Positivity are reported in Table 4.3. Point-biserial correlations were calculated for the INMI mode variable due to its dichotomous nature; all other reported correlations are Pearson's correlations.

Table 4.3: *Correlations of INMI Musical Features and Mood Variables*

	Tempo	Mode	Arousal	Positivity
Tempo	1.00			
Mode	-.003	1.00		
Arousal	.14*	.10	1.00	
Positivity	.15*	-.09	.07	1.00

Note: * signifies a significant correlation at the level of $p < .05$. Coding for the INMI mode variable is: 1 = minor, 2 = major.

Next, a linear mixed effect models was fitted with both Arousal and Positivity as predictors of INMI tempo. This analysis made use of all 228 INMI episodes with usable tempo data. A mixed effects model was employed in order to take account of the individual variations among participants and the multiple observations recorded from each participant by including "Participant" as a random effect in the model. In this model,

Arousal was a significant positive predictor of INMI tempo and no significant relationship was found between Positivity and INMI tempo. It should be noted that this non-significant relationship between Positivity and INMI tempo emerged despite an initial significant correlation between these two variables when examining all INMI episodes together (disregarding that there were multiple observations taken from each participant; see Table 4.3). This indicates that the initial significant correlation is actually driven primarily by specific individuals. Thus, when individual variations are taken into account by allowing a random effect in the model for "Participant", the initial significant relationship disappears.

The previously non-significant effect of Positivity was then removed and the model was refitted with only Arousal as a predictor of INMI tempo. Arousal was again a significant predictor and the reduced model provided a better fit to the data—based on the Bayesian Information Criterion (BIC) —than the full model with both mood variables included as predictors (BIC of full model = 2163.98, BIC of Arousal-only model = 2160.62). As standard R^2 measures for mixed effects models are not presently available, a pseudo-R^2 value was calculated for the effect of arousal on INMI tempo by squaring the correlation between the INMI tempo values predicted from the mixed effects model and the observed values of INMI tempo. The pseudo-R^2 value obtained was 0.34. Results of this analysis are presented in Table 4.4.

Table 4.4: *Linear Mixed Effects Models with Mood Variables as Predictors of INMI Tempo*

Model 1: Arousal and Positivity as Predictors of INMI Tempo

	Coefficient	S.E.	*t*-value	*p*-value
Intercept	81.24	12.99	6.26	.00
Arousal	1.87	0.67	2.78	.01*
Positivity	0.08	1.11	0.08	.94

Note: * signifies a significant predictor at the level of *p* < .05.

Model 2: Arousal as a Predictor of INMI Tempo

	Coefficient	S.E.	*t*-value	*p*-value
Intercept	82.06	7.23	11.35	.00
Arousal	1.87	0.67	2.79	.01*

Note: * signifies a significant predictor at the level of *p* < .05.

A second mixed effects analysis was conducted with Arousal and Positivity as predictors of INMI mode and "Participant" as a random effect. This analysis was conducted using the 203 INMI episodes for which musical mode data was ascertained by the two independent raters, minus two episodes for which participants did not complete the mood ratings, resulting in a total of 201 episodes for analysis. These 201 episodes comprised 42 INMI episodes in the minor mode and 159 episodes in the major mode. A binomial mixed effects model was fitted for this analysis due to the binary nature of the INMI mode variable. Neither of the mood variables were significant predictors of INMI mode. These results are presented in Table 4.5.

Table 4.5: *Binomial Mixed Effects Model with Mood Variables as Predictors of INMI Mode*

	Coefficient	S.E.	*z*-value	*p*-value
Intercept	1.47	1.43	1.03	.30
Arousal	0.12	0.07	1.67	.10
Positivity	-0.10	0.13	-0.76	.45

Note: Coding for the INMI mode variable is: 0 = minor, 1 = major.

A final exploratory analysis tested whether just the Happy/Sad dimension of the Positivity mood factor might relate to INMI mode, as "happy-sad" judgements have been used in previous research on musical mode in perceived melodies (e.g., Gagnon & Peretz, 2003). A binomial mixed effects model was fitted with Happy/Sad ratings as a predictor of INMI mode. Again, no significant effect was found of these mood ratings on INMI mode (see Table 4.6).

Table 4.6: *Binomial Mixed Effects Model with Happy/Sad Ratings as a Predictor of INMI Mode*

	Coefficient	S.E.	*z*-value	*p*-value
Intercept	0.98	1.05	0.93	.35
Happy/Sad	0.11	0.20	0.56	.57

Note: The Happy/Sad mood rating is scored such that higher scores indicate happier mood. Coding for the INMI mode variable is: 0 = minor, 1 = major.

4.4 Discussion

The present study has contributed a number of novel results, demonstrating that the combination of naturalistic diary methods with a quantitative measurement device—in this case an accelerometer—can add a new dimension to research on ephemeral phenomena such as INMI. These data represent, to the author's knowledge, the first attempt towards an objective marker related to the occurrence and tempo of INMI during everyday life. Despite the lesser degree of experimental control over the research environment as compared to a laboratory setting, over 80% of the acquired tapping data was usable for analysis. A wide range of INMI tempi, from approximately 40 to 200 bpm, provided a rich source of data that suggest a wide variety of personal inner music experiences.

The four main hypotheses of this study, as presented at the beginning of the chapter were as follows:

- Hypothesis 1: The tempo of INMI for canonical songs will be recalled highly veridically from memory.
- Hypothesis 2: The tempo of INMI will be recalled highly consistently from memory (although the number of instances of recurring INMI episodes within the same participant is expected to be fairly low).
- Hypothesis 3: INMI tempo will be positively related to concurrent ratings of arousal.
- Hypothesis 4: Ratings of emotional valence will be more positive during INMI in the major mode than minor mode INMI.

The present results provide support for both Hypotheses 1 and 3. The following sections offer some interpretation of these significant results and possible explanations as to why evidence in support of Hypotheses 2 and 4 remains somewhat lacking.

4.4.1 Research Question 1: The Temporal Precision of Memory Representations in INMI

INMI for music that exists in a canonical version was generally experienced at a tempo quite close to the original, recorded tempo of the song. In the present study, the mean absolute deviation from the original tempo for INMI experiences of canonical songs was 10.8%, and 77.4% of these songs were recalled within 15% of the original tempo. A highly significant correlation was also found between tapped and original, recorded tempo. Overall, this finding of high temporal veridicality within INMI is particularly striking given that 1) the songs reported as INMI were recalled spontaneously with no instruction for veridical recall (in contrast to Study 3 of this thesis) and 2) the INMI occurred within the context of the external distractions of everyday life (in contrast to a controlled laboratory setting).

For comparison, in Study 3 of this thesis, the mean absolute deviation from the original, recorded tempo for songs that were deliberately imagined in a tapping task (the Imagery (motor) task) was 18.4%. Additionally, Levitin and Cook (1996) reported that 72% of their participants' trials in a sung recall paradigm were within 8% of the original, recorded tempo. These figures suggest that tempo representations within INMI may be equally or potentially even more veridical than those within musical imagery that is deliberately imagined within a laboratory context, although both forms of imagery appear to be less temporally veridical than songs produced in a sung recall paradigm (Levitin & Cook, 1996). The superior performance in Levitin and Cook's study might be partially explained by the presence of auditory feedback within singing, which may have aided participants in refining their tempo judgements during the course of their sung productions. Additionally, as Levitin and Cook asked participants to sing only two self-selected songs that they knew very well, familiarity or overlearning may have played a role in the higher level of veridical recall. The same argument might also explain the difference in results obtained between the present study and

Study 3 of this thesis: participants in the present study may have been more familiar with the songs they spontaneously experienced as INMI than participants in Study 3 were with the experimenter-selected songs that they were asked to imagine in the laboratory. This explanation cannot be directly tested at present and further research is needed in which tempo recall for involuntarily and voluntarily retrieved musical memories are compared directly within the same participants, in order to control for individual differences in familiarity, as well as factors such as musical training. Nevertheless, as they stand, these findings suggest several parallels between voluntarily and involuntarily recalled musical memories, as well as parallels to involuntary autobiographical memories, which tend to be as or more specific and detailed in comparison to voluntary autobiographical memories (Berntsen, 1998; Mace, 2006; Schlagman & Kvavilashvili, 2008).

In 40% of INMI episodes in the present study, participants reported that they had not heard the song that was experienced as INMI in over one week. However, INMI for such songs was not experienced at a less veridical tempo than INMI for songs that had been heard more recently. This suggests that INMI that are recalled from long-term memory are temporally precise, and that the overall finding of high veridicality for INMI tempo is not explained solely by those episodes for which an INMI song was heard minutes ago on the radio and might still be held in short-term memory.

Evidence was also found for a regression to the mean for INMI tempo, such that faster songs tended to be recalled slower than their original tempi and slower songs tended to be recalled faster than their original tempi. This parallels previous findings on tempo for voluntary musical imagery (Halpern, 1988b). Taken together, these findings suggest that both voluntarily and involuntarily imagined music tends to occupy a narrower range of tempi than analogous perceptual experiences of the same songs (in terms of hearing a recording or adjusting the tempo of a song to one's preferred perceptual tempo). This regression to the mean

also occurs despite the fact that participants are deliberately able to imagine songs at fairly extreme slow and fast tempi when specifically instructed to do so (Halpern, 1988b, Study 2). Further research should be conducted to explore the mechanisms underlying this regression toward a mid-range tempo in both spontaneous and deliberate musical imagery, and whether this tendency is related to factors such as one's natural, spontaneous tapping rate or preferred perceptual tempo (e.g., McAuley et al., 2006).

Finally, an exploratory analysis was conducted to examine the temporal consistency for the eight INMI tunes with usable tempo data that were reported at least three times. The majority of these INMI tunes differed in tempo between the slowest and fastest rendition by less than 20%, however two songs varied in tempo by approximately 40%. In regard to one of these less consistent songs ("For Unto Us a Child is Born"), the participant who reported this song informed the experimenter in the post-experiment debrief session that he was a pianist who often practiced musical pieces at different tempi when learning them, and that this seemed to influence his subsequent INMI tempi. This is just one example of a variety of reasons that might affect the consistency of INMI tempi that could not be accounted for in the present study. In future research, it may be fruitful to employ a design that aims specifically to examine the issue of temporal consistency. A possible design might involve priming participants with a recording of a song for which only one version exists and then asking participants to record the tempo of all subsequent INMI episodes related to that particular song. Controlling the exposure phase of the INMI tune would consequently provide more control over the version of the song that comes to participants' minds as INMI.

One potential limitation of the present research design is that, as the study was completed outside of the experimenter's supervision, participants could not be prevented from voluntarily manipulating the tempo of their INMI. However, the present study aimed to combat this issue in several ways: 1) by providing a clear definition of "earworms" to

participants that emphasized their involuntary recall as a key feature, 2) by instructing participants specifically to "please tap the beat of the tune as closely as possible as to what you hear in your head," 3) by not requiring any sort of button press or marker *before* the tapping period, so as to make it easy to start tapping as soon as possible when an INMI episode began, and 4) by not revealing the purposes of the study until after it was completed. Indeed, when asked what they thought might be the purpose of the study in the debrief session, none of the present participants surmised that an aim of the study was to investigate the veridicality and/or consistency of INMI tempo.

4.4.2 Research Question 2: The Relationships Between INMI Features and Affective State

Examining the relationships between musical features of INMI and concurrent affective states allowed the present research to begin to unravel the complex question of how endogenous bodily and mental states may interact with spontaneous cognitions. The present results revealed a modest, yet significant, positive relationship between subjective arousal and INMI tempo. This parallels and extends findings on the relationship between musical tempo and arousal in the domain of music listening (Edworthy & Waring, 2006; Husain et al., 2002; North & Hargreaves, 2000) by demonstrating that even music that is 1) experienced only as imagery and 2) spontaneously generated displays this significant tempo-arousal relationship. The relationship between INMI tempo and subjective arousal suggests some first evidence that INMI might be functionally linked to mood regulation in a similar manner to perceived music. That is, in the absence of an iPod or other music-generating device, spontaneous imagery for music may be able to fill in as a mood regulatory mechanism. It is not possible to deduce from these results a direction of causality of this effect (i.e., whether INMI tempo influences arousal level or arousal level influences the tempo of a song within INMI). However, previous studies of perceived music suggest a bidirectional relationship, such that tempo can influence one's arousal

level (e.g., Husain et al., 2002) and arousal can influence one's tempo preferences (North & Hargreaves, 2000) or the tempo at which a piece of music sounds "right" (Study 1 of this thesis; Jakubowski et al., 2015). As such, a similar bidirectional relationship might exist between musical tempo recalled within INMI and concurrent arousal.

No significant relationship was found between INMI mode and positivity of mood. This was unexpected, given previous evidence of an association between emotional valence and musical mode (Gagnon & Peretz, 2003; Husain et al., 2002; Webster & Weir, 2005). One possible explanation for this is that musical mode may be a less prominent or distinctive feature of INMI, such that it is less discernable by participants than other musical features of the imagery. Another possible explanation concerns a key distinction made in the literature on music and emotions between the emotional connotation that one perceives to be portrayed by the music and the induction of actual felt emotion (e.g., Gabrielsson, 2002; Juslin & Laukka, 2004). The present study asked participants to rate felt emotions during INMI episodes using the mood rating scales, rather than to evaluate the emotions portrayed by the music experienced as INMI. Previous research has revealed that these two questions do not necessarily elicit the same response, as, for instance, many people experience feelings of pleasure and enjoyment when listening to music that expresses sadness (Huron, 2011; Vuoskoski, Thompson, McIlwain, & Eerola, 2012). As such, it may also be the case that this dissociation between perceived and felt emotions can be made within the INMI experience.

One reason that INMI tempo may have a more direct relationship to mood-related variables than INMI mode is that INMI tempo is more likely to influence motor responses or, conversely, to be influenced *by* movements that were occurring before the INMI episode began, thus making it more observable to one's conscious awareness. This idea is supported by evidence of a close relationship between musical beat perception and motor areas of the brain (Grahn & Brett, 2007; Grahn &

Rowe, 2009), as well as by the finding in the present study that approximately one quarter of INMI episodes occurred during a repetitive movement. The design of the present study, in which participants were asked to move to the beat of the music by tapping, may also have increased the salience of INMI tempo over other musical features.

The present findings provide a first glimpse of the interactions between musical features of INMI and concurrent mood state. Further research should be conducted to examine the relationships between affective state and other musical features of INMI (e.g., lyrical content, timbre, loudness) in order to gain a more complete understanding of the possible interactions between different types of music experienced as INMI and concurrent mood. Physiological measurements of arousal (e.g., electrodermal activity) could also be included in future research.

4.4.3 The INMI Experience: Descriptive Findings

Finally, several descriptive findings from the diary data corroborated and extended previous research. The diverse variety of songs reported (182 in total) and the lack of overlap in INMI tunes between participants is in line with several previous studies, suggesting that almost any song can become INMI and that the INMI experience is very personal and idiosyncratic (Beaman & Williams, 2010; Williamson & Müllensiefen, 2012). The phenomenon has also been found to be short-lived (Byron & Fowles, 2013), as affirmed by the relatively few songs reported three or more times as INMI within each participant.

The descriptive results revealed that the number of INMI episodes at different periods of the day was fairly stable. This supports some previous findings that INMI does not appear to be more frequent at certain times of the days than others (Bailes, 2007; Byron & Fowles, 2013). The results also provided some preliminary evidence that INMI tempo might vary depending on the time of day, although this question needs to be systematically addressed in future research with more data points per

participant and per time period. A future study that is able to capture more INMI episodes per participant per day could investigate individual profiles of INMI tempo fluctuations and investigate whether INMI tempo varies systematically in relation to one's activities, circadian rhythms, etc.

The present study also served as one of the first diary-based investigations into the triggers of the INMI experience. Recent exposure to music was the most commonly reported INMI trigger, which is in line with a retrospective questionnaire-based method employed by Williamson et al. (2012). However, it should be noted that, as the present study also included a question about how recently a participant had heard a song aloud, participants might have been somewhat primed towards reporting recent exposure to a song as an INMI trigger. The study also revealed that conscious awareness of an external or internal trigger for the experience was absent in about one quarter of INMI episodes, those in which a participant reported, "I have no idea why this tune came into my head". Additionally, 10 reports of an INMI episode being triggered in relation to a mood state were collected. In subsequent larger-scale studies, an investigation could be conducted into whether these specifically mood-triggered INMI episodes display a stronger relationship between mood ratings and the musical features of INMI than INMI triggered by other sources.

Finally, data on concurrent movements during INMI indicated that approximately one quarter of INMI episodes occurred during a repetitive movement, such as walking. These data provide some impetus for future accelerometer-based research, in which one could investigate how concurrent movement and INMI tempo might influence one another, e.g., if INMI tempo changes to match a movement or if a movement becomes more regular when it is made in time with INMI, by comparing acceleration patterns before and during INMI. Such research could provide valuable insights into the interactions between musical imagery and the sensorimotor system.

4.4.4 Conclusion

Study 4 has introduced a novel methodology that opens new avenues for research into dynamic aspects of INMI and other types of spontaneous and self-generated thoughts. This study specifically investigated tempo-related aspects of the INMI experience within the course of everyday life. The results demonstrated that INMI is often a highly veridical experience in terms of tempo, even when a song experienced as INMI has not recently been heard aloud. A significant positive relationship between subjective arousal and INMI tempo was also revealed, suggesting a first link between spontaneous musical imagery and mood that parallels findings in the perceived music domain. Research into temporal aspects of INMI and involuntary memories in general opens an array of important questions that will help to further our understanding of endogenous thought processes.

Chapter 5. Investigating Melodic Features of Involuntary Musical Imagery

Abstract

This chapter investigates the melodic commonalities of pop songs frequently reported as involuntary musical imagery (INMI). A preliminary study found effects of song popularity and recency (measured using UK Music Chart data) on the frequency with which a song was reported as INMI by participants in an online survey. In light of this finding, songs frequently named as INMI were matched in the main study in terms of popularity and recency to tunes that were never named as INMI. These two sets of matched tunes were then compared in terms of a large number of pitch and rhythmic statistical features using automated melodic feature extraction software and powerful statistical classification techniques. The commonness of the song contour, in relation to a large pop music corpus, and the tempo of a song were among the most predictive features of INMI tunes. Results are discussed in relation to existing literature on the INMI experience, musical memory, and melodic "catchiness".

5.1 Introduction

Following on from Study 4, which investigated one specific musical feature of INMI episodes as they occurred (the overall tempo of each episode), the research in the present chapter continues to explore the musical content of the INMI experience by examining the melodic features of songs that were frequently reported as INMI by participants in an online survey. Specifically, a detailed computational analysis of 83 pitch- and rhythm-related melodic features of songs frequently reported as INMI is conducted. This research examines these melodic features in conjunction with several extra-musical variables related to the popularity

of the tunes in order to provide new evidence for factors that contribute to increasing the propensity of a certain tune toward becoming INMI.

5.1.1 Relevant Research on INMI

Why do certain songs get stuck in our heads over others? This is a question that is regularly asked to music researchers, yet remains highly elusive. The reason this question is so difficult to answer may be that the likelihood of a tune becoming INMI is influenced by a wide variety of both intra- and extra-musical factors, from the musical features and lyrics of a song to the context in which it appears as INMI to previous personal associations with a song.

Several researchers have examined extra-musical features that increase the likelihood that a song will become INMI. For instance, evidence of a recency effect has been found in some laboratory-based studies, such that the song that has been heard aloud most recently is more likely to become INMI than a song heard less recently (Hyman et al., 2013; Liikkanen, 2012b). Recent exposure to a tune is also the most commonly reported trigger of INMI experiences in several previous studies (Bailes, 2015; Floridou & Müllensiefen, 2015; Williamson et al., 2012). Familiarity can also increase the likelihood that a song will become INMI. In a study by Byron and Fowles (2013), participants who were exposed to a previously unfamiliar song six times were more likely to experience that song as INMI than participants who had only heard the song twice. It is also generally uncommon to experience completely unfamiliar music as INMI, although a handful of reports of self-composed music have been found in previous work (Beaman & Williams, 2010; Beaty et al., 2013).

In terms of features of the actual music that increase a song's propensity toward becoming INMI, one study reported that a piece of music with lyrics was more likely to become INMI than a piece of music without lyrics (Floridou et al., 2014). However, since only two pieces of music were

used as stimuli in this particular experiment (one with and one without lyrics) it is difficult to tell whether this effect will generalize to all music.

In terms of the melodic content of INMI, a pilot study first presented by Finkel, Jilka, Williamson, Stewart, and Müllensiefen (2010) and further developed by Williamson and Müllensiefen (2012) represents the first empirical investigation in this realm. In this study, 29 songs were collated that had been frequently or recently experienced as INMI by more than one participant in an online survey. Then, 29 non-INMI songs (songs that had never been named as INMI in the online survey) that were similar to the 29 INMI songs (based on Gower's similarity coefficient; Gower, 1971) were compared to the INMI songs in terms of melodic features. Statistical melodic summary features of all 58 songs were computed using the melody analysis software FANTASTIC (Müllensiefen, 2009) and a binary logistic regression was used to predict INMI versus non-INMI tunes based on these features. The results of this analysis indicated that INMI tunes generally contained notes with longer durations and smaller pitch intervals than non-INMI tunes. Williamson and Müllensiefen (2012) suggest that these two features might make songs easier to sing along with, which connects to results reported earlier in the same paper— specifically, that people who sing more often experience more frequent and longer INMI. The present study builds on the initial findings of Finkel et al. (2010) and Williamson and Müllensiefen (2012) and extends this work by using 1) a larger sample of tunes, 2) a larger set of melodic features (including second-order, corpus-based features and m-type features in addition to simple summary features to describe a melody), and 3) more powerful statistical modelling techniques for both matching of the INMI tunes to non-INMI tunes and classifying INMI versus non-INMI tunes based on their features.

It should also be noted that several previous diary and questionnaire studies, including Study 4 of this thesis, have found a diverse variety of songs reported as INMI, with few repetitions of the same song across multiple participants (Beaman & Williams, 2010; Halpern & Bartlett, 2011;

Williamson & Müllensiefen, 2012). One might be tempted to conclude that this finding suggests that any song is equally susceptible to becoming INMI and that there are no underlying commonalities that can be derived across participants in terms of the musical content of their INMI experiences. However, it is quite plausible that the participants in these previous studies varied in their listening habits and musical interests, and that their highly idiosyncratic INMI experiences were a result of the different types of music they preferred and were exposed to during the course of their daily lives. This does not necessarily preclude the idea that there are some underlying melodic commonalities amongst this at first seemingly diverse sample of songs. In order to detect such commonalities, the present study employed intensive computational analyses based on a wide array of pitch- and rhythm-related melodic features alongside powerful statistical modelling techniques designed to extract and combine the most important variables from a large number of predictors.

5.1.2 Research on Musical Catchiness

A related area to the present line of research deals with the concept of musical "catchiness". This is a new area of empirical enquiry and, as such, scientific investigations are sparse. However, a large degree of anecdotal evidence has accumulated over many years suggesting that certain songs—or certain sections within the same song—are catchier than others. This concept has long been intertwined with the concept of musical "hooks". A song hook, as defined by a musician writing on the topic, is "that part of the song meant to catch the ear of the listener" (Covach, 1996). One of the first musicological investigations of hooks was conducted by Burns (1987), who compiled detailed qualitative descriptions of how hooks might be constructed using rhythmic, melodic, lyrical, timbral, temporal, dynamic, and recording-based features of a tune.

A currently ongoing research project in the Netherlands represents the first systematic attempt to quantify the concept of hooks and musical

catchiness. The authors of this research offer a definition of a hook as "the most salient, easiest-to-recall fragment of a piece of music" (Burgoyne et al., 2013, p. 1) and a definition of catchiness as "long-term musical salience, the degree to which a musical fragment remains memorable after a period of time" (Burgoyne et al., 2013, p. 1). This research project was distributed in the form of an Internet-based game called "Hooked". The premise of the study relied on participants judging as quickly as possible whether they recognized a song that was played aloud. Different structural sections of the same song were played on different trials. The faster a participant responded, the catchier that particular section of the song was presumed to be. Initial results from this project indicate that different sections even within the same song differ significantly in the amount of time required to recognize them, thus suggesting some sections serve as better hooks than others (Burgoyne et al., 2013). More recent work from this group has examined audio and symbolic musical features that predict the catchiness of a song section (Van Balen, Burgoyne, Bountouridis, Müllensiefen, & Veltkamp, in press). A principal component analysis revealed a number of features related to melodic repetitiveness, melodic "conventionality" in comparison to a corpus of pop music, and prominence of the vocal line as predictors of catchiness. The findings from this research suggest various potential relationships to INMI research. For example, if a section of a song is the easiest section to recognize in the "Hooked" game, it might also be the section that most easily comes to mind when *involuntarily* retrieved from memory. Additionally, the melodic features that were found to enhance long-term salience (i.e., catchiness) of a melody might also bear certain similarities to the features that enhance involuntary mental retrieval and rehearsal.

5.1.3 Research on Musical Features of Song Memorability

Another related area of research has examined the melodic features that enhance recognition or recall of tunes from memory. Müllensiefen and Halpern (2014) investigated the melodic features that aid in making a

tune memorable in terms of both explicit and implicit memory. They conducted a laboratory-based study in which participants heard 40 novel melodies in an encoding phase and were then assessed on their memory for these melodies in a subsequent recognition task that included the 40 previously heard melodies as well as 40 previously unheard melodies. Explicit memory was assessed by asking participants to judge if they had heard the melody before and implicit memory was assessed using pleasantness judgements for the melodies (based on the "mere exposure effect" assumption that preference for an item increases with increasing exposure to it; Zajonc, 1968).

The study used the same feature extraction software that will be used in the present research (FANTASTIC; Müllensiefen, 2009). A relevant feature of Müllensiefen and Halpern's study to the present work is that it made use of both first- and second-order melodic features. First-order features are features that are calculated based on the intrinsic content of a melody itself, such as the average note duration, average interval size, or pitch range of the melody. Second-order features, also called corpus-based features, are features that compare a melody to a larger collection or corpus of melodies (determined by the user, generally this corpus is comprised of music from the same genre or style as the melodies that are being analysed, such as pop songs or folk songs). For instance, one example of a second-order feature might measure whether the average interval size within a particular melody is smaller or larger than the average interval size of all melodies within the corpus. The use of second-order features allows one to determine whether particular features of a melody are highly common or highly distinctive in comparison to a corpus of music that is intended to be representative of the genre from which the melody is taken. Further details of the first- and second-order features employed in the present study will be presented in the *Method* section of Study 5b of this chapter.

Müllensiefen and Halpern (2014) conducted a number of detailed analyses of their data using partial least squares regression and found

somewhat different patterns of results for predicting explicit and implicit memory for tunes. Explicit memory was enhanced for tunes that included melodic motives that were rare in terms of their occurrence in the corpus and that repeated all motives frequently. In terms of implicit memory, the usage of unique motives in comparison to the corpus was also important, similar to the findings on explicit memory. However *less* repetition of motives, a smaller average interval size, simple contour, and complex rhythms were also important to implicit memory recognition.

Although this study is relevant to the present research, several differences are inherent. Müllensiefen and Halpern's work tested whether certain features of a melody can increase memorability for previously unfamiliar tunes that had only been heard *once* before, in terms of both explicit and implicit memory. In the case of INMI, however, tunes that are often *highly familiar* to participants (and have been heard aloud many times before) are retrieved in a *spontaneous* fashion from memory. Therefore, although it is plausible that some of these melodic features related to explicit and implicit memory for previously unfamiliar music might be implicated in INMI, it is also likely that other features might serve to enhance the spontaneous recall of well-known tunes and looping nature of the INMI experience.

Other studies have investigated the musical features that contribute to memory for melodies through the use of paradigms that seek to identify the point at which familiar songs are identified. Schulkind, Posner, and Rubin (2003) conducted such a study in which familiar songs were played to participants on a note-by-note basis. The positions in a song in which participants were most likely to identify the song correctly included notes located at phrase boundaries, notes that completed alternating sequences of rising and falling pitches, and metrically accented notes. Using a similar paradigm, Bailes (2010) explained around 85% of the variance in her participants' data with second-order features that measured timing distinctiveness and pitch distinctiveness in comparison to a large corpus of Western melodies. The results of these studies

indicate that assessing memory for melodies based on structural and melodic features can be useful in modelling aspects of music cognition and provide impetus for conducting similar research in the domain of *involuntarily* retrieved musical memories.

5.1.4 Research on Musical Features of Hit Songs

A final relevant body of literature has investigated the commercial success of songs, that is, whether certain musical features of a song predispose it toward becoming a "hit". This literature is sometimes referred to as "Hit Song Science".

One common approach in this research area has been to analyse the acoustic features from recordings of songs in an attempt to predict hits versus non-hits based on these features (Dhanaraj & Logan, 2005; Ni et al., 2011). However, the approach of predicting songs based solely on acoustic features has received some criticism, due to the generally low prediction accuracy rates that have been reported. For instance, Pachet and Roy (2008) conducted a large-scale analysis of a database of 32,000 songs and reported that classifiers built using state-of-the-art machine learning techniques were not able to significantly predict the commercial success of pop songs based on the songs' acoustic features. The researchers thus concluded that "Hit Song Science" is not yet a science, but that there is scope to uncover new features that more accurately relate to human aesthetic judgements, which may be more useful for uncovering commonalities in commercially successful songs.

An alternative approach that has been employed is to investigate features of the compositional structure of hit songs. Kopiez and Müllensiefen (2011) conducted a first exploration into this area by attempting to predict the commercial success of cover versions of songs from the Beatles' album "Revolver". They were able to accurately predict 100% of cases using a logistic regression model with just two melodic features as predictors—pitch range and pitch entropy—thereby indicating that

compositional features can be useful in predicting hit song potential. However, as the sample of songs used in this study (14 songs all composed by the same band) is very specific, it is unlikely that such a simple classifier would be able to cope with the wide diversity of styles and artists represented across all of the "popular music" genres. As such, a subsequent study was conducted by Frieler, Jakubowski, and Müllensiefen (in press) in order to investigate the compositional features contributing to the commercial success of a more diverse sample of 266 pop songs from a variety of artists and genres. The study used a wide range of first-order melodic features to predict hits versus non-hits. The three most predictive variables for hit songs all related to the interval content of the melodies, which indicated that 1) hit songs use all nine possible combinations of three contour directions (up, down, or repetition) between two subsequent intervals in more uniform combinations than non-hits, 2) the interval distribution within hit songs comprises a smaller number of different intervals than non-hits, and 3) hit songs were more uniform in terms of intervallic direction than non-hit songs. However, the classifier used in this work only achieved a classification accuracy rate of 52.6%. This finding suggests that extra-musical factors (such as artist popularity) and audio features (such as timbre) may play a large role in the commercial success of pop music, but also leaves open the question as to whether second-order, corpus-based features might help to further capture the unexplained variance in the data. Hence, the present research employed a similar approach for predicting INMI tunes, but included second-order features in addition to simple summary features.

5.1.5 Aims of the Research

To summarize, although some previous research has investigated the musical features of catchiness, memorability, and hit songs, only one small-scale study (Finkel et al., 2010; summarized in Williamson & Müllensiefen, 2012) has examined melodic features that might increase the likelihood of a tune to become INMI. It was hypothesized that the melodic features that served as predictors of INMI quality in the present

study might include similar features to those reported by Finkel et al. (2010) and Williamson and Müllensiefen (2012) —small pitch intervals and long note durations—as well as other features (including second-order features) that might enhance the "singability" of a tune, as proposed by Williamson and Müllensiefen. Features that enhance the memorability or long-term salience of a melody (the definition of catchiness provided by Burgoyne et al., 2013) might also serve to enhance the ease with which a melody is retrieved within the mind in an involuntary fashion. Additionally, hit songs and INMI might share some common melodic features. Songs that are commercially successful "hits" may also be likely to become stuck in the mind as INMI; in turn, frequent reappearances of a song as INMI might motivate the experiencer to seek out and replay the actual tune aloud (as reported in Williamson et al., 2014), thereby increasing the song's popularity in terms of downloads of the tune, YouTube plays, etc., in a cyclical fashion. Therefore, it was predicted that the musical features of INMI might bear some similarities to existing literature on the melodic features contributing to catchiness, memorability, and commercial success of a song. It was also predicted that additional melodic features to those revealed within this related literature might be implicated as predictors of INMI tunes, in order to account, for instance, for the looping of short sections of music within the mind that is a unique feature of the INMI experience.

The present work collated a large number of reports of participants' most recent and most frequent INMI tunes from a pre-existing questionnaire, in order to investigate aspects of their melodic features and popularity that contributed to their INMI propensity. The research will be reported in two parts. The first part of the research (Study 5a) was a preliminary study that investigated the extent to which the number of times a song was named as an INMI tune in the questionnaire could be explained by the song's popularity and recency, as measured using data from the UK Music Charts. As recency and familiarity effects have been demonstrated in previous studies of INMI (Byron & Fowles, 2013; Hyman et al., 2013; Liikkanen, 2012b), it was hypothesized that variables that described a

song in terms of its recency and popularity could be used to predict the number of times a song had been named as an INMI tune. Building on the findings from this preliminary study, the second part of the research (Study 5b) examined the extent to which the propensity of a song to become INMI could be explained by melodic features of the song, while controlling for popularity- and recency-related variables.

Study 5a: Modelling Reports of INMI Tunes Based on Popularity and Recency

5.2 Method

5.2.1 Participants

A database of commonly reported INMI tunes was compiled using responses to a pre-existing online questionnaire on INMI experiences ("The Earwormery", Williamson et al., 2012). The questionnaire had been completed by 5,989 participants in total. For the present project, responses regarding participants' most recent and most frequent INMI tunes were compiled from the first 3,000 of these participants.[15] These 3,000 participants ranged in age from 12 to 81 years ($M = 35.9$, $SD = 13.1$). Of these participants, 1,338 were male, 1,644 were female, and 18 did not provide gender information.

5.2.2 Ethics Statement

The two studies within this chapter, as well as the online "Earwormery" questionnaire from which the data were collected, were approved by the Ethics Committee of Goldsmiths, University of London.

---

[15] Responses to only the first 3,000 participants were compiled as manual verification of each response was required to ensure the response was indeed an existing song and that artist names and song titles were spelled correctly. Due to the time-consuming nature of this manual verification step, only the first 3,000 entries in the online survey could be processed within the time scale of this project.

5.2.3 Materials and Data Preparation

The online "Earwormery" questionnaire contained questions about the features and phenomenology of participants' INMI experiences, such as how often they experienced INMI, whether they found INMI disturbing or distracting, and their reactions and attitudes toward INMI. The survey questions relevant to the current project were two open-ended questions: one that asked for the name, artist, and section of the tune (e.g., chorus, verse) experienced as the participant's most recent INMI tune, and one that asked for the name, artist, and section of the participant's most frequent INMI tune.[16] Data related to additional aspects of this online questionnaire are reported in various other existing papers (Williamson et al., 2012, 2014; Williamson & Müllensiefen, 2012).

The first step in preparing the data for the present analysis involved compiling participant responses to the two relevant survey questions. Not all participants responded to these questions, and some only responded to one of the two questions. The analysis in the present project did not distinguish between tunes listed as a "most frequent" and a "most recent" INMI tune, as the aim was to include as many different pieces of music in one sample as possible. Some participants provided incomplete information, such as an artist without a song title, and this information was not included in the data unless the missing information could be readily reconstructed (e.g., a participant gave an artist and a line of lyrics, from which the song title could be deduced via Internet search). As a result of the design of the current study, songs in the sample were limited to only "popular music" genres (e.g., pop, rock, rap, rhythm & blues, etc.), while excluding such music types as classical, children's songs, and TV jingles. This is due to the fact that the project utilized popularity and recency variables for each of the songs that were measured in terms of

[16] Within the actual online survey, the term *earworm* was used in all instructions and questions directed to participants, rather than *INMI,* as *earworm* was deemed a more familiar and colloquial term.

data obtained from the UK Music Charts. Hence, only songs that may have appeared in the music charts were included within the present sample.

Overall, the data compilation process for the 3,000 participants resulted in 3,806 responses. In total, 2,817 different songs were named, with 2,311 songs being named only once, and 506 being named more than once. This small overlap amongst songs named more than once provides evidence for striking individual differences in terms of the INMI experience, which is in accordance with previous studies that have also reported a small degree of overlap in INMI tunes reported across multiple participants (e.g., Beaman & Williams, 2010; Williamson & Müllensiefen, 2012).

For each of the 2,817 songs in the sample, information on the song's popularity and recency was acquired. These variables were measured in terms of the number of weeks the song had spent in the UK Music Charts (popularity measure), the highest position the song had attained in the charts (popularity measure), and the date the song had exited the charts (recency measure). This information was acquired via the UK Music Charts database at polyhex.com. The exit date variable was then converted from a date to the number of days since exiting the charts for use in subsequent analyses. This number of days was calculated from February 22, 2013—the end date of data collection for the present project. Not all songs named by participants in the online questionnaire had been in the charts; the resulting number of songs from the original sample that had been listed in the charts was 1,558. From this list of 1,558 songs, the most frequently mentioned song was named 33 times as an INMI tune (Lady Gaga's "Bad Romance"). The nine most frequently named INMI tunes are listed in Table 5.1.

Table 5.1: *Songs Most Frequently Named as INMI*

Song Title & Artist	# of Times Named as INMI
1) Bad Romance- Lady Gaga	33
2) Can't Get You Out of My Head- Kylie Minogue	24
3) Don't Stop Believing- Journey	21
4) Somebody That I Used to Know- Gotye	19
5) Moves Like Jagger- Maroon 5	17
6) California Gurls- Katy Perry	15
7) Bohemian Rhapsody- Queen	14
8) Alejandro- Lady Gaga	12
9) Poker Face- Lady Gaga	11

The count data for these 1,558 songs are visualized in Figure 5.1. In this dataset, 1,144 songs were named once as INMI and 414 songs were named more than once. Figure 5.2 displays the song count plotted against the recency variable (number of days since exiting the charts) and one of the popularity variables (highest chart position attained).
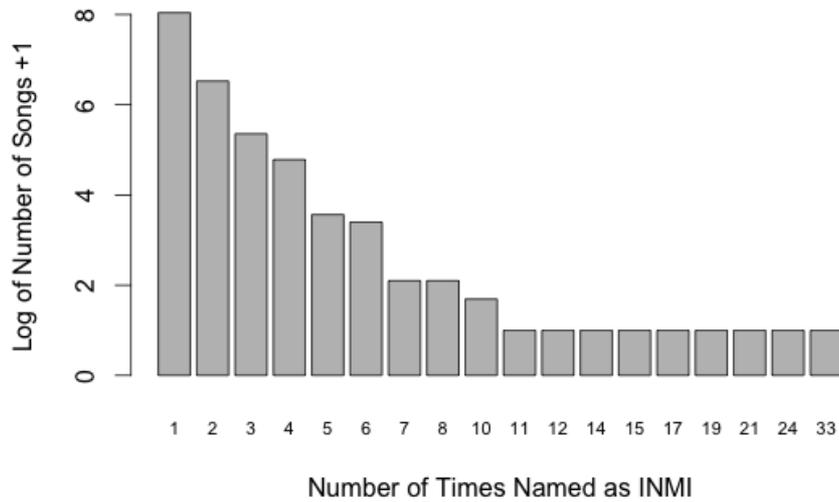
Figure 5.1. Count data for number of times a song was named as INMI in the "Earwormery" questionnaire (counts displayed as logarithm of counts + 1 to compress data for visualization).
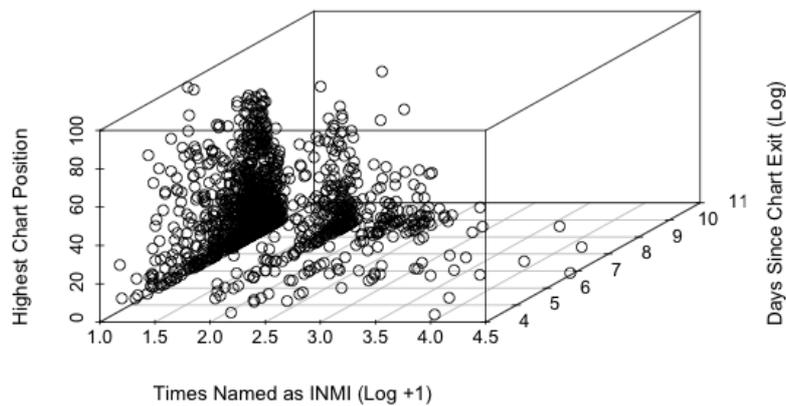
Figure 5.2. Number of times a song was named as INMI by highest chart position attained by number of days since exiting the charts (times named as INMI displayed as logarithm of count +1 and days since exiting the charts displayed as logarithm of days since exiting the charts to compress data for visualization).

5.2.4 Analysis

The next step was to prepare the data for the main analysis, which aimed to investigate whether the number of times a song was named as INMI could be predicted by popularity and recency data. The popularity and recency variables were highly correlated, particularly in terms of highest chart entry and weeks in the charts (see Table 5.2). As such, these three variables were subjected to a one-component principal component analysis (PCA). As the recency variable (days since exiting the charts) did not load highly onto the component, it was excluded and a second one-component PCA was performed on the two popularity variables (highest chart entry and weeks in the charts). The highest chart entry variable loaded negatively onto this component and the weeks in the charts variable loaded positively onto the component. The component scores from this PCA were extracted for use in subsequent regression models as a composite measure of the two popularity variables. The recency variable was subjected to a square root transformation due a non-normal distribution.

Table 5.2: *Correlations of Popularity and Recency Measures*

	Highest Entry	Weeks in Charts	Days Since Exit
Highest Entry	1.00		
Weeks in Charts	- .51**	1.00	
Days Since Exit	- .15**	- .08**	1.00

Note: ** signifies a significant correlation at the level of $p < .001$.

The data in the present project require a class of modelling techniques that can account for count data comprised of whole numbers. Poisson regression is the most traditional method for modelling count data that assume a distribution similar to that displayed in Figure 5.1. Therefore, a Poisson regression model was generated for predicting the data based

on the recency and popularity variables. This model was then compared to several Poisson-related models to find the best fit to the present data. As the Poisson and related models are designed to address data in which the count distribution begins at zero, the present data was transformed by subtracting 1 from each song count so that the 1-counts for the 1,144 songs named once became zeros, the 2-counts became ones, and so on.

One assumption of the Poisson distribution is that the observed mean and the observed variance of the data are equal, as the dispersion parameter (ø) is fixed at 1. Since the observed variance of the present data (2.89) for the song counts is much greater than the observed mean (1.55), this indicates that over-dispersion is present in the data. Over-dispersion is a common problem in count data, and can be dealt with by adopting a quasi-Poisson or negative binomial regression model instead of the more restrictive Poisson model. The quasi-Poisson model deals with over-dispersion by leaving the dispersion parameter unrestricted so that it is estimated by the data, rather than fixing this parameter at 1. The negative binomial model addresses over-dispersion by including an extra parameter specifically to model the over-dispersion.

Additionally, a related family of distributions were tested as possible models for the data, due to their ability to account for excess zero-counts. The hurdle and zero-inflated models can both cope with data that display an inflated zero-count in a more efficient way than Poisson or negative binomial distributions (Mullahy, 1986; Zeileis, Kleiber, & Jackman, 2007). The hurdle model comprises two components: 1) a truncated count component that models the positive counts in a Poisson or negative binomial distribution and 2) a hurdle component that models the zero-counts against larger counts. The zero-inflated model addresses the excess zero-count through two components as well, the first of which models the zeros and the second of which models all of the data in a Poisson or negative binomial distribution. In both the hurdle and zero-inflated models used in the present project, the count component was

specified as a negative binomial distribution (rather than a Poisson distribution), in order to account for the presence of over-dispersion.

5.3 Results

For each of the five methods of modelling detailed above (Poisson, quasi-Poisson, negative binomial, hurdle, and zero-inflated), a model was fitted using the recency variable (square root of days since exiting the charts) and the component scores from the PCA of the popularity variables as predictors of the number of times a song was named as INMI. These five models are presented in Table 5.3.

Table 5.3: *Parameters and Model Summaries for the 5 Reduced Models*

		Estimate	S.E.	z- or t-value	p-value
Poisson	Intercept	0.123	0.082	1.493	.135
	Popularity	0.548	0.038	14.576	< .001 **
	Sqrt Exit Days	-0.011	0.000	-11.432	< .001 **
Quasi-	Intercept	0.123	0.139	0.884	.377
Poisson	Popularity	0.548	0.064	8.627	< .001 **
	Sqrt Exit Days	-0.011	0.002	-6.766	<.001 **
Negative	Intercept	-0.039	0.149	-0.262	.794
Binomial	Popularity	0.467	0.058	5.134	< .001 **
	Sqrt Exit Days	-0.009	0.002	-5.134	< .001 **
Hurdle:	Intercept	-6.815	51.424	-0.133	.895
-Count	Popularity	0.603	0.103	5.845	< .001 **
component	Sqrt Exit Days	-0.013	0.003	-4.681	< .001 **
	Log(theta)	-8.415	51.442	-0.164	.870
-Hurdle	Intercept	-0.695	0.158	-4.411	< .001
component	Popularity	0.310	0.060	5.204	< .001 **
	Sqrt Exit Days	-0.004	0.002	-2.320	.020 *
Zero-Infl.:	Intercept	-0.039	0.140	-0.280	.780
-Count	Popularity	0.467	0.056	8.395	< .001 **
component	Sqrt Exit Days	-0.009	0.002	-5.529	<.001 **
	Log (theta)	-1.015	0.094	-10.826	< .001
-Zero-	Intercept	-9.750	105.684	-0.092	.926
inflated	Popularity	-0.249	29.685	-0.008	.993
component	Sqt Exit Days	-0.013	1.026	-0.012	.990

Note: * = a significant predictor at the level of $p < .05$, ** = a significant predictor at the level of $p < .001$. z-values are presented for all models except the quasi-Poisson, for which a t-value is presented. The variable named here as "Popularity" represents the component scores from the PCA of the two popularity variables: highest chart entry and weeks in the charts.

The five models were then compared to one another using the predict() function in R to predict the distribution that each of these models would assume given the present data as input. Table 5.4 displays the number of responses predicted for each response category in each model for comparison to the present data. Figures 5.3 through 5.6 display the predicted distribution for each model graphically, in comparison to the actual distribution of the present data.

Table 5.4: *Present and Predicted Values for Each Response Category*

Response	Present Data	Poisson	Quasi-Poisson	Negative Binomial	Hurdle	Zero-Inflated
0	1144	1059	1122	1134	1134	1134
1	251	448	390	369	364	369
2	78	35	29	28	29	28
3	44	8	10	10	13	10
4	13	4	3	5	6	5
5	11	1	1	5	5	5
6	3			2	2	2
7	3	1	1	1		1
8				1	1	1
9	2				1	
10	1	1	1			
11	1					
13	1					
14	1			1	1	1
16	1					
17		1				
18	1					
20	1		1			
23	1					
24					1	
25				1		1
32	1					
129					1	
162				1		1
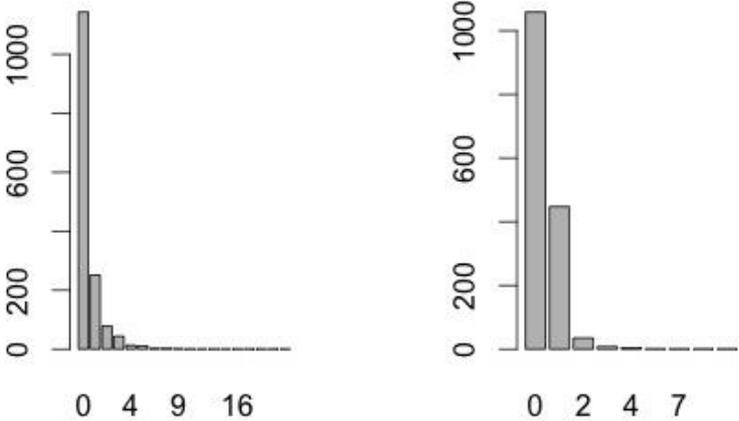
Figure 5.3. Present data (left) vs. predicted Poisson distribution (right).
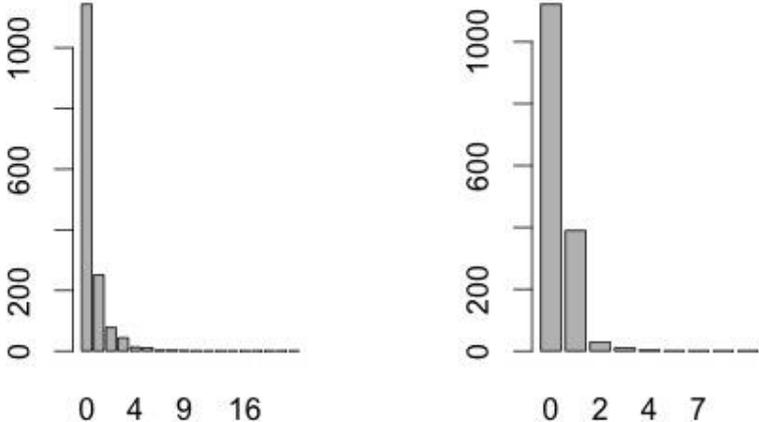
Figure 5.4. Present data (left) vs. predicted quasi-Poisson distribution (right).
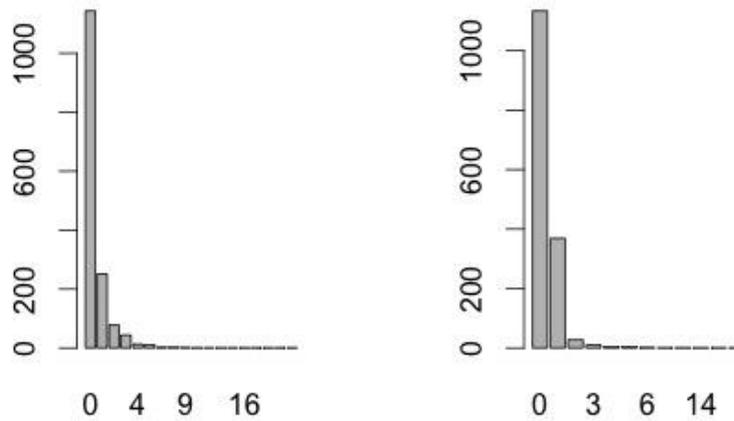
Figure 5.5. Present data (left) vs. predicted negative binomial/ zero-inflated distribution (right).
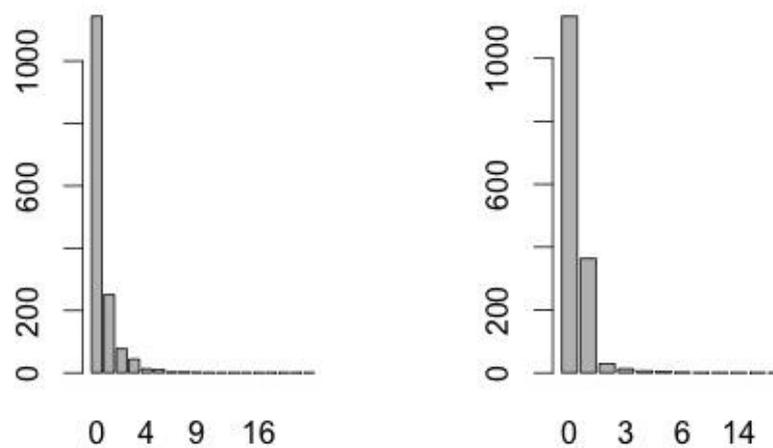
Figure 5.6. Present data (left) vs. predicted hurdle distribution (right).

Finally, the five models were compared in terms of the Akaike Information Criterion (AIC), the Bayesian Information Criterion (BIC), and log-likelihood. These three values provide measures of the goodness-of-fit of each model, with smaller values of the AIC and BIC criteria and higher values of the log-likelihood indicating better model fits. Results of these model comparisons are presented in Table 5.5.

Table 5.5: *Goodness-of-Fit Measures for All Potential Models*

	df	AIC	BIC	Log-Likelihood
Poisson	3	3586.935	3602.989	-1790.468
Quasi-Poisson	3	NA	NA	NA
Neg. Binomial	4	2843.822	2865.226	-1417.911
Hurdle	7	2834.526	NA	-1410.263
Zero-Inflated	7	2849.822	NA	-1417.911

The hurdle model appears on several levels to be the most parsimonious solution for modelling the present data. It achieves the lowest AIC value and highest log-likelihood, while visually resembling the present data more closely than the other models tested (see Figure 5.6 and Table 5.5). This model is also able to account for both the over-dispersion and excess zero-counts present in the current data. This model includes both the popularity and recency variables as significant predictors of the number of times a tune was named as INMI. Specifically, songs that had attained higher chart positions (with number 1 as the highest chart position) and longer runs in the charts and songs that had exited the charts more recently were named as INMI more frequently than less successful and less recent songs.

5.4 Discussion

The results of Study 5a indicate that the frequency with which a tune is reported as INMI can be positively predicted by measures of its popularity and recency. The hurdle model provided the best fit to the present data, but all five of the models tested revealed both the popularity and recency variables to be significant predictors of INMI count. This suggests a key

contribution of these extra-musical features related to the commercial success of a song to the generation of an INMI experience. In light of this finding of a strong effect of extra-musical factors on INMI propensity, Study 5b matched INMI tunes to a "control group" of non-INMI tunes in terms of their popularity and recency. This allowed for an investigation of the role of melodic content in the generation of an INMI experience while controlling for these important extra-musical factors.

Study 5b: Modelling the Melodic Features of INMI Tunes

The main aim of Study 5b was to reveal any underlying melodic features that significantly influence the likelihood that a song would become INMI. The study made use of symbolic (compositional) data—data that is obtained from a score-based representation rather than an audio-based representation of a song—in order to compute 83 statistical features to describe each song's melodic content. The contribution of these 83 melodic features to predicting the likelihood that a song was named as INMI was then assessed using three powerful statistical techniques.

5.5 Method

5.5.1 Matching INMI and Non-INMI Tunes

A subset of the 1,558 songs with chart data[17] was taken that 1) were named as INMI by at least three separate participants in the "Earwormery" questionnaire and 2) had a corresponding high-quality MIDI transcription available from the Geerdes MIDI music database (http://www.geerdes.com/). Step 1 above was included to ensure that the songs included in the present sample had "INMI quality" that transferred

---

[17] Full details on the compilation of this dataset are provided in the *Method* section of Study 5a.

across multiple participants and Step 2 was included as MIDI transcriptions were required as input for the automated melodic feature analysis algorithms. After implementing these two data filtering steps, a dataset of 129 INMI tunes remained for subsequent analyses.

A "control group" of 438 non-INMI tunes was compiled to be from similar genres, artists, and chart positions to the 129 INMI tunes, with the condition that these non-INMI tunes had never been named as INMI by any of the 3,000 questionnaire participants. The INMI and non-INMI tunes were then subjected to a multivariate matching procedure using the GenMatch function from the R package "Matching" (Sekhon, 2011). The two sets of songs were matched on the basis of five variables: highest chart entry, weeks in the charts, days since exiting the charts, artist, and genre. Matching was performed without replacement in order to provide one-to-one matching between INMI and non-INMI tunes. The caliper argument was set to 1.2 standard deviations, meaning that the values for all of the continuous variables defined above for each non-INMI tune were required to be within 1.2 standard deviations of the values of those same variables for its matched INMI tune. Overall, this matching procedure was able to generate matches for 101 of the INMI tunes. Some examples of INMI tunes and their matches, along with the highest chart position attained and days since exiting the charts for each of these songs, are presented in Table 5.6.

Table 5.6: *Most Frequent INMI Tunes and Their Matched Non-INMI Tunes*

INMI Tune	Highest Position	Chart Exit	Non-INMI Tune	Highest Position	Chart Exit
Bad Romance- Lady Gaga (33)	1	1322	Just Dance- Lady Gaga	1	1504
Can't Get You Out of My Head- Kylie Minogue (24)	1	4164	Princess of China- Rihanna	4	475
Somebody that I Used to Know- Gotye (19)	1	398	In For The Kill- La Roux	2	1427
California Gurls- Katy Perry (15)	1	1083	Who's That Chick- Rihanna	6	804
Bohemian Rhapsody- Queen (14)	1	13621	Pretty Woman- Roy Orbison	1	17695

Note: Two of the five variables on which the songs were matched are displayed above (highest chart position attained and number of days since exiting the charts). The number of times each INMI tune was named by a participant in the online questionnaire is displayed in brackets after the artist name.

5.5.2 Melodic Feature Extraction

MIDI transcriptions for all 202 songs (INMI and non-INMI tunes) were obtained from the Geerdes MIDI music database. The melody line of the section of the song (e.g. chorus, verse, instrumental) reported as INMI by the participants in the "Earwormery" questionnaire was extracted for all 101 INMI tunes. If more than one section was reported by different participants, the section that was reported most frequently was extracted. For cases in which no particular song section was reported, the chorus was extracted, as this has been shown in previous research to be the section of a song that is most commonly experienced as INMI (Beaman & Williams, 2010; Hyman et al., 2013). The chorus of each of the 101 non-INMI tunes was then extracted for comparison to the INMI tune excerpts. The full melody line of each of the 202 songs was also extracted for use in some of the subsequent feature analyses. Thus, two versions (excerpt version and full version) of each INMI and non-INMI tune were compiled. During this process, it was noted that one INMI tune (Funky Cold Medina) was comprised primarily of spoken words rather than a melody line. As such, this song and its matched non-INMI tune were excluded from subsequent analysis. This left a dataset of 100 INMI and 100 non-INMI tunes.

Next, all MIDI files were converted to MCSV files, as this is the required input format for the FANTASTIC melodic feature extraction software (Müllensiefen, 2009). MCSV files are tabular representations of monophonic musical events that were created in the present work using the conversion program MELCONV (Frieler, 2005).

A total of 82 melodic structural features were computed using FANTASTIC. These features included both first-order and second-order features of the melodies. Most first-order FANTASTIC features have a second-order counterpart. For instance, the feature *p.range* is a first-

order feature that summarizes the pitch range of a melody by subtracting the lowest pitch in the melody from the highest pitch. The second-order counterpart of this feature is *dens.p.range*, which compares the pitch range of the melody in question to the pitch ranges of all melodies in the reference corpus and computes a measure of how unique the pitch range of the melody is in comparison to the full corpus. The reference corpus used for computing second-order features in the present study was a collection of 14,063 MIDI transcriptions of commercially successful Western pop songs from the Geerdes MIDI music database.

In addition to the distinction made between first- and second-order features, several of the features in FANTASTIC are classified as "m-type features". M-type features are features that aim to capture the usage and repetition of melodic motives in a phrase by taking account of the note order. M-types are calculated within FANTASTIC through the use of a "moving window" that slides over small sections of notes in a melody and records the content of each position of the window. The size of the window can vary, and in the present analysis the window was set to vary from containing two notes at a time up to six notes at a time (the default setting in FANTASTIC). Examples of m-type features include mean m-type entropy (*mean.entropy* in FANTASTIC), which calculates the average entropy value across all computed m-types for a melody, and mean Yules K (*mean.Yules.K* in FANTASTIC), which is a feature taken from linguistics that measures the rate at which words are repeated in a text—or, in this case, the rate at which musical m-types are repeated in the melody.

In the present analysis, all second-order m-type features were computed using the full melody version of the INMI and non-INMI tunes, in order to take account of the repetition of motives throughout each song as a whole. All other FANTASTIC features were computed using the song excerpts as input (the section that had been reported as INMI or the chorus of each song).

In the initial data preparation stages, a logarithmic transformation was applied to 7 of the 82 melodic feature variables computed with FANTASTIC due to non-normality of the variable distributions[18]. Finally, the tempo of each song excerpt was added to the dataset as an additional predictor of interest, resulting in a total of 83 predictor variables for the subsequent analyses. Tempo information, in beats per minute (bpm), for each song excerpt was obtained from the Geerdes MIDI music database.

5.5.3 Analysis

The data were analysed using three different classification techniques, in which the aim was to classify INMI versus non-INMI tunes based on their melodic features. The three techniques will be summarized below.

The random forest method was the first technique utilized in the present study (Breiman, 2001). This method has several advantages over other traditional data classification procedures, as it can handle a large number of predictor variables and can cope with non-linear relationships between variables. In this method, a large number of classification trees are grown using different randomly sampled observations from the dataset for each tree. A classification tree is a machine-learning method that recursively partitions the dataset and fits a prediction model based on the values of the independent variables to classify each observation into a finite set of classes (in this case two classes- INMI or non-INMI; Loh, 2011). In the random forest, the use of averaging across a large number of trees helps to improve predictive accuracy and control over-fitting in comparison to other tree-based methods. The random forest method also provides a measure of variable importance for each of the predictor variables, which is useful in cases in which one wishes to select the most predictive of a large number of variables for further evaluation. Additionally, classification

---

[18] The FANTASTIC variables that were transformed were: *mtcf.mean.log.TFDF, mtcf.norm.log.dist, mtcg.mean.gl.weight, mtcf.mean.productivity, mtcf.mean.Simpsons.D, mtcf.mean.Yules.K,* and *mtcf.TDIDF.m.entropy.*

trees and random forests can model complex higher order interactions between variables that may be difficult to capture using more traditional regression methods. The random forest method was implemented in the present research using the "party" package in R (Hothorn, Hornik, & Zeileis, 2006).

In contrast to random forests, which generally select and combine a small number of individual features for prediction, partial least squares regression (PLSR) (Wold, Ruhe, Wold, & Dunn, 1984) is a method that can be used to group many correlated features into a few components, which are then used to predict the values of the dependent variable. As such, this method combines features of principal component analysis with features of regression. PLSR is commonly used with data for which there are a large number of predictor variables and at least some of these variables are highly collinear, as is the case with the present dataset. Thus, a partial least squares binary logistic regression was implemented using the R package "plsRglm" (Bertrand, Magnanensi, Meyer, & Maumy-Bertrand, 2014).

Finally, a binary logistic regression model was applied using the elastic net method (Zou & Hastie, 2005). The elastic net is a compromise between two other methods: the lasso (Tibshirani, 1996) and ridge regression (Hoerl & Kennard, 1970). The lasso method employs a regularization penalty to set the coefficients of variables that are not highly predictive to zero, which is useful in variable selection, while the ridge method handles correlated predictor variables by treating them as groups of variables. Hence, this method combines strengths of the random forest, in terms of its ability to perform variable selection, *and* the PLSR method, in terms of its ability to group correlated predictors. The optimal balance between the lasso and ridge methods is specified during the elastic net analysis by the α parameter. The elastic net analysis was implemented in the present work using the "glmnet" package in R (Friedman, Hastie, Simon, & Tibshirani, 2015).

Three models were fitted to the present data, using each of these three analysis methods. Then, a leave-one-out cross-validation analysis was conducted to evaluate all three models and obtain an unbiased assessment of their classification accuracy. These steps are reported in the following sections.

5.6 Results

5.6.1 Model Fitting: Random Forest

In a first analysis step, a random forest model was fitted including all 83 predictor variables and INMI versus non-INMI tune as a binary response variable in order to select a subset of variables with the most predictive power. Parameters were set using the "cforest_unbiased" function to ensure that the variable selection and variable importance values were unbiased (Strobl, Boulesteix, Zeileis, & Hothorn, 2007). The total number of trees to be grown was set to 1,000, the number of randomly preselected variables for each split was set to 20, and the minimum number of observations per node necessary for splitting was set to 5. A variable importance score was obtained for each predictor variable, as an indicator of the variable's ability to classify INMI versus non-INMI tunes. The variable importance scores for the top 12 performing variables are visualized in Figure 5.7.

A "confidence interval" criterion was applied to select the top performing variables for further analysis. This criterion specified that only the variables whose (positive) variable importance scores were greater than the absolute value of the lowest negative variable importance score (from the worst performing predictor) would be chosen (Strobl, Malley, & Tutz, 2009). Only the top three performing variables had variable importance scores that met this criterion (*dens.step.cont.glob.dir*, *tempo*, and *dens.int.cont.grad.mean*). These three variables were used in the subsequent cross-validation evaluation, as reported in Section 5.6.4.
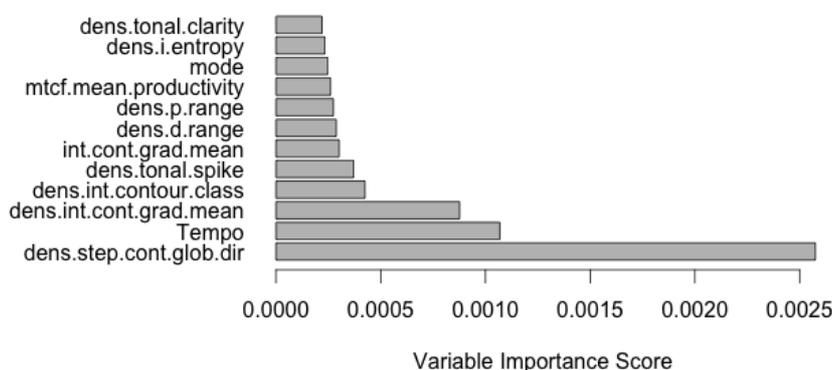
Figure 5.7. Variable importance scores for the 12 most important predictors in the random forest model.

5.6.2 Model Fitting: Partial Least Squares

The first step in the partial least squares regression analysis was to determine the optimal number of components needed for grouping related variables. To accomplish this, a 10-fold cross-validation was run. The maximum number of components to be computed was set to 10 and the cross-validation was run 100 times. The results indicated a two-component PLSR model as the most common solution, with 32% of the cross-validation runs indicating a two-component solution as optimal. As such, a two-component partial least squares binary logistic regression model was fitted to the data. The top 13 predictors from each of the two components are presented in Table 5.7.

5.6.3 Model Fitting: Elastic Net Regularization Method

In the first step of fitting an elastic net regularized model, a 10-fold cross-validation was run 100 time in order to identify the optimal values for the parameters alpha (α) and lambda (λ), which control the balance and the overall strength of the two penalties in the elastic net model. The α

parameter is a value between 0 and 1, where a value of 0 corresponds to a ridge model and a value of 1 corresponds to a lasso model. The ridge model operates by shrinking the coefficients of correlated predictors towards each other, while the lasso tends to pick one of the correlated predictors and discard the others. An α value between 0 and 1 implicates an elastic net model, which works as a comprise between the ridge and lasso methods. The optimal α value obtained for the present data (for which the misclassification rate was lowest) was 0.1.

The λ parameter is a tuning parameter that controls the overall strength of the elastic net penalty. An optimal λ parameter was also obtained through the cross-validation procedure by finding the λ value at which the misclassification rate was lowest with the minimum number of necessary musical feature variables as predictors. The value obtained for λ was 0.43. This model included 13 predictor variables with non-zero regression coefficients, which are listed in Table 5.7. As the response variable was coded as 1=non-INMI tune and 2=INMI tune, positive coefficients indicate an increase in the value of a musical feature for INMI versus non-INMI tunes. For example, INMI tunes are generally faster in tempo than non-INMI tunes (standardized regression coefficient= 0.044).

Table 5.7: *Most Predictive Variables for Classifying INMI vs. Non-INMI*

Random Forest (importance score)	PLS Component 1 (loading)	PLS Component 2 (loading)	Elastic Net (standardized regression coefficient)
dens.step.cont.glob.dir (0.003)	dens.step.cont.glob.dir (0.27)	dens.step.cont.glob.dir (0.37)	dens.step.cont.glob.dir (0.059)
tempo (0.001)	tempo (0.25)	dens.d.punct.trans (-0.31)	tempo (0.044)
dens.int.cont.grad. mean (0.0009)	dens.int.cont.grad. mean (-0.25)	tempo (0.30)	dens.d.punct.trans (-0.044)
dens.int.contour.class (0.0004)	glob.duration (-0.24)	dens.i.entropy (-0.28)	dens.int.cont.grad. mean (-0.039)
dens.tonal.spike (0.0004)	dens.d.punct.trans (-0.23)	dens.d.range (-0.27)	glob.duration (-0.026)
int.cont.grad.mean (0.0003)	int.cont.grad.mean (0.22)	glob.duration (-0.26)	dens.i.entropy (-0.021)
dens.d.range (0.0003)	dens.int.contour.class (0.19)	p.range (-0.25)	dens.tonal.clarity (-0.017)

dens.p.range (0.0003)	dens.i.entropy (-0.18)	dens.tonal.clarity (-0.24)	int.cont.grad.mean (0.013)
mtcf.mean.productivity (0.0003)	dens.tonal.clarity (-0.18)	dens.p.range (-0.23)	tonal.clarity (0.009)
mode (0.0002)	d.dotted.trans (0.17)	dens.step.cont.loc.var (-0.23)	dens.int.contour.class (0.009)
dens.i.entropy (0.0002)	tonal.clarity (0.17)	i.abs.std (-0.22)	dens.p.range (-0.009)
dens.tonal.clarity (0.0002)	len (-0.16)	tonal.clarity (0.22)	dens.d.range (-0.0004)
tonal.clarity (0.0002)	dens.p.range (-0.16)	step.cont.glob.var (-0.21)	dens.d.half.trans (0.00017)

Note: The variables are listed in each column in descending order of importance in the models. Importance scores, factor loadings, or regression coefficients are listed in brackets following the FANTASTIC feature name.

5.6.4 Cross Validation and Model Evaluation

All three classes of models were then evaluated using the same leave-one-out cross-validation to obtain an unbiased assessment of their classification accuracy. The random forest and the elastic net were cross-validated using only the three and 13 variables respectively that were selected in the first stage, while the PLSR model was evaluated with only two components as determined previously. The elastic net parameters were also set to $\alpha = 0.1$ and $\lambda = 0.43$ as determined in the model fitting stage. The cross-validated classification rate for the random forest was 62.5%, while the elastic net achieved only a classification rate of 52% and the PLSR model achieved a rate of 49%.

As the random forest achieved the highest classification rate in the cross-validation analysis, the directionality of the relationships between the three variables included in the random forest and the dependent variable (INMI vs. non-INMI) were investigated further. These three variables were thus entered into a classification tree (Figure 5.8) and a binary logistic regression analysis (Table 5.8). The classification tree indicates that tunes with a common global melodic contour are more likely to become INMI (*dens.step.cont.glob.dir* values greater than 0.326) and approximately 80% of the tunes fulfilling this criterion were classified as INMI tunes. For tunes with a global melodic contour value less than or equal to 0.326, the feature *dens.int.cont.grad.mean* is then taken into account by the classification tree. This feature measures the commonness of the average gradient of the interpolation lines between melodic contour turning points. Tunes that are less similar to the corpus in terms of this features (*dens.int.cont.grad.mean* values less than or equal to 0.421) are more likely to become INMI than tunes that have a more common average contour gradient. Within this particular model tempo was not selected as a deciding variable, however this single classification tree should not be interpreted as representing the outcomes of the full random forest, which averages across a large number of

classification trees. As such, the results of the binary logistic regression add further clarity to the question of directionality of effects. This regression model indicated that INMI tunes had faster tempi than non-INMI tunes and also indicated the same directionality of effects of the contour variables as the classification tree, although only the *dens.step.cont.glob.dir* variable reached the conventional significance level of $p < .05$. No significant two- or three-way interactions were found between any of the predictor variables in the logistic regression analysis.
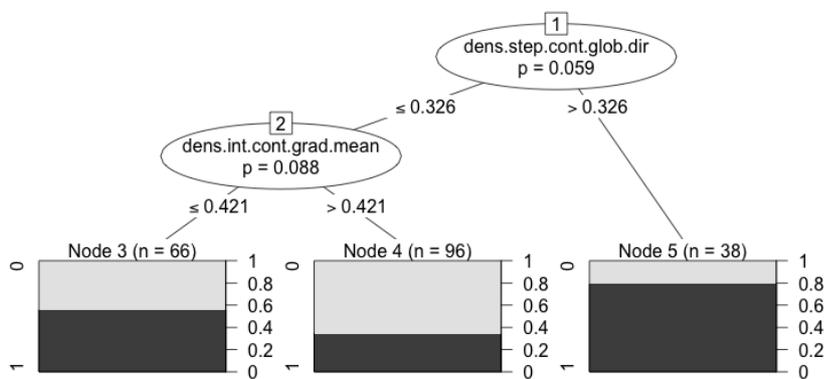
Figure 5.8. Classification tree resulting from the three predictors used in the random forest.

Table 5.8: *Binary Logistic Regression Analysis Using the Three Random Forest Predictors*

	Estimate	S.E.	*z*-value	*p*-value
Intercept	-2.42	1.28	-1.89	.06
Tempo	0.01	0.01	1.66	.10
dens.int.cont.grad.mean	-1.93	1.07	-1.81	.07
dens.step.cont.glob.dir	7.08	3.31	2.14	.03*

Note: * = significant predictor at the level of $p < .05$.

5.6.4.1 Comparisons to Human INMI Classifications

As a final query in this line of research, the model classification rates were compared to evaluations made by human participants. A subset of 25 INMI tunes and their matched non-INMI tunes were randomly selected. These pairs of tunes were presented to 31 volunteers (mean age = 38.1 years (SD = 11.5); 10 males, 18 females, 3 participants did not provide gender information) who were attending an "Earworm Symposium" at Goldsmiths, University of London. As such, this sample included INMI researchers from around the globe, other music researchers and professional musicians, and members of the general public. For each pair of tunes that was presented, the volunteers were asked to try to detect which tune they thought had been named as an earworm. They were asked to focus in particular on the melodic content of each tune and how likely they thought a tune was to get stuck in someone's head. The same excerpts of the tunes were presented as had been used for the FANTASTIC feature analysis. Unlike in the FANTASTIC analysis, the full audio from the original recording of each song was presented to the participants. Thus, these recordings included not only the main melodic lines but also the harmonic content, timbral features, lyrics, etc. Overall, the average accuracy rate on this task

across all participants was 65.8%. In comparison to the cross-validated random forest classification rate of 62.5%, it appears that human participants are not much more accurate in classifying INMI versus non-INMI tunes than a computer algorithm. Interestingly, this is the case even though the original recordings were presented to the participants, which included audio data that was not available to the computer algorithms.

5.7 Discussion

In the cross-validation evaluation, the random forest emerged as the best performing model, with a 62.5% classification rate, while the other two models performed very close to chance level of 50% classification accuracy. The random forest made use of only three melodic features, specifically song tempo and two second-order features that express how common the contour of a song is with respect to the reference corpus. The same three features were also included among the most important features in the PLSR and elastic net models, providing further evidence for their general importance in predicting whether a tune is likely to become INMI. Thus, the propensity of pop songs to be spontaneously recalled and repeated within the mind seems to be related to their tempo and the commonness of their melodic contour.

Tunes that were more likely to become INMI were generally faster in tempo than non-INMI tunes, although this predictor was only marginally significant in the logistic regression analysis. This finding could potentially relate to the fact that certain musical tempo ranges are more easily entrained to than others (McAuley et al., 2006), particularly considering the finding from Study 4 of this thesis that around 25% of INMI episodes reported by participants occurred during repetitive movements. In terms of the FANTASTIC feature *dens.step.cont.glob.dir,* the findings of the present research indicate that tunes with *more common* global contour shapes are more likely to become INMI than those with less common contours. Some examples of tunes from the present dataset with the least

and most common global contour directions with respect to the reference corpus are provided in Figure 5.9. In these particular examples, the tunes with more common global contours (section B of Figure 5.9) assume fairly arch-shaped phrases. This is in line with previous research citing the melodic arch as one of the most common contour shapes in Western music traditions (Huron, 1996). Tunes in the present dataset with less common global contours (section A of Figure 5.9) appear to take on contours other than arch shapes, such as ascending melodic lines that do not descend again. Finally, the findings related to the feature *dens.int.cont.grad.mean* indicate that tunes with a *less common* average gradient (slope) of the melodic lines between contour turning points are more likely to become INMI. Some examples of the extreme values of this particular variable within the present dataset are presented in Figure 5.10. The tunes with more common average contour gradients within this sample appear to comprise mostly stepwise intervallic motion or repetitions of the same note, whereas the tunes with less common average contour gradients tend to contain many melodic leaps (as in A1 of Figure 5.10) or uncharacteristically large melodic leaps (as in A2 of Figure 5.10). However, further research is needed to examine whether such contour patterns hold across larger datasets and other musical genres, in order to provide additional systematic evidence for the role of these melodic features in predicting INMI tunes. Future research could also involve composing novel tunes that incorporate various dimensions and combinations of these melodic features (e.g., slow versus fast tunes, tunes with common versus uncommon global contour shapes, etc.) and presenting them in a laboratory context in order to test whether participants experience some of these tunes more frequently as INMI than others. Such research would provide further control over extraneous variables such as popularity and familiarity.

A) Uncommon Global Contours (lowest values of *dens.step.cont.glob.dir*):

A1) Owner of a Lonely Heart (Yes) (INMI tune)

A2) Rock 'N' Me (Steve Miller Band) (non-INMI tune)

B) Common Global Contours (highest values of *dens.step.cont.glob.dir*):

B1) Smoke on the Water (Deep Purple) (INMI tune)

B2) Plug in Baby (Muse) (INMI tune)

Figure 5.9. Examples from the present dataset with the lowest (A) and highest (B) values of the variable *dens.step.cont.glob.dir.*

A) Uncommon Avg. Contour Gradient (lowest values of *dens.int.cont.grad.mean):*

A1) In the Mood (Glenn Miller) (INMI tune)
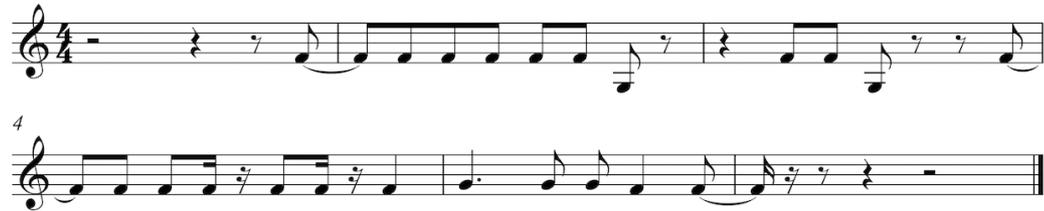
A2) My Sharona (Knack) (INMI tune)

B) Common Avg. Contour Gradient (highest values of *dens.int.cont.grad.mean*):
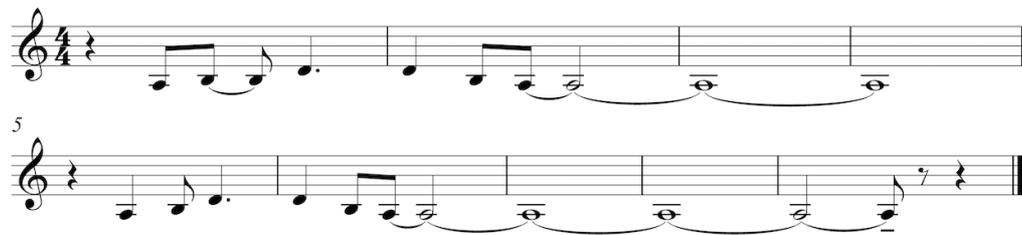
B1) Lucky Man (Verve) (non-INMI tune)
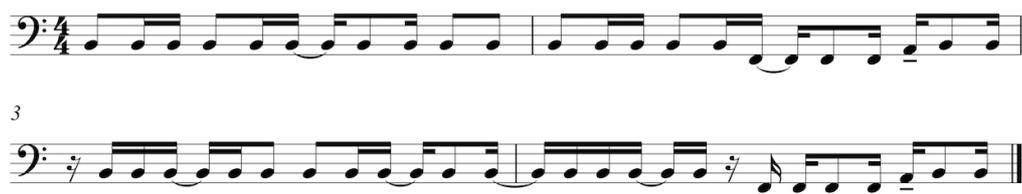
B2) Intergalactic (Beastie Boys) (INMI tune)

Figure 5.10. Examples from the present dataset with the lowest (A) and highest (B) values of the variable *dens.int.cont.grad.mean.*

It should also be noted that several additional melodic features appeared in the list of the most predictive features across all three of the models tested in Study 5b (see Table 5.7). Although the potential effects of these variables are likely fairly small, it is noteworthy that all three models

tested in the present study come to fairly similar solutions in terms of the features that are most predictive of INMI quality. Additional features that were present in the top 13 variables across all three models are *dens.int.contour.class, int.cont.grad.mean, dens.d.range, dens.p.range, i.entropy, dens.tonal.clarity,* and *tonal.clarity*. This set of additional features includes several second-order features (those features beginning with the prefix *dens*), such as measures of the duration range, pitch range, and tonal clarity of a tune in comparison to the corpus (*dens.d.range, dens.p.range*, *dens.tonal.clarity*; note that the first-order equivalent *tonal.clarity* is also featured). The tonal clarity features measure the extent to which a melody can be defined as fitting within a single key. The features *dens.int.contour.class* and *int.cont.grad.mean* describe a tune's melodic contour, thus further highlighting the importance of melodic contour in INMI prediction, and *i.entropy* describes the entropy (i.e., uncertainty) of the interval content of a melody.

5.8 Extra- and Intra-Musical Features That Predict INMI: Combined Analysis of Data from Studies 5a and 5b

As a final step in understanding the influence of both extra- and intra-musical features on the likelihood of a tune to become INMI, a combined analysis was conducted that included features of both analyses from Studies 5a and 5b. The aim of this analysis was to predict the number of times a tune had been named as INMI based on both the popularity and recency variables employed in Study 5a and the predictions generated by the random forest model of melodic features from Study 5b. This combined analysis was conducted using the dataset of 200 tunes from Study 5b (100 INMI and 100 non-INMI tunes).

As in Study 5a, the highest UK chart entry and weeks in the charts variables were highly correlated. These variables were thus subjected to a one-component principal component analysis and the component scores were extracted for use as a combined measure of these two

variables. The highest chart entry variable loaded negatively onto this component and the weeks in the charts variable loaded positively onto the component. A square root transformation was also applied to the recency variable (number of days since a song exited the charts), as this variable was positively skewed. Thus, three predictor variables were included in the subsequent analysis: the combined measure of the two popularity variables, the square root-transformed recency variable, and the binary predictions of the final random forest model that made use of three melodic features (*tempo*, *dens.int.cont.grad.mean,* and *dens.step.cont.glob.dir*). The dependent variable of interest was the number of times a tune had been named as INMI. Non-INMI tunes were coded as a value of 0 for this variable. As the dataset from Study 5b included only INMI tunes that had been named at least 3 times, the count variable for each INMI tune was transformed by subtracting 2 from each count, so that there was no gap in the distribution between the zero-counts of the non-INMI tunes and the count data for the INMI tunes.

As in Study 5a, a hurdle model provided the best fit to the present data, in comparison to similar methods such as Poisson regression, based on several goodness-of-fit criteria (see Table 5.9). The hurdle model resulting from the combined analysis is presented in Table 5.10. The results indicate that the popularity and recency variables were significant predictors in the count component of the model, which models the positive (non-zero) counts from the dataset, whereas the random forest predictions were a significant predictor in the hurdle component, which models the zero-counts against the larger (non-zero) counts. This result can be interpreted to indicate that the melodic feature data from the random forest significantly predicts whether a tune is named as INMI (whether its INMI count is zero or non-zero) and that the popularity and recency of a song serves to predict *how many times* a tune is named as INMI (for non-zero INMI counts).

Table 5.9: *Goodness-of-Fit Measures for All Potential Models: Combined Analysis*

	df	AIC	BIC	Log-Likelihood
Poisson	4	736.18	749.37	-364.09
Quasi-Poisson	4	NA	NA	NA
Neg. Binomial	5	583.70	600.19	-286.85
Hurdle	9	543.22	NA	-262.61
Zero-Inflated	9	568.51	NA	-275.26

Table 5.10: *Hurdle Model with Extra- and Intra-Musical Predictors of INMI Likelihood*

		Estimate	S.E.	z-value	p-value
-Count component	Intercept	0.712	0.697	1.021	.307
	Popularity	0.629	0.162	3.885	.0001**
	Sqrt Exit Days	-0.012	0.005	-2.554	.011*
	RF Predictions	-0.101	0.370	-0.273	.785
	Log(theta)	-1.037	0.911	-1.138	.255
-Hurdle component	Intercept	-0.932	0.443	-2.105	.035
	Popularity	0.111	0.170	0.651	.515
	Sqrt Exit Days	-0.001	0.005	-0.227	.820
	RF Predictions	2.302	0.336	6.842	< .0001**

Note: * = a significant predictor at the level of $p < .05$, ** = a significant predictor at the level of $p < .001$. The variable named here as "Popularity" represents the component scores from the PCA of the two popularity variables: highest chart entry and weeks in the charts.

The predictions from the hurdle model (obtained using the predict() function in R) are compared to the count distribution of the actual data in Figure 5.11. This figure indicates that, although the hurdle model is informative in terms of offering an interpretation of the contributions of both extra- and intra-musical features to the INMI nature of a tune, it is not entirely accurate in forming predictions that resemble the original data distribution. However, the overestimation of one-counts and underestimation of zero-counts seen in Figure 5.11 could be at least partially due to the fact that a somewhat artificial distribution was imposed on the present data (by subtracting 2 from each count for the INMI tunes but not the non-INMI tunes). As all instances of INMI tunes being named once or twice were excluded during the data compilation process in Study 5b, it is difficult to estimate how these counts would feature in the model had they been included.
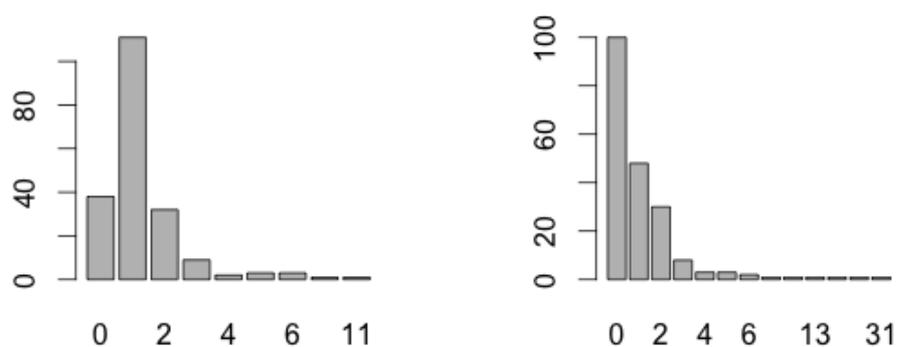
Figure 5.11. Predicted hurdle distribution for combined data (left) vs. present data (right).

A likelihood ratio test between the hurdle model presented in Table 5.10 (log-likelihood = -262.61) and an analogous model that includes only the popularity and recency variables (log-likelihood = -290.37) revealed a

significant difference between the two models, $X^2(2) = 55.52$, $p < .001$, indicating that the inclusion of the melodic feature-based predictor (predictions from the random forest) provides substantial explanatory power over the model including only extra-musical features. Thus, it appears that both extra- and intra-musical features contribute to the generation of an INMI experience.

5.9 General Discussion

The results of the present chapter indicate that measures of a song's popularity as well as features of its melodic structure can be used to predict INMI propensity, thereby demonstrating that both extra- and intra-musical variables can play a role in increasing the likelihood that a particular tune will be recalled as INMI. These findings contribute to the growing literature on the INMI experience and serve to increase our general understanding of why certain songs are spontaneously recalled in the mind over others.

The extra-musical features that played a significant role in predicting the number of times a song was named as INMI by participants in Study 5a were the song's highest UK Chart entry, the number of weeks the song spent in the charts, and the number of days since the song had exited the charts. These findings bear some similarities to previous work on extra-musical features that influence INMI occurrence. For instance, Byron and Fowles (2013) reported that previously unfamiliar songs were more likely to become INMI if participants were exposed to them six rather than two times, thereby suggesting a role of familiarity in the INMI experience. Additionally, recency of exposure seems to play a key role, such that songs that have been heard aloud more recently are more likely to become INMI than songs that were heard less recently (Hyman et al., 2013; Liikkanen, 2012b). The UK chart variables used in the present work are conceptually related to familiarity and recency measures, as songs that achieve higher positions and longer runs in the charts may be played

more often on the radio and online platforms such as YouTube and Spotify, thereby increasing their familiarity in the listener. Songs that have more recently been in the charts may also have a greater chance of having been recently heard by participants, due to the bias of many radio stations towards playing currently popular hits.

The three melodic features that were most predictive of the likelihood of a tune to become INMI are somewhat related to previous literature as well. The "Hooked" project on musical catchiness has revealed audio features related to melodic "conventionality" as positive predictors of the long-term salience of a melody (Van Balen et al., in press), which bears some conceptual similarity to the finding in the present work that common global melodic contours were positive predictors of INMI tunes. Müllensiefen and Halpern (2014) reported a variety of features that predicted explicit and implicit memory for previously unfamiliar songs. Interestingly, there is little overlap between the features revealed by their research and the present study, with the possible exception of the fact that they found simple contours to be predictive of implicit melodic memory. However, as Müllensiefen and Halpern's study investigated recall of novel tunes after only a single exposure it is not surprising that the melodic features implicated in their work are rather different to the present findings on features that enhance the spontaneous recall and repetition of familiar music within musical imagery. Additionally, the present work did not replicate the findings of Finkel et al. (2010) and Williamson and Müllensiefen (2012), who reported that, on average, INMI tunes made use of longer note durations and smaller pitch intervals than non-INMI tunes. The present findings may be at least somewhat related to Williamson and Müllensiefen's account that INMI tunes may be easier to sing, as, for instance, common global contours may be easier to sing than less common contours. This is, however, a speculative explanation that needs further evidence from experiments that test the "singability" of frequently named INMI versus non-INMI tunes in a controlled setting. Finally, the finding in the present work that INMI tunes tended to be faster in tempo than non-INMI tunes does not appear to have a previous

precedent in existing literature. It is of interest to explore this finding further in terms of potential relationships of this tempo variable to sensorimotor and entrainment processes.

5.9.1 Limitations and Future Directions

In Study 5a, data related to a song's position and entry in the UK charts were used as measures of popularity and recency. However, there are various other ways in which song popularity and recency might be operationalized. For instance, one could measure song popularity in terms of the number of plays of a song on YouTube or the number of downloads of the song on iTunes or other online music stores. Recency of exposure to a song could also be measured in terms of the amount of recent airplay on radio stations or even on an individual basis by examining participants' recent listening history on platforms such as Spotify. In light of previous findings suggesting that INMI is often a highly idiosyncratic experience (Beaman & Williams, 2010; Williamson & Müllensiefen, 2012) this last point might be especially relevant—that is, taking account of participants' individual listening histories, although methodologically challenging, could provide potentially fruitful avenues for exploring the triggers and causes of INMI experiences. If additional measures of popularity and recency, such as those detailed above, could be obtained in future research, one could match INMI and non-INMI tunes on these variables as well in order to provide even stricter control over the matching procedure used in Study 5b.

In regard to Study 5b, there are several potentially promising avenues for future investigation and expansion of the present research that should be highlighted here. First, the present work comprises only symbolic data analysis and does not include measures of audio features derived from the actual song recordings, such as loudness, timbral content, rhythmic clarity, etc. Additionally, there may be other key compositional features not represented within the single-line melodic analysis implemented in FANTASTIC, such as the harmonic content or chord structure of the

music, articulation, and expressive timing, which could contribute to the INMI nature of a tune. An expanded version of the present study that includes analysis of audio features as well as other structural features of the INMI tunes would be highly beneficial in terms of identifying additional features that can increase the classification accuracy of the models. Finally, an analysis of the *lyrical* content of INMI tunes could be beneficial, in terms of investigating whether linguistic features, such as rhyme or alliteration, play a role in increasing the likelihood of a song toward becoming INMI.

Future research should also compare the results of the current study to data from music styles not included in the present sample in order to identify commonalities and/or differences in INMI features across diverse genres. Due to the use of UK Music Chart data, a large number of reports of INMI experiences from the "Earwormery" questionnaire were excluded from consideration in the present work, including INMI experiences of classical music, TV theme tunes, and children's songs. Methodological extensions that would allow this excluded set of data to be examined in a similar way to the present data could help to shed light on whether the melodic features revealed in Study 5b may be genre invariant.

A final question for future research is whether the melodic features that increase a tune's INMI propensity might actually be related to two subcomponents of the experience, specifically, 1) features that enhance the spontaneous retrieval of a tune from memory and 2) features that contribute to the looping nature of INMI within the mind. It is somewhat surprising that in the present research features related to the repetitiveness of melodies were not particularly predictive, as repetitiveness of a melody could plausibly be related to the repetitiveness of a tune within the mind. However, this may be due to the way in which the research was conducted, in which participants reported their most recent and/or most frequent INMI tunes via a retrospective questionnaire. Future research could examine more specifically tunes that come back frequently as INMI or tunes that are experienced as INMI for long

durations of time (using diary or ESM methods to assess INMI duration) in order to more directly isolate a dataset of tunes that tend to repeat over and over in the mind. This dataset could then be used to assess whether specific melodic features contribute to the incessant repetitiveness that sometimes (but not always) accompanies the INMI experience.

5.9.2 Conclusion

The outcomes of the present research indicate that both extra- and intra-musical features of a pop song can be used to predict the likelihood that the song will become INMI. The results of this work may be of interest to researchers of musical memory, as well as to music composers and advertisers interested in writing music that will continue to be spontaneously replayed in one's head long after the initial music exposure period. It is possible that the melodic features revealed in this work as predictors of INMI might also serve more general functions in terms of increasing the ease with which a tune can be retrieved from memory, although further research needs to be conducted to test this possibility.

It is unlikely that a model including only melodic features of tunes will ever achieve a classification rate of 100% for classifying INMI tunes, as it is clear from the present research that both extra- and intra-musical features play a role in the genesis of an INMI experience. It would be highly beneficial in future research to begin to construct a model that takes account of not only acoustic, melodic, harmonic, and lyrical features of melodies, but also participant-level factors such as personal listening habits, personal associations with music, and endogenous states (e.g., mood) in order to provide a more comprehensive overview of the factors that contribute to the generation of an INMI experience.

Chapter 6. General Discussion

Abstract

This chapter summarizes the main research findings of the work presented within the thesis. Implications of the research within the psychological literature and the realm of music performance are discussed. General limitations of the work and possible future extensions are also outlined. In sum, the present research has provided several novel insights regarding the temporal and melodic features of voluntarily and involuntarily generated musical imagery.

6.1 Introduction

Previous research has established that temporal and melodic elements of music are represented within voluntary musical imagery (VMI) (e.g., Farah & Smith, 1983; Halpern, 1988a; Weber & Brown, 1986) and that many parallels exist between the representation of these musical features within VMI and perceived music (e.g., Halpern, 1988b; Janata & Paroo, 2006; Weber & Brown, 1986). Existing work also suggests that representations of tempo and pitch within VMI are fairly veridical as compared to the original version of a song and fairly consistent across multiple trials (Halpern, 1989, 1992; Levitin, 1994; Levitin & Cook, 1996). However, some studies have revealed degradations of temporal information within musical imagery as compared to pitch information (Bailes, 2002; Janata & Paroo, 2006; Weir et al., 2015) and multiple studies have revealed training-related effects, such that musicians perform better on certain musical imagery tasks than non-musicians (Halpern, 1989, 1992; Janata & Paroo, 2006).

Previous research on *involuntary* musical imagery (INMI) has primarily examined the phenomenology of INMI (e.g., duration, genre of music experienced, section of the song experienced; Beaman & Williams, 2010;

Halpern & Bartlett, 2011; Hyman et al., 2013), the situations surrounding the genesis of an INMI episode (e.g., concurrent task, environmental triggers of the episode; Floridou & Müllensiefen, 2015; Hyman et al., 2013; Williamson et al., 2012), and individual differences related to the propensity to experience INMI (e.g., personality traits or cognitive style; Beaman & Williams, 2013; Beaty et al., 2013; Floridou et al., 2012; Müllensiefen et al., 2014a). However, previous research has not probed the musical content of the INMI experience in an objective manner, as in the studies of VMI described above, in order to detect whether features of the music itself can play a role in the genesis and underlying functions of an INMI experience.

The research presented in this thesis provides several new insights into the temporal and melodic content of both voluntary and involuntary musical imagery. The first studies within the thesis aimed to resolve some of the issues highlighted above regarding the potential degradation of tempo information within VMI by revealing several psychological factors and individual differences between participants that can influence the precision and consistency of tempo representations within VMI. The findings of this research were then used to inform the first known study on the tempo of *involuntary* musical imagery. This study revealed several parallels between involuntary and voluntary imagery, in terms of memory precision and potential functions. Finally, the empirical section of this thesis closed by examining a number of melodic features that characterize the INMI experience in order to understand how such features might contribute to the genesis of an INMI episode. The following section will recapitulate the main research findings of each of the five studies reported in this thesis and discuss their relationship to existing psychological literature.

6.2 Summary of Findings

Study 1 investigated whether one factor that has been found to affect time perception—arousal—could influence tempo choices for familiar songs recalled from long-term memory. The results indicate that an increase in physiological arousal can significantly increase the tempo at which a familiar song sounds "right", in both perceived and imagined musical contexts. In this study, participants' average chosen tempi for familiar music increased following an exercise condition that increased heart rate in comparison to a control condition that did not significantly increase heart rate (anagram completion). Additionally, no effects of prior musical training or self-rated auditory imagery abilities were found on task performance, suggesting that the arousal manipulation affected more and less musically experienced participants similarly. Arousal is a factor that has previously been shown to influence time perception in a variety of time estimation and time production tasks (Grommet et al., 2011; Miró et al., 2003; Ozel et al., 2004; Wearden et al., 1999). The results of Study 1 extend this area of research by demonstrating that arousal can also influence mental representations for familiar and complex time-related stimuli such as music that are recalled from long-term memory. Similar results were found in a perception and a VMI task, thereby revealing an additional parallel between the mechanisms implicated in perceiving and imagining music.

Study 2 tested the influence of age, another factor known to affect time-related judgements, on tempo choices for familiar melodies. As previous studies had revealed a general slowing of preferred musical tempo and spontaneous tapping rates with age (Drake et al., 2000; McAuley et al., 2006; Vanneste et al., 2001), it was hypothesized that age would be negatively related to average chosen tempo for familiar songs in both a musical perception and a VMI condition. However, the results of Study 2 indicated that the average tempo at which familiar, non-canonical songs

sounded "right" was *not* significantly related to age in the perception condition. Additionally, a quadratic effect of age on chosen tempo was found in the imagery condition, such that both younger and older adults chose faster average tempi for familiar songs than middle-aged adults. This finding suggests that age-related differences in temporal representations within musical imagery cannot be explained solely by a slowing of the internal clock, and highlights the need to investigate additional measures such as working memory capacity and attentional demands of the imagery task.

Study 3 provided evidence that veridicality of recall for the tempo of familiar music can be directly affected by the task used to measure tempo judgements. Recall of the tempo of familiar pop songs was significantly more veridical in a perception condition—in which all musical cues from the original song were provided—than in two VMI conditions. However, recall was also significantly more veridical in a VMI condition that involved tapping to the beat of imagined music than in a VMI condition that involved adjusting a click track to the beat of imagined music, thus suggesting that motor engagement with imagery can play a role in refining tempo judgements. This finding has parallels to a handful of previous studies that have found that performance on imagined timing tasks can be improved when a motor component is involved (Clynes & Walker, 1982; Manning & Schutz, 2013). Study 3 also revealed that performance on the tapping to imagined music task was positively related to formal musical training, whereas performance on the task involving adjusting a click track to the beat of imagined music was positively related to active musical engagement (e.g., active listening, concert attendance, etc.). On the other hand, performance on the perception task was positively influenced by prior familiarity with a stimulus but not musical background. These findings suggest that musical experience does not necessarily aid performance on tempo judgement tasks in which a perceived stimulus is presented, but that it can play a key role when such a task is abstracted into the imagined music domain. Additionally, significant positive effects of the original, recorded tempo of a song

stimulus were found on performance in the perception and tapping to imagined music tasks, suggesting that certain tempo ranges of pop songs may be recalled more easily from memory than others.

Study 4 introduced a novel method for measuring the tempo of INMI episodes during the course of everyday life. The results of this work revealed several parallels between the findings on temporal features of VMI reported in this thesis and the temporal features of INMI. In particular, INMI tempo was generally recalled with a high degree of veridicality in comparison to the original version of a tune, thus paralleling existing findings on the deliberate recall of musical tempo (Study 3; see also Levitin & Cook, 1996). Evidence for a "regression to the mean" for INMI tempo was found, which bears similarities to the findings reported by Halpern (1988b) that slow tempo songs tended to be imagined faster than their usual tempo and fast songs tended to be imagined slower than their usual tempo in a VMI task. Finally, a positive relationship was found between INMI tempo and concurrent arousal level, which parallels the findings of Study 1 of this thesis in the VMI and music perception domains. This significant relationship between arousal and INMI tempo provides some initial evidence for potential mood regulatory functions of INMI, which should be explored further in future research.

Study 5 provided evidence that the likelihood that a pop song will be reported as INMI can be influenced by both extra- and intra-musical features of the song. It was found that songs that had more recently been in the UK Music Charts and had had more success within the charts were more frequently named as INMI than less recent and less successful tunes. Additionally, the likelihood that a tune would become INMI could also be predicted by inherent features of the musical structure of the tune, including its tempo and the commonness of its melodic contour. The findings on the extra-musical predictors of INMI relate to previous findings suggesting that recency and familiarity play a role in the likelihood that a song will become INMI (Byron & Fowles, 2013; Liikkanen, 2012b; Williamson et al., 2012). Previous research has also revealed that

specific musical features of a tune can be used to predict its memorability (Müllensiefen & Halpern, 2014) and melodic "catchiness" (Burgoyne et al., 2013). The melodic predictors that featured most highly in the present work on INMI bear some similarities but do not overlap entirely with features revealed in this previous work, suggesting that the melodic features that contribute to the spontaneous and repetitive recall of a tune within the mind are somewhat divergent from features that increase its explicit and implicit memorability and "catchiness".

6.3 Implications of the Research

6.3.1 Implications for Psychological Research

The present research has several key implications for the psychological literature on musical imagery and musical memory. The first two studies of this thesis have revealed that factors known to influence time perception in general can influence tempo judgements for both perceived and imagined music. These findings indicate that factors such as the age range of the sample and the present arousal level of participants (e.g., has a participant just consumed a double espresso?) should be taken into careful consideration when conducting research requiring perceived and/or imagined musical tempo judgements. It is also plausible that other factors that have been shown to affect time perception (e.g., psychological disorders, drugs) might affect musical tempo judgements. The influence of such additional factors should be investigated in future research in order to provide a more complete picture of how mental tempo representations can be preserved or distorted within musical imagery. Finally, the present research has indicated that arousal also relates significantly to temporal representations within *involuntarily* recalled musical memories. This result suggests that INMI tempo can be similarly affected by factors known to influence time perception, although the causal direction of the relationship between INMI tempo and arousal requires further investigation.

Previous work in the imagined pitch domain by Janata and Navarro Cebrian (2010) and Janata (2012) has revealed that, in general, the greater the degree of sensory support afforded by an imagery task, the more accurately participants performed on the task. The findings of Study 3 of this thesis have revealed parallels to this research within the imagined *tempo* domain, as performance in a tempo judgement task where a perceived musical stimulus was present was significantly more accurate than in pure imagery tasks. The present findings have also revealed a significant influence of an additional type of "sensory support", in the form of motor feedback, in refining imagined tempo judgements. Future work that further varies the amount of sensory support afforded by imagined tempo tasks should be conducted to provide additional evidence as to how mental tempo representations may be reinforced by perceptual cues.

The results of this work have also revealed a number of individual differences that serve to predict performance on perceived and imagined tempo tasks. Previous research had indicated that musical training can aid performance in imagined tempo and pitch tasks (Halpern, 1989, 1992; Janata & Paroo, 2006), but evidence for any other participant- or song-level factors that affect imagery task performance was lacking. One key finding of the present work in regard to individual differences is that formal musical training had a significant effect on certain imagined tempo tasks (tapping to imagined music in Study 3), but not others (Study 1 and the click track adjustment task in Study 3). Additionally, the presence of a significant effect of active musical engagement on task performance in the click track adjustment task in Study 3 indicates that not all musical experience-related effects on imagery tasks may be a result of formal lessons on an instrument. Finally, an effect of the original, recorded tempo of a musical stimulus was found in both a tapping to imagined music and a perception task, indicating that faster tempi songs were recalled more veridically in both tasks. This factor has never been accounted for in previous research and could serve to at least partially

explain performance differences found between different studies and different stimuli used within previous imagined tempo research. Taken together, these findings highlight the need for researchers to take into account the previous experience of participants, in terms of both formal and informal musical expertise, as well as properties of a musical stimulus itself, when administering both perceived and imagined musical tempo tasks.

The parallels revealed by the present work in terms of the content and interactions with endogenous states between VMI and INMI suggest that these two types of imagery share similar underlying mechanisms. It could be posited that VMI and INMI differ primarily in terms of the retrieval mechanisms implicated in each type of imagery, while the content and maintenance of the two types of imagery appear to be highly similar once an image has been generated. An analogous finding has been reported in research on episodic memories, in that different retrieval networks were implicated in the brain for involuntary compared to voluntary episodic memories, but similar regions were activated during the maintenance of both types of memories (Hall et al., 2014). Such an idea could be tested future in research on musical imagery, in particular as a potentially informative approach to neuroscientific investigations of the neural mechanisms underlying VMI versus INMI.

The results of Study 4 also provide evidence that musical features of INMI can interact with affective states—in this case subjective arousal—in a similar manner to features of perceived music. As such, further research should investigate whether musical imagery (both voluntary and involuntary) can serve in similar mood regulatory capacities to perceived music (Juslin & Laukka, 2004; Saarikallio & Erkkilä, 2007; Sloboda et al., 2001; Tarrant et al., 2000). Such findings could, for instance, help to inform potential coping strategies in everyday situations requiring mood regulation, as musical imagery is a mental process that can be activated in any environment for free and is relatively easy for both musicians and non-musicians to generate.

6.3.2 Implications for Musicians

In addition to its contributions to the psychological literature, the research presented in this thesis has several implications within the realm of music performance. The results of Study 1 indicate that physiological arousal can affect both perceived and imagined musical tempo judgements. Such a finding highlights the importance of effective musical performance anxiety management. As anxiety during performance is often accompanied by an increase in arousal (LeBlanc et al., 1997) and musicians are often expected to maintain steady and consistent tempi in performance and audition situations, strategies for maintaining an optimal level of arousal could be beneficial to training programs that aim to deal with performance anxiety. The findings of Study 2, though less clear, suggest that ageing can also affect at least imagined tempo representations, which may translate to the realm of anticipatory imagery generated before and during a musical performance. Further research, including longitudinal as well as cross-sectional work, on the effect of a performer's age on the consistency of tempo across multiple performances and changes in performance tempo throughout the course of one's life could be fruitful in understanding and combating any age-related effects on the temporal precision of musical performances. Finally, Study 3 has revealed that musical training and musical engagement can increase the veridicality of imagined recall of familiar music from within long-term memory. Such findings suggest that representations of tempo within VMI can be refined with more musical experience and training, although a definite causal link cannot be identified using the present study design.

The findings from Study 5 have potential applications for music composers and advertisers, in terms of revealing a handful of musical features that increase the likelihood of a tune to become INMI. As some composers and jingle writers aim to create songs that return quite easily and spontaneously to the mind, the findings that both the commonness of

musical contour features and the tempo of songs can influence their likelihood of becoming INMI could be of use in aiding deliberate compositional strategies. However, additional experimental research should also be conducted to further explore whether music specifically chosen to exhibit these compositional features returns to the mind as INMI more easily or frequently than other music.

6.4 Limitations of the Research

Various limitations of the work in this thesis have already been addressed throughout the course of the individual chapters. However, a couple broader points will be discussed here in relation to musical imagery research in general.

It should be noted that tasks used to measure imagined tempo may be subject to a certain degree of measurement error and, therefore, might not always capture the exact tempo of one's mental representation. For instance, the tapping to imagined music tasks utilized in Studies 3 and 4 rely on motor constraints of the participant, whereas the click track adjustment task utilized in Studies 1, 2, and 3 creates a sound that may result in a degree of auditory interference with the mental image. Additionally, both of these imagined tempo tasks rely on a participant's ability to extract a beat from imagined music, which may present a challenge to some participants. Further development of methods to measure "pure" imagined musical representations (imagery that is not reinforced by auditory feedback such as singing) is necessitated to reduce measurement error as much as possible and create tasks that are easy and understandable for all participants.

In regard to INMI research in particular, one limitation that presents various difficulties is the issue of intentionality (voluntary versus involuntary retrieval of an image). Although there are some instances in which a tune clearly comes into one's mind completely "out of the blue"

and some instances in which a person quite purposefully brings a song to mind that he/she desires to imagine, there exist other situations in which participants report that it can be difficult to judge whether a tune was initially recalled via voluntary or involuntary means. This leads to the question as to how useful it is to study VMI and INMI in entirely separate contexts, as has been done in most previous literature (see Bailes, 2006, 2007a, 2015 and Beaty et al., 2013 for exceptions). Instead, it might be useful in future research to study both types of imagery in conjunction (in both laboratory and naturalistic contexts). The degree of intentionality with which a musical image was recalled could be measured on a continuous scale in this research, in order to further investigate the proportion of episodes for which an image is clearly voluntarily or involuntarily recalled, as well as how often intentionality is less clear.

6.5 Future Directions

The results of the present work open a variety of avenues for future research on musical imagery. The inconclusive findings of Study 2 leave open several questions regarding the effects of ageing on the representation of music within imagery. To the author's knowledge, no previous research has systematically investigated the musical imagery capacities of younger versus older adults and, as such, it is yet to be determined whether imagery representations change with age in regard to vividness, consistency, etc. It should be restated here that the brain regions implicated in VMI are similar but not entirely equivalent to those implicated during music perception (e.g., Zatorre & Halpern, 2005). In particular, frontal cortical activations have been found during VMI, which are likely related to the memory component inherent in bringing a tune to mind from within long-term memory (Herholz et al., 2012; Zatorre et al., 1996, 2010). As previous neuroscientific findings indicate that the volume of the frontal cortex can decrease with age (e.g., Raz et al., 1997; Tisserand et al., 2002) and that this can affect memory retrieval processes (Schacter, Savage, Alpert, Rauch, & Albert, 1996; West,

1996), future research should investigate the influence of age-related changes in frontal cortical volume on both the voluntary and involuntary recall of musical imagery, as well as the maintenance and rehearsal of these images within working memory. Such research could administer both self-report questionnaires on imagery vividness and control (e.g., using the BAIS) and objective imagery tasks, such as the mental melody reversal task of Zatorre et al. (2010), in order to investigate the relationships between performance on musical imagery measures, performance on working memory tasks, and age-related changes in frontal cortical volume. Future research should also explore the relationship between age and INMI, in terms of both the developmental trajectory of the INMI experience during childhood as well as the prevalence and evaluation of INMI in older adults. At the time of writing, research on this aspect of the INMI experience is virtually non-existent and could provide valuable insights regarding the prevalence, functions, and degree of interference attributed to involuntary musical memories across the lifespan.

The findings of Study 3 of this thesis indicate that musical background, in terms of both formal training and active musical engagement, was related to imagery task performance. However, as mentioned in Section 6.3, no definite causal link can be established from these findings to indicate that receiving training and/or engaging with music can subsequently result in more accurate performance on imagery tasks. In light of the emphasis of many music education curricula on the importance of being able to deliberately imagine a piece of music before playing it (Bailes, 2002; Gordon, 1999, 2003), it is key that future research investigates whether training in music and/or imagery tasks can directly result in more precise representations of music within VMI. Such research could employ a longitudinal design in which imagery abilities are tested at different time points throughout a training period. This research could be conducted in collaboration with school or university music educators, in order to collaboratively assess and develop effective curricula for training imagery abilities in student musicians.

Another related question raised by the results of Study 3 is whether participants with higher musical engagement/training scores performed more accurately on the imagery tasks than participants with less musical experience because they were truly generating more veridical mental images or because these imagery tasks somehow favoured more musically experienced participants (e.g., due to the necessity to be able to extract a beat from imagined music). The development of a battery of additional tests to measure imagined tempo representations in different ways could aid in providing further insight into this question. Indeed, the findings of Study 3 highlight the general need to develop more standardized measures for assessing various aspects of VMI content, in order to allow for comparability and generalizability across different studies of the same musical dimension, such as imagined tempo.

The results of Studies 4 and 5 indicate that objective investigations of the musical content of INMI can be useful in furthering our understanding of the genesis and functions of the experience. The development of new methods for probing additional musical content of the INMI experience (e.g., pitch, loudness) during the course of everyday life could provide further insights into why certain songs become INMI and how the musical features of INMI interact with both internal states and environmental surroundings. Such work could employ voice recorders into which participants sing the pitches of their INMI or a perceptual stimulus that can be adjusted to match the pitch, loudness, or timbre of one's mental image. This research could test such questions as whether the absolute pitch level of tunes are preserved within INMI, whether the perceived loudness of an INMI episode affects ratings of interference with a concurrent task, and whether songs in a higher or lower pitch range relate to aspects of concurrent mood and arousal. Additionally, the work presented in Study 5 should be extended in order to explore how additional features of a tune, such as the lyrics, harmony, and acoustic features, might influence its likelihood to become INMI.

Finally, future research should be implemented to compare VMI and INMI within the same participants. Such a study could reveal the clearest evidence to date regarding the parallels between VMI and INMI, such as whether emotional reactions to the same song within both types of imagery are equivalent and whether the same song is recalled with equal levels of precision in both VMI and INMI. A potential design could involve 1) collecting tempo measurements and affective evaluations of INMI episodes over the course of several days and then 2) bringing participants into the laboratory for a session in which they voluntarily imagine, tap to the beat, and affectively evaluate the same songs that they had previously experienced as INMI.

6.6 Conclusion

The work conducted for this thesis has produced several novel findings regarding the representation of music within imagery. This research has revealed influences of arousal, age, musical experience, song tempo, and familiarity on perceived and imagined tempo judgements for familiar music. A number of parallels were found between the precision of tempo representations within VMI and INMI, and a handful of musical features were uncovered that served to increase the likelihood that a tune would become INMI. The research in this thesis contributes to furthering our understanding as to how consistent and precise mental representations of music are generated, the role of individual differences in the ability to mentally represent music, how imagined music interacts with endogenous states, and how certain intrinsic features of the music itself can contribute to the generation of an imagery experience.

References

Aleman, A., Nieuwenstein, M.R., Böcker, K.B., & de Haan, E. H. (2000). Music training and mental imagery ability. *Neuropsychologia, 38*(12), 1664–1668. doi:10.1016/S0028-3932(00)00079-8

Baddeley, A.D., & Andrade, J. (2000). Working memory and the vividness of imagery. *Journal of Experimental Psychology: General, 129*(1), 126–145. doi:10.1037/0096-3445.129.1.126

Baddeley, A., & Logie, R. (1992). Auditory imagery and working memory. In D. Reisberg (Ed.), *Auditory imagery* (pp. 179–197). Hillsdale, NJ: Lawrence Erlbaum Associates.

Bailes, F.A. (2002). *Musical imagery : Hearing and imagining music.* PhD thesis, University of Sheffield. Retrieved from http://etheses.whiterose.ac.uk/3452/

Bailes, F. (2006). The use of experience-sampling methods to monitor musical imagery in everyday life. *Musicae Scientiae, 10*, 173–190. doi:10.1177/102986490601000202

Bailes, F. (2007a). The prevalence and nature of imagined music in the everyday lives of music students. *Psychology of Music, 35*, 555–570. doi:10.1177/0305735607077834

Bailes, F. (2007b). Timbre as an elusive component of imagery for music. *Empirical Musicology Review, 2*, 21–34.

Bailes, F. (2010). Dynamic melody recognition: Distinctiveness and the role of musical expertise. *Memory & Cognition, 38*(5), 641–50. doi:10.3758/MC.38.5.641

Bailes, F. (2012). Arousal, valence, and the involuntary musical image. In: Cambouropoulos, E., Tsougras, C., Mavromatis, K., Pastiadis, K., editors. *Proceedings of ICMPC-ESCOM 12* (Thessaloniki, Greece), 86.

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., & Barwick, L. (2011). Absolute tempo in multiple performances of Aboriginal songs: Analyzing recordings of Djanba 12 and Djanba 14. *Music Perception*, *28*(5), 473–491.

Bangert, M., Peschel, T., Schlaug, G., Rotte, M., Drescher, D., Hinrichs, H., Heinze, H., & Altenmüller, E. (2006). Shared networks for auditory and motor processing in professional pianists: Evidence from fMRI conjunction. *NeuroImage*, *30*(3), 917–26. doi:10.1016/j.neuroimage.2005.10.044

Baudouin, A., Vanneste, S., Isingrini, M., & Pouthas, V. (2006). Differential involvement of internal clock and working memory in the production and reproduction of duration: A study on older adults. *Acta Psychologica*, *121*, 285–296. doi:10.1016/j.actpsy.2005.07.004

Beaman, C.P., Powell, K., & Rapley, E. (2015). Want to block earworms from conscious awareness? B(u)y gum! *Quarterly Journal of Experimental Psychology*, *68*(6), 1049–57. doi:10.1080/17470218.2015.1034142

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. doi:10.1348/000712609X479636

Beaman, C.P., & Williams, T.I. (2013). Individual differences in mental control predict involuntary musical imagery. *Musicae Scientiae*, *17*, 398–409. doi:10.1177/1029864913492530

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*(4), 1163–1173. doi:10.1016/j.concog.2013.07.006

Bennett, S. (2002). *Musical Imagery Repetition (MIR)*. M. Phil., Cambridge, United Kingdom. Retrieved from www.zingle.org/ Musical_Imagery_Repetition/

Benoit, C.-E., Dalla Bella, S., Farrugia, N., Obrig, H., Mainka, S., & Kotz, S. A. (2014). Musically cued gait-training improves both perceptual and motor timing in Parkinson's disease. *Frontiers in Human Neuroscience*, *8*, 494. doi:10.3389/fnhum.2014.00494

Bergeson, T.R., & Trehub, S.E. (2002). Absolute pitch and tempo in mothers' songs to infants. *Psychological Science*, *13*(1), 72–75. doi:10.1111/1467-9280.00413

Berntsen, D. (1998). Voluntary and involuntary access to autobiographical memory. *Memory*, *6*(2), 113–41. doi:10.1080/741942071

Berntsen, D., & Hall, N.M. (2004). The episodic nature of involuntary autobiographical memories. *Memory & Cognition*, *32*(5), 789–803. doi:10.3758/BF03195869

Bertrand, F., Magnanensi, J., Meyer, N., & Maumy-Bertrand, M. (2014). plsRglm: Algorithmic insights and applications. Technical Documetation. Retrieved from https://cran.r-project.org/web/packages/plsRglm/vignettes/plsRglm.pdf

Bharucha, J.J. (1984). Anchoring effects in music: The resolution of dissonance. *Cognitive Psychology, 16*(4), 485-518. doi:10.1016/0010-0285(84)90018-5

Bharucha, J., & Krumhansl, C.L. (1983). The representation of harmonic structure in music: Hierarchies of stability as a function of context. *Cognition, 13*(1), 63–102. doi:10.1016/0010-0277(83)90003-3

Bishop, L., Bailes, F., & Dean, R.T. (2013). Musical expertise and the ability to imagine loudness. *PLoS ONE, 8*(2). doi:10.1371/journal.pone.0056052

Bishop, L., Bailes, F., & Dean, R.T. (2014). Performing musical dynamics. *Music Perception*, *32*(1), 51–66. doi:10.1525/mp.2014.32.1.51

Block, R.A., Zakay, D., & Hancock, P.A. (1998). Human aging and duration judgements: A meta-analytic review. *Psychology & Aging*, *13*(4), 584–596. doi:10.1037/0882-7974.13.4.584

Boltz, M.G. (1994). Changes in internal tempo and effects on the learning and remembering of event durations. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *20*(5), 1154–1171. doi:10.1037/0278-7393.20.5.1154

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5–32. doi:10.1023/A:1010933404324

Brodsky, W., Henik, A., Rubinstein, B.-S., & Zorman, M. (2003). Auditory imagery from musical notation in expert musicians. *Perception & Psychophysics*, *65*(4), 602–612. doi:10.3758/BF03194586

Brown, R.M., & Palmer, C. (2012). Auditory–motor learning influences auditory memory for music. *Memory & Cognition, 40*(4), 567-78. doi:10.3758/s13421-011-0177-x

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

Burgoyne, J.A., Bountouridis, D., Van Balen, J., & Honing, H. (2013). Hooked: A game for discovering what makes music catchy. In *Proceedings of the International Society for Music Information Retrieval (ISMIR)*, 245-250.

Burns, G. (1987). A typology of "hooks" in popular records. *Popular Music, 6*(1), 1–20. doi:10.1017/S0261143000006577

Byron, T.P., & Fowles, L.C. (2013). Repetition and recency increases involuntary musical imagery of previously unfamiliar songs. *Psychology of Music, 43*(3), 375-89. doi:10.1177/0305735613511506

Callard, F., Smallwood, J., Golchert, J., & Margulies, D.S. (2013). The era of the wandering mind? Twenty-first century research on self-generated mental activity. *Frontiers in Psychology, 4*, 1–11. doi:10.3389/fpsyg.2013.00891

Chen, J.L., Penhune, V.B., & Zatorre, R.J. (2008). Moving on time: Brain network for auditory-motor synchronization is modulated by rhythm complexity and musical training. *Journal of Cognitive Neuroscience, 20*(2), 226–39. doi:10.1162/jocn.2008.20018

Clark, D.A., & Rhyno, S. (2005). Unwanted intrusive thoughts in nonclinical individuals: Implications for clinical disorders. In D.A. Clark (Ed.), *Intrusive thoughts in clinical disorders: Theory, research, and treatment* (pp. 1–29). New York: The Guilford Press.

Clark, T., Williamon, A., & Aksentijevic, A. (2011). Musical imagery and imagination: The function, measurement, and application of imagery skills for performance. In D. Hargreaves, D. Miell, & R. MacDonald (Eds.). *Musical imaginations: Multidisciplinary persepectives on creativity, performance and perception* (pp. 351-368). Oxford: Oxford University Press.

Clynes, M., & Walker, J. (1982). Neurobiologic functions of rhythm, time, and pulse in music. In M. Clynes (Ed.), *Music, mind, and brain: The neuropsychology of music* (pp.171-216)*.* New York: Plenum.

Clynes, M., & Walker, J. (1986). Music as time's measure. *Music Perception, 4*(1), 85–120.

Coffman, D.D. (1990). Effects of mental practice, physical practice, and knowledge of results on piano performance. *Journal of Research in Music Education, 38*(3), 187-96. doi:10.2307/3345182

Collier, G.L., & Collier, J.L. (1994). An exploration of the use of tempo in jazz. *Music Perception, 11*(3), 219–242.

Connolly C. (2002). Mental skills to optimize musical performance. In Stevens, C., Burnham, D., McPherson, G., Schubert, E., & Renwick,

J., editors, *Proceedings of the 7th International Conference on Music Perception and Cognition*. Adelaide: Casual Productions, 97-100.

Conover, W.J., & Iman, R.L. (1981). Rank transformations as a bridge between parametric and nonparametric statistics. *The American Statistician, 35*(3), 124–129.

Covach, J. (2005). Form in rock music: A primer. In D. Stein (Ed.), *Engaging music: Essays in music analysis* (pp 65-76). New York: Oxford University Press.

Craik, F.I.M, & Hay, J.F. (1991). Aging and judgements of duration: Effects of task complexity and method of estimation. *Perception & Psychophysics, 61*(3), 549–560. doi:10.3758/BF03211972

Crowder, R.G. (1989). Imagery for musical timbre. *Journal of Experimental Psychology: Human Perception and Performance, 15*(3), 472-8. doi:10.1037/0096-1523.15.3.472

Cuddy, L.L., Cohen, A.J., & Miller, J. (1979). Melody recognition: The experimental application of musical rules. *Canadian Journal of Psychology, 33*(3), 148–157. doi:10.1037/h0081713

Dhanaraj, R., & Logan, B. (2005). Automatic prediction of hit songs. *Proceedings of the International Symposium on Music Information Retrieval (ISMIR),* pp. 488–491.

Drake, C., Jones, M.R., & Baruch, C. (2000). The development of rhythmic attending in auditory sequences: Attunement, referent period, focal attending. *Cognition, 77*(3), 251–288. doi:10.1016/S0010-0277(00)00106-2

Droit-Volet, S., & Wearden, J. (2002). Speeding up an internal clock in children? Effects of visual flicker on subjective duration. *The Quarterly Journal of Experimental Psychology, 55*(3), 193–211.

Edworthy, J., & Waring, H. (2006). The effects of music tempo and loudness level on treadmill exercise. *Ergonomics, 49*(15), 1597–1610.

Espinosa-Fernández, L., Miró, E., Cano, M., & Buela-Casal, G. (2003). Age-related changes and gender differences in time estimation. *Acta Psychologica, 112*(3), 221–232. doi:10.1016/S0001-6918(02)00093-8

Farah, M.J., & Smith, A.F. (1983). Perceptual interference and facilitation with auditory imagery. *Perception & Psychophysics, 33*(5), 475–478. doi:10.3758/BF03202899

Farrugia, N., Jakubowski, K., Cusack, R., & Stewart, L. (2015). Tunes stuck in your brain: The frequency and affective evaluation of involuntary musical imagery correlate with cortical structure. *Consciousness and Cognition, 35*, 66–77. doi:10.1016/j.concog.2015.04.020

Finke, R.A. (1980). Levels of equivalence in imagery and perception. *Psychological Review, 87*(2), 113–132. doi:10.1037/0033-295X.87.2.113

Finke, R.A. (1986). Some consequences of visualization in pattern identification and detection. *The American Journal of Psychology, 99*(2), 257–274.

Finkel, S., Jilka, S.R, Williamson, V.J., Stewart, L., & Müllensiefen, D. (2010). Involuntary musical imagery: Investigating musical features that predict earworms. Paper presented at the *Third International Conference of Students of Systematic Musicology (SysMus10)*, University of Cambridge, UK.

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*, 472–86. doi:10.1016/j.concog.2015.02.012

Floridou, G., Williamson, V.J., & Müllensiefen, D. (2012). Contracting earworms: The roles of personality and musicality. In Cambouropoulos, E., Tsougras, C., Mavromatis, K., Pastiadis, K., (Eds), *Proceedings of ICMPC-ESCOM 12* (pp. 302–310).

Thessaloniki, Greece.

Floridou, G.A., Williamson, V.J., & Stewart, L. (2014). How does cognitive load affect the appearance of involuntary musical imagery (INMI)? Inducing earworms in the lab. In: Jakubowski, K., Farrugia, N., Floridou, G.A., & Gagen, J., editors. *Proceedings of the 7th International Conference of Students of Systematic Musicology.* (London, UK).

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.

Friberg, A., & Sundberg, J. (1995). Time discrimination in a monotonic, isochronous sequence. *The Journal of the Acoustical Society of America*, *98*(5), 2524. doi:10.1121/1.413218

Friedman, J., Hastie, T., Simon, N., & Tibshirani, R. (2015). Package 'glmnet.' Technical Documentation. Retrieved from https://cran.r-project.org/web/packages/glmnet/glmnet.pdf

Frieler, K. (2005). Melody - csv file format (mcsv). Technical Documentation (unpublished).

Frieler, K., Fischinger, T., Schlemmer, K., Lothwesen, K., Jakubowski, K., & Mullensiefen, D. (2013). Absolute memory for pitch: A comparative replication of Levitin's 1994 study in six European labs. *Musicae Scientiae*, *17*(3), 334–349. doi:10.1177/1029864913493802

Frieler, K., Jakubowski, K. & Müllensiefen, D. (in press). Is it the song and not the singer? Hit song prediction using structural features of melodies. *Yearbook of Music Psychology (Jahrbuch Musikpsychologie).*

Gabrielsson, A. (2002). Emotion perceived and emotion felt: Same or different? *Musicae Scientiae, 5*, 123–147. doi:10.1177/10298649020050S105

Gagnon, L., & Peretz, I. (2003). Mode and tempo relative contributions to "happy-sad" judgements in equitone melodies. *Cognition & Emotion*, *17*(1), 25–40. doi:10.1080/02699930302279

Ganis, G., Thompson, W.L., & Kosslyn, S.M. (2004). Brain areas underlying visual mental imagery and visual perception: An fMRI study. *Cognitive Brain Research*, *20*(2), 226–241. doi:10.1016/j.cogbrainres.2004.02.012

Gibbon, J., Church, R.M., & Meck, W.H. (1984). Scalar timing in memory. *Annals of the New York Academy of Sciences*, *423*, 52–77. doi:10.1111/j.1749-6632.1984.tb23417.x

Gordon, E.E. (1999). All about audiation and music aptitudes. *Music Educators Journal, 86*(2), 41-44.

Gordon, E.E. (2003). *Learning sequences in music: Skill, content, and patterns.* Chicago: GIA Publications.

Gower, J.C. (1971). A General coefficient of similarity and some of its properties. *Biometrics*, *27*(4), 857–871. doi:10.2307/2528823

Grahn, J.A., & Brett, M. (2007). Rhythm and beat perception in motor areas of the brain. *Journal of Cognitive Neuroscience*, *19*(5), 893–906. doi:10.1162/jocn.2007.19.5.893

Grahn, J.A., & Rowe, J.B. (2009). Feeling the beat: Premotor and striatal interactions in musicians and nonmusicians during beat perception. *Journal of Neuroscience*, *29*(23), 7540–8. doi:10.1523/JNEUROSCI.2018-08.2009

Griffiths, T.D. (2000). Musical hallucinosis in acquired deafness: Phenomenology and brain substrate. *Brain, 123 (1)*, 2065–2076. doi:10.1093/brain/123.10.2065

Griffiths, T.D., & Warren, J.D. (2002). The planum temporale as a computational hub. *Trends in Neurosciences*, *25*(7), 348–353. doi:10.1016/S0166-2236(02)02191-4

Grommet, E.K., Droit-Volet, S., Gil, S., Hemmes, N.S., Baker, A.H., & Brown, B.L. (2011). Time estimation of fear cues in human observers. *Behavioural Processes*, *86*(1), 88–93. doi:10.1016/j.beproc.2010.10.003

Haddon, E. (2007). What does mental imagery mean to university music students and their professors? In *Proceedings of the International Symposium on Performance Science (ISPS)*, 301-306.

Hall, S.A., Rubin, D.C., Miles, A., Davis, S.W., Wing, E.A., Cabeza, R., & Berntsen, D. (2014). The neural basis of involuntary episodic memories. *Journal of Cognitive Neuroscience*, *26*(10), 2385–99. doi:10.1162/jocn_a_00633

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

Halpern, A.R. (1988b). Perceived and imagined tempos of familiar songs. *Music Perception*, *6*(2), 193–202.

Halpern, A.R. (1989). Memory for the absolute pitch of familiar songs. *Memory & Cognition*, *17*(5), 572–581. doi:10.3758/BF03197080

Halpern, A.R. (1992). Musical aspects of auditory imagery. In D. Reisberg (Ed.), *Auditory imagery* (pp. 1–27). Hillsdale, NJ: Erlbaum.

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., & Bartlett, J.C. (2011). The persistence of musical memories: A descriptive study of earworms, *Music Perception, 28*(4), 425–432.

Halpern, A.R., Zatorre, R.J., Bouffard, M., & Johnson, J.A. (2004). Behavioral and neural correlates of perceived and imagined musical timbre. *Neuropsychologia, 42*(9), 1281–1292. doi:10.1016/j.neuropsychologia.2003.12.017

Herholz, S.C., Halpern, A.R., & Zatorre, R.J. (2012). Neuronal correlates of perception, imagery, and memory for familiar tunes. *Journal of Cognitive Neuroscience, 24*(6), 1382-97. doi:10.1162/jocn_a_00216

Highben, Z., & Palmer, C. (2004). Effects of auditory and motor mental practice in memorized piano performance. *Bulletin of the Council for Research in Music Education*, 159, 58–67.

Hoerl, A.E., & Kennard, R.W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics, 12*(1), 55–67.

Hothorn, T. Hornik, K., & Zeileis, A. (2006). Unbiased recursive partitioning: A conditional inference framework. *Journal of Computational and Graphical Statistics, 15*(3), 651-674.

Hubbard, T.L. (2010). Auditory imagery: Empirical findings. *Psychological Bulletin, 136*(2), 302–329. doi:10.1037/a0018436

Hubbard, T.L., & Stoeckig, K. (1988). Musical imagery: Generation of tones and chords. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*(4), 656-67. doi:10.1037/0278-7393.14.4.656

Huron, D. (1996). The melodic arch in Western folksongs. *Computing in Musicology, 10*, 3-23.

Huron, D.B. (2006). *Sweet anticipation: Music and the psychology of expectation*. Cambridge, MA: MIT Press.

Huron, D. (2011). Why is sad music pleasurable? A possible role for prolactin. *Musicae Scientiae, 15*(2), 146–158. doi:10.1177/1029864911401171

Husain, G., Thompson, W.F., & Schellenberg, E.G. (2002). Effects of musical tempo and mode on arousal, mood, and spatial abilities. *Music Perception, 20*(2), 151–171. doi:10.1525/mp.2002.20.2.151

Hyman, I.E., Burland, N.K., Duskin, H.M., Cook, M.C., Roy, C.M., McGrath, J.C., & Roundhill, R.F. (2013). Going Gaga: Investigating,

creating, and manipulating the song stuck in my head. *Applied Cognitive Psychology*, *27*, 204–215. doi:10.1002/acp.2897

Hyman, I.E., Cutshaw, K.I., Hall, C.M., Snyders, M.E., Masters, S.A., Au, V.S.K., & Graham, J.M. (2015). Involuntary to intrusive: Using involuntary musical imagery to explore individual differences and the nature of intrusive thoughts. *Psychomusicology: Music, Mind and Brain*, *25*(1), 14–27.

Intons-Peterson, M.J. (1980). The role of loudness in auditory imagery. *Memory & Cognition*, *8*(5), 385–393. doi:10.3758/BF03211134

Intons-Peterson, M.J. (1992). Components of auditory imagery. In D. Reisberg (Ed.), *Auditory imagery* (pp. 45-72). Hillsdale, NJ: Lawrence Erlbaum Associates.

Iwanaga, M. (1995). Relationship between heart rate and preference for tempo of music. *Perceptual and Motor Skills*, *81*(2), 435–40. doi:10.2466/pms.1995.81.2.435

Jakubowski, K., Halpern, A.R., Grierson, M., & Stewart, L. (2015). The effect of exercise-induced arousal on chosen tempi for familiar melodies. *Psychonomic Bulletin & Review*, *22*(2), 559–565.

Jakubowski, K., & Müllensiefen, D. (2013). The influence of music-elicited emotions and relative pitch on absolute pitch memory for familiar melodies. *Quarterly Journal of Experimental Psychology*, *66*(7), 1259–67. doi:10.1080/17470218.2013.803136

Janata, P. (2001). Neuropsychlogical mechanisms underlying auditory image formation. In R. Godoy & H. Jorgensen (Eds.), *Musical imagery* (pp. 27-42). New York: Taylor & Francis.

Janata, P. (2012). Acuity of mental representations of pitch. *Annals of the New York Academy of Sciences*, *1252*, 214–21. doi:10.1111/j.1749-6632.2011.06441.x

Janata, P., & Paroo, K. (2006). Acuity of auditory images in pitch and time. *Perception & Psychophysics*, *68*(5), 829–844. doi:10.3758/BF03193705

Jeannerod, M. (1995). Mental imagery in the motor context. *Neuropsychologia, 33*(11)*,* 1419–1432. doi:10.1016/0028-3932(95)00073-C

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. doi:10.1177/1321103X11400501

Juslin, P.N., & Laukka, P. (2004). Expression, perception, and induction of musical emotions: A review and a questionnaire study of everyday listening. *Journal of New Music Research*, *33*(3), 217–238. doi:10.1080/0929821042000317813

Kalakoski, V. (2001). Musical imagery and working memory. In R. Godoy & H. Jorgensen (Eds.), *Musical imagery* (pp. 43-55). New York: Taylor & Francis.

Keller, P. E. (2012). Mental imagery in music performance: Underlying mechanisms and potential benefits. *Annals of the New York Academy of Sciences*, *1252*, 206–13. doi:10.1111/j.1749-6632.2011.06439.x

Keller, P.E., & Appel, M. (2010). Individual differences, auditory imagery, and the coordination of body movements and sounds in musical ensembles. *Music Perception*, *28*(1), 27–46. doi:10.1525/mp.2010.28.1.27

Keller, P.E., Dalla Bella, S., & Koch, I. (2010). Auditory imagery shapes movement timing and kinematics: Evidence from a musical task. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(2), 508–513. doi:10.1037/a0017604

Keller, P.E., & Koch, I. (2006). The planning and execution of short auditory sequences. *Psychonomic Bulletin & Review*, *13*(4), 711–716. doi:10.3758/BF03193985

Keller, P.E., & Koch, I. (2008). Action planning in sequential skills: Relations to music performance. *Quarterly Journal of Experimental Psychology*, *61*(2), 275–291. doi:10.1080/17470210601160864

Kleber, B., Birbaumer, N., Veit, R., Trevorrow, T., & Lotze, M. (2007). Overt and imagined singing of an Italian aria. *NeuroImage*, *36*(3), 889–900. doi:10.1016/j.neuroimage.2007.02.053

Kopiez, R. & Müllensiefen, D. (2011). Auf der Suche nach den „Popularitätsfaktoren" in den Song-Melodien des Beatles-Albums Revolver: eine computergestützte Feature-Analyse. In S. Meine & N. Noeske (Eds.), *Musik und Popularität. Beiträge zu einer Kulturgeschichte zwischen 1500 und heute* (pp. 207–225). Münster: Waxmann Verlag.

Kosslyn, S.M. (1980). *Image and Mind*. Cambridge, MA: Harvard University Press.

Kosslyn, S.M., Ball, T.M., & Reiser, B.J. (1978). Visual images preserve metric spatial information: Evidence from studies of image scanning. *Journal of Experimental Psychology: Human Perception and Performance*, *4*(1), 47–60. doi:10.1037/0096-1523.4.1.47

Kozlov, M.D., Hughes, R.W., & Jones, D.M. (2012). Gummed-up memory: Chewing gum impairs short-term recall. *Quarterly Journal of Experimental Psychology*, *65*(3), 501–13. doi:10.1080/17470218.2011.629054

Kraemer, D.J.M., Macrae, C.N., Green, A.E., & Kelley, W.M. (2005). Musical imagery: Sound of silence activates auditory cortex. *Nature*, *434*, 158. doi:10.1038/434158a

Kvavilashvili, L., & Anthony, S. (2012). When do Christmas songs pop into your mind? Testing a long-term priming hypothesis. Poster

presented at the Annual Meeting of Psychonomic Society, Minneapolis, Minnesota, US.

Kvavilashvili, L., & Mandler, G. (2004). Out of one's mind: A study of involuntary semantic memories. *Cognitive Psychology*, *48*(1), 47–94. doi:10.1016/S0010-0285(03)00115-4

Larson, R., & Csikszentmihalyi, M. (1983). The Experience Sampling Method. *New Directions for Methodology of Social and Behavioral Science*, *15,* 41–56. doi:10.1177/0894439305281844

LeBlanc, A., Jin, Y.C., Obert, M., & Siivola, C. (1997). Effect of audience on music performance anxiety. *Journal of Research in Music Education*, *45*(3), 480–496. doi:10.2307/3345541

Levitin, D.J. (1994). Absolute memory for musical pitch: Evidence from the production of learned melodies. *Perception & Psychophysics*, *56*(4), 414–423. doi:10.3758/BF03206733

Levitin, D.J. (2006). *This is your brain on music.* New York: Dutton.

Levitin, D.J., & Cook, P.R. (1996). Memory for musical tempo: Additional evidence that auditory memory is absolute. *Perception & Psychophysics*, *58*(6), 927–935. doi:10.3758/BF03205494

Liikkanen, L.A. (2008). Music in everymind: Commonality of involuntary musical imagery. *Proceedings of the 10th International Conference on Music Perception and Cognition (ICMPC10)*, Sapporo, Japan. Retrieved from http://l.kryptoniitti.com/lassial/files/publications/080904-Music_in_everymind_pdf.pdf

Liikkanen, L.A. (2012a). Musical activities predispose to involuntary musical imagery. *Psychology of Music*, *40,* 236–256. doi:10.1177/0305735611406578

Liikkanen, L.A. (2012b). Inducing involuntary musical imagery: An experimental study. *Musicae Scientiae*, *16*(2), 217–234. doi:10.1177/1029864912440770

Liikkanen, L.A., Jakubowski, K. & Toivanen, J. (in press). Catching earworms on Twitter: Using big data to study involuntary musical imagery. *Music Perception.*

Loh, W.Y. (2011). Classification and regression trees. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *1*(1), 14–23. doi:10.1002/widm.8

Lucas, B.J., Schubert, E., & Halpern, A.R. (2010). Perception of emotion in sounded and imagined music. *Music Perception, 27*(5), 399-412. doi:10.1525/mp.2010.27.5.399

Mace, J.H. (2006). Episodic remembering creates access to involuntary conscious memory: Demonstrating involuntary recall on a voluntary recall task. *Memory*, *14*(8), 917–24. doi:10.1080/09658210600759766

Manning, F., & Schutz, M. (2013). "Moving to the beat" improves timing perception. *Psychonomic Bulletin & Review*, *20*(6), 1133–1139. doi:10.3758/s13423-013-0439-7

McAuley, J.D., Jones, M.R., Holub, S., Johnston, H.M., & Miller, N.S. (2006). The time of our lives: Life span development of timing and event tracking. *Journal of Experimental Psychology: General*, *135*(3), 348–367. doi:10.1037/0096-3445.135.3.348

McDowd, J.M., & Craik, F.I. (1988). Effects of aging and task difficulty on divided attention performance. *Journal of Experimental Psychology: Human Perception and Performance*, *14*(2), 267–280. doi:10.1038/ng749

McKinney, M.F., & Moelants, D. (2006). Ambiguity in tempo perception: What draws listeners to different metrical levels? *Music Perception*, *24*(2), 155–166. doi:10.1525/mp.2006.24.2.155

Meyer, L.B. (1956). *Emotion and meaning in music.* Chicago: The University of Chicago Press.

Miró, E., Cano, M.C., Espinosa-Fernández, L., & Buela-Casal, G. (2003). Time estimation during prolonged sleep deprivation and its relation to activation measures. *Human Factors*, *45*(1), 148–159. doi:10.1518/hfes.45.1.148.27227

Moelants, D. (2002). Preferred tempo reconsidered. In: Stevens, C., Burnham, D., McPherson, G., Schubert, E., & Renwick, J., editors. *Proceedings of the 7th International Conference on Music Perception and Cognition* Adelaide: Causal Productions, 580-583.

Mountain, R. (2001). Composers & imagery: Myths & realities. In R. Godoy & H. Jorgensen (Eds.), *Musical imagery* (pp. 271-288). New York: Taylor & Francis.

Mullahy, J. (1986). Specification and testing of some modified count data models. *Journal of Econometrics*, *33*(3), 341–365. doi:10.1016/0304-4076(86)90002-3

Müllensiefen, D. (2009). FANTASTIC: Feature ANalysis Technology Accessing STatistics (In a Corpus). Technical Report. URL: http://www.doc.gold.ac.uk/isms/m4s/FANTASTIC_docs.pdf.

Müllensiefen, D., Jones, R., Jilka, S., Stewart, L., & Williamson, V.J. (2014a). Individual differences predict patterns in spontaneous involuntary musical imagery, *Music Perception, 31*(4), 323–338.

Müllensiefen, D., Gingras, B., Musil, J., & Stewart, L. (2014b). The musicality of non-musicians: An index for assessing musical sophistication in the general population. *PloS One, 9*(2), e89642. doi:10.1371/journal.pone.0089642

Müllensiefen, D., & Halpern, A.R. (2014). The role of features and context in recognition of novel melodies. *Music Perception*, *31*(5), 418–435. doi:10.1525/mp.2014.31.5.418

Murray, N.P., & Russoniello, C. (2012). Acute physical activity on cognitive function: A heart rate variability examination. *Applied*

Psychophysiology and Biofeedback, *37*(4), 219–27.
doi:10.1007/s10484-012-9196-z

Navarro Cebrian, A., & Janata, P. (2010). Electrophysiological correlates of accurate mental image formation in auditory perception and imagery tasks. *Brain Research*, *1342*, 39–54. doi:10.1016/j.brainres.2010.04.026

Neisser, U. (1972). Changing conceptions of imagery. In P.W. Sheehan & J.S. Antrobus (Eds.), *The function and nature of imagery* (pp. 233-251). New York: Academic Press.

Ni, Y., Santos-Rodríguez, R., Mcvicar, M., & De Bie, T. (2011). Hit song science once again a science? *4th International Workshop on Machine Learning and Music.*

North, A.C., & Hargreaves, D.J. (2000). Musical preferences during and after relaxation and exercise. *American Journal of Psychology*, *113*(1), 43–67. doi:10.2307/1423460

Okada, H., & Matsuoka, K. (1992). Effects of auditory imagery on the detection of a pure-tone in white-noise: Experimental evidence of the auditory Perky effect. *Perceptual and Motor Skills*, *74*(2), 443–448.

Ozel, S., Larue, J., & Dosseville, F. (2004). Effect of arousal on internal clock speed in real action and mental imagery. *Canadian Journal of Experimental Psychology*, *58*(3), 196–205.

Pachet, F., & Roy, P. (2008). Hit song science is not yet a science. In: *Proceedings of the International Symposium on Music Information Retrieval (ISMIR)*, 355–360.

Pastor, M.A., Artieda, J., Jahanshahi, M., & Obeso, J. A. (1992). Time estimation and reproduction is abnormal in Parkinson's disease. *Brain*, *115*(1), 211–225. doi:10.1093/brain/115.1.211

Patel, A.D. (2007). *Music, Language, and the Brain*. Oxford: Oxford University Press.

Pecenka, N., & Keller, P.E. (2009). Auditory pitch imagery and its relationship to musical synchronization. *Annals of the New York Academy of Sciences*, *1169,* 282–286. doi:10.1111/j.1749-6632.2009.04785.x

Perbal, S., Droit-Volet, S., Isingrini, M., & Pouthas, V. (2002). Relationships between age-related changes in time estimation and age-related changes in processing speed, attention, and memory. *Aging, Neuropsychology, and Cognition*, *9*(3), 201–216.

Pitt, M.A., & Crowder, R.G. (1992). The role of spectral and dynamic cues in imagery for musical timbre. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(3), 728–738. doi:10.1037/0096-1523.18.3.728

Ponds, R.W.H.M., Brouwer, W.H., & Van Wolffelaar, P.C. (1988). Age differences in divided attention in a simulated driving task. *Journal of Gerontology*, *43*(6), 151–156. doi:10.1093/geronj/43.6.P151

Pontifex, M.B., & Hillman, C.H. (2007). Neuroelectric and behavioral indices of interference control during acute cycling. *Clinical Neurophysiology*, *118*(3), 570–80. doi:10.1016/j.clinph.2006.09.029

Raz, N., Gunning, F.M., Head, D., Dupuis, J.H., McQuain, J., Briggs, S.D., Loken, W.J., Thornton, A.E., & Acker, J.D. (1997). Selective aging of the human cerebral cortex observed in vivo: Differential vulnerability of the prefrontal gray matter. *Cerebral Cortex*, *7*(3), 268–282. doi:10.1093/cercor/7.3.268

Reisberg, D., & Heuer, F. (2005). Visuospatial images. In P. Shah & A. Miyake (Eds.), *The Cambridge handbook of visuospatial thinking* (pp. 35-80). Cambridge, UK: Cambridge University Press.

Repp, B. H. (1999). Control of expressive and metronomic timing in pianists. *Journal of Motor Behavior*, *31*(2), 145–164. doi:10.1080/00222899909600985

Repp, B. H. (2001). Expressive timing in the mind's ear. In R. Godoy & H. Jorgensen (Eds.), *Musical imagery* (pp. 185-198). New York: Taylor & Francis.

Ross, S. L. (1985). The effectiveness of mental practice in improving the performance of college trombonists. *Journal of Research in Music Education*, *33*(4), 221. doi:10.2307/3345249

Royet, J.-P., Zald, D., Versace, R., Costes, N., Lavenne, F., Koenig, O., & Gervais, R. (2000). Emotional responses to pleasant and unpleasant olfactory, visual, and auditory stimuli: A Positron Emission Tomography study. *Journal of Neuroscience*, *20*(20), 7752–7759.

Saarikallio, S., & Erkkila, J. (2007). The role of music in adolescents' mood regulation. *Psychology of Music*, *35*(1), 88–109. doi:10.1177/0305735607068889

Sacks, O. (2007). *Musicophilia: Tales of music and the brain.* New York: Alfred A. Knopf.

Salthouse, T.A. (1991). Mediation of adult age differences in cognition by reductions in working memory and speed of processing. *Psychological Science*, *2*(3), 179–183. doi:10.1111/j.1467-9280.1991.tb00127.x

Salthouse, T.A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, *103*(3), 403–428.

Salthouse, T.A., & Babcock, R.L. (1991). Decomposing adult age differences in working memory. *Developmental Psychology*, *27*(5), 763–776.

Schacter, D.L., Savage, C.R., Alpert, N.M., Rauch, S.L., & Albert, M.S. (1996). The role of hippocampus and frontal cortex in age-related memory changes: A PET study. *Neuroreport*, *7*(6), 1165–9.

Schaefer, R.S. (2014). Images of time: Temporal aspects of auditory and movement imagination. *Frontiers in Psychology*, *5*, 877. doi:10.3389/fpsyg.2014.00877

Schaefer, R.S., Vlek, R.J., & Desain, P. (2011). Decomposing rhythm processing: Electroencephalography of perceived and self-imposed rhythmic patterns. *Psychological Research*, *75*(2), 95–106. doi:10.1007/s00426-010-0293-4

Schellenberg, E.G., & von Scheve, C. (2012). Emotional cues in American popular music: Five decades of the Top 40. *Psychology of Aesthetics, Creativity, and the Arts*, *6*(3), 196-203. doi:10.1037/a0028024

Schlagman, S., & Kvavilashvili, L. (2008). Involuntary autobiographical memories in and outside the laboratory: How different are they from voluntary autobiographical memories? *Memory & Cognition*, *36*(5), 920–932. doi:10.3758/MC.36.5.920

Schulkind, M.D., Posner, R.J., & Rubin, D.C. (2003). Musical features that facilitate melody identification: How do you know it's "your" song when they finally play it? *Music Perception*, *21*(2), 217–249. doi:10.1525/mp.2003.21.2.217

Schwarz, M.A., Winkler, I., & Sedlmeier, P. (2013). The heart beat does not make us tick: The impacts of heart rate and arousal on time perception. *Attention, Perception & Psychophysics*, *75*, 182–193. doi:10.3758/s13414-012-0387-8

Sekhon, J.S. (2011). Multivariate and propensity score matching software with automated balance optimization: The Matching package for R. *Journal of Statistical Software*, *42*(7), 1–52.

Shepard, R.N., & Cooper, L.A. (1982). *Mental images and their transformations.* Cambridge, MA: MIT Press.

Sloboda, J.A., O'Neill, S.A., & Ivaldi, A. (2001). Functions of music in everyday life: An exploratory study using the Experience Sampling Method. *Musicae Scientiae*, *5*(1), 9–32. doi:10.1177/102986490100500102

Smallwood, J., & Schooler, J.W. (2006). The restless mind. *Psychological Bulletin*, *132*(6), 946–958. doi:10.1037/0033-2909.132.6.946

Smallwood, J., & Schooler, J.W. (2015). The science of mind wandering : Empirically navigating the stream of consciousness, *Annual Review of Psychology, 66*, 487-518. doi:10.1146/annurev-psych-010814-015331

Smith, J.D., Wilson, M., & Reisberg, D. (1995). The role of subvocalization in auditory imagery. *Neuropsychologia*, *33*(11), 1433–1454. doi:10.1016/0028-3932(95)00074-D

Sowiński, J., & Dalla Bella, S. (2013). Poor synchronization to the beat may result from deficient auditory-motor mapping. *Neuropsychologia*, *51*(10), 1952–63. doi:10.1016/j.neuropsychologia.2013.06.027

Strobl, C., Boulesteix, A.-L., Zeileis, A., & Hothorn, T. (2007). Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC Bioinformatics*, *8*(1), 25. doi:10.1186/1471-2105-8-25

Strobl, C., Malley, J., & Tutz, G. (2009). An introduction to recursive partitioning: Rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychological Methods*, *14*(4), 323–348.

Tarrant, M., North, A.C., & Hargreaves, D.J. (2000). English and American adolescents' reasons for listening to music. *Psychology of Music*, *28*(2), 166–173. doi:10.1177/0305735600282005

Taylor, S., McKay, D., Miguel, E.C., De Mathis, M.A., Andrade, C., Ahuja, N., … Storch, E. A. (2014). Musical obsessions: A comprehensive review of neglected clinical phenomena. *Journal of Anxiety Disorders*, *28*(6), 580–9. doi:10.1016/j.janxdis.2014.06.003

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, *58*(1), 267–288.

Tisserand, D.J., Pruessner, J.C., Sanz Arigita, E.J., van Boxtel, M.P.J., Evans, A.C., Jolles, J., & Uylings, H.B.M. (2002). Regional frontal cortical volumes decrease differentially in aging: An MRI study to compare volumetric approaches and voxel-based morphometry. *NeuroImage*, *17*(2), 657–669. doi:10.1006/nimg.2002.1173

Trainor, L.J., Wu, L., & Tsang, C.D. (2004). Long-term memory for music: Infants remember tempo and timbre. *Developmental Science*, *7*(3), 289–296. doi:10.1111/j.1467-7687.2004.00348.x

Treisman, M. (1963). Temporal discrimination and the indifference interval: Implications for a model of the "internal clock". *Psychological Monographs: General and Applied*, *77*(13), 1–31.

Van Balen, J. Burgoyne, J.A., Bountouridis, D., Müllensiefen, D., & Veltkamp, R. (in press). Corpus analysis tools for computational hook discovery. *Proceedings of the International Symposium on Music Information Retrieval (ISMIR 2015)*.

Van Noorden, L., & Moelants, D. (1999). Resonance in the perception of musical pulse. *Journal of New Music Research*, *28*(1), 43–66.

Vanneste, S., Pouthas, V., & Wearden, J.H. (2001). Temporal control of rhythmic performance: A comparison between young and old adults. *Experimental Aging Research*, *27*(1), 83–102. doi:10.1080/03610730125798

Vuoskoski, J.K., Thompson, W.F., McIlwain, D., & Eerola, T. (2012). Who enjoys listening to sad music and why? *Music Perception*, *29*(3), 311–317. doi:10.1525/mp.2012.29.3.311

Wammes, M., & Barušs, I. (2009). Characteristics of spontaneous musical imagery. *Journal of Consciousness Studies*, *16*(1), 37–61.

Warren, J.E., Wise, R.J.S., & Warren, J.D. (2005). Sounds do-able: Auditory-motor transformations and the posterior temporal plane. *Trends in Neurosciences*, *28*(12), 636–43. doi:10.1016/j.tins.2005.09.010

Weber, R.J., & Brown, S. (1986). Musical imagery. *Music Perception*, *3*(4), 411–426. doi:10.2307/40285346

Webster, G.D., & Weir, C.G. (2005). Emotional responses to music: Interactive effects of mode, texture, and tempo. *Motivation and Emotion*, *29*(1), 19–39. doi:10.1007/s11031-005-4414-0

Wegner, D.M., Schneider, D.J., Carter, S.R., & White, T.L. (1987). Paradoxical effects of thought suppression. *Journal of Personality and Social Psychology*, *53*(1), 5–13. doi:10.1037/0022-3514.53.1.5

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.

West, R.L. (1996). An application of prefrontal cortex function theory to cognitive aging. *Psychological Bulletin*, *120*(2), 272–292.

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

Williamson, V.J. & Jakubowski, K. (2014). Earworms. In W. F. Thompson (Ed.), *Music in the Social and Behavioral Sciences*. Sage.

Williamson, V.J., & Jilka, S.R. (2013). Experiencing earworms: An interview study of Involuntary Musical Imagery. *Psychology of Music*, *42*, 653–670. doi:10.1177/0305735613483848

Williamson, V.J., Jilka, S., Fry, J., Finkel, S., Mullensiefen, D., & Stewart, L. (2012). How do earworms start? Classifying the everyday circumstances of Involuntary Musical Imagery, *Psychology of Music*, *40*(3), 259-84. doi:10.1177/0305735611418553

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).doi: 10.1371/journal.pone.0086170

Williamson, V.J. & Müllensiefen, D. (2012). Earworms from three angles. In: Cambouropoulos, E., Tsougras, C., Mavromatis, K., Pastiadis, K., editors. *Proceedings of ICMPC-ESCOM 12* (Thessaloniki, Greece), 1124–1133.

Wittmann, M., Leland, D.S., Churan, J., & Paulus, M.P. (2007). Impaired time perception and motor timing in stimulant-dependent subjects. *Drug and Alcohol Dependence*, *90*(2-3), 183–92. doi:10.1016/j.drugalcdep.2007.03.005

Wold, S., Ruhe, A., Wold, H., & Dunn, III, W.J. (1984). The collinearity problem in linear regression: The partial least squares (PLS) approach to generalized inverses. *SIAM Journal on Scientific and Statistical Computing*, *5*(3), 735–743. doi:10.1137/0905052

Wolpert, D.M., Ghahramani, Z., & Jordan, M.I. (1995). An internal model for sensorimotor integration. *Science*, *269*(5232), 1880–1882. doi:10.1126/science.7569931

Zajonc, R.B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology*, *9*(2), 1-27. doi:10.1037/h0025848

Zakay, D., & Block, R.A. (1996). The role of attention in time estimation processes. *Advances in Psychology*, *115*, 143–164. doi:10.1016/S0166-4115(96)80057-4

Zatorre, R.J., Chen, J.L., & Penhune, V.B. (2007). When the brain plays music: Auditory-motor interactions in music perception and production. *Nature Reviews Neuroscience*, *8*(7), 547–58. doi:10.1038/nrn2152

Zatorre, R.J., & Halpern, A.R. (1993). Effect of unilateral temporal-lobe excision on perception and imagery of songs. *Neuropsychologia*, *31*(3), 221–232. doi:10.1016/0028-3932(93)90086-F

Zatorre, R.J., & Halpern, A.R. (2005). Mental concerts: Musical imagery and auditory cortex. *Neuron*, *47*(1), 9-12. doi:10.1016/j.neuron.2005.06.013

Zatorre, R.J., Halpern, A.R., & Bouffard, M. (2010). Mental reversal of imagined melodies: A role for the posterior parietal cortex. *Journal of Cognitive Neuroscience*, *22*(4), 775–789. doi:10.1162/jocn.2009.21239

Zatorre, R.J., Halpern, A.R., Perry, D.W., Meyer, E., & Evans, A.C. (1996). Hearing in the mind's ear: A PET investigation of musical imagery and perception. *Journal of Cognitive Neuroscience*, *8*(1), 29-46. doi:10.1162/jocn.1996.8.1.29

Zeileis, A., Kleiber, C., & Jackman, S. (2007). Regression models for count data in R. Retrieved from http://epub.wu.ac.at/1168/

Zelaznik, H.N., & Rosenbaum, D.A. (2010). Timing processes are correlated when tasks share a salient event. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(6), 1565–1575.

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *67*(2), 301–320. doi:10.1111/j.1467-9868.2005.00503.

Appendix 1

Goldsmiths Musical Sophistication Index: Musical Training Subscale (Müllensiefen et al., 2014b)

Please circle the most appropriate category:	1 Completely Disagree	2 Strongly Disagree	3 Disagree	4 Neither Agree nor Disagree	5 Agree	6 Strongly Agree	7 Completely Agree
I have never been complimented for my talents as a musical performer.	1	2	3	4	5	6	7
I would not consider myself a musician.	1	2	3	4	5	6	7

Please circle the most appropriate category:

I engaged in regular, daily practice of a musical instrument (including voice) for **0 / 1 / 2 / 3 / 4-5 / 6-9 / 10 or more** years.

At the peak of my interest, I practiced **0 / 0.5 / 1 / 1.5 / 2 / 3-4 / 5 or more** hours per day on my primary instrument.

I have had formal training in music theory for **0 / 0.5 / 1 / 2 / 3 / 4-6 / 7 or more** years.

I have had **0 / 0.5 / 1 / 2 / 3-5 / 6-9 / 10 or more** years of formal training on a musical instrument (including voice) during my lifetime.

I can play **0 / 1 / 2 / 3 / 4 / 5 / 6 or more** musical instruments.

Appendix 2

Goldsmiths Musical Sophistication Index: Active Engagement Subscale (Müllensiefen et al., 2014b)

Please circle the most appropriate category:	1 Completely Disagree	2 Strongly Disagree	3 Disagree	4 Neither Agree nor Disagree	5 Agree	6 Strongly Agree	7 Completely Agree
I spend a lot of my free time doing music-related activities.	1	2	3	4	5	6	7
I enjoy writing about music, for example on blogs and forums.	1	2	3	4	5	6	7
I'm intrigued by musical styles I'm not familiar with and want to find out more.	1	2	3	4	5	6	7
I often read or search the internet for things related to music.	1	2	3	4	5	6	7

I don't spend much of my disposable income on music.	1	2	3	4	5	6	7
Music is kind of an addiction for me - I couldn't live without it.	1	2	3	4	5	6	7
I keep track of new of music that I come across (e.g. new artists or recordings).	1	2	3	4	5	6	7

Please circle the most appropriate category:

I have attended **0 / 1 / 2 / 3 / 4-6 / 7-10 / 11 or more** live music events as an audience member in the past twelve months.

I listen attentively to music for **0-15 min / 15-30 min / 30-60 min / 60-90 min / 2 hrs / 2-3 hrs / 4 hrs or more** per day.

Appendix 3

Bucknell Auditory Imagery Scale (Halpern, 2015)

Bucknell Auditory Imagery Vividness Scale

The following scale is designed to measure auditory imagery, or the way in which you "think about sounds in your head." For the following items you are asked to do the following:

Read the item and consider whether you think of an image of the described sound in your head. Then rate the vividness of your image using the following "Vividness Rating Scale." If no image is generated, give a rating of 1.

Please feel free to use all of the levels in the scale when selecting your ratings.

<u>Vividness Rating Scale</u>

1	2	3	4	5	6	7

No Image   Fairly   As Vivid As
Present At All   Vivid   Actual Sound

1. For the first item, consider the beginning of the song "Happy Birthday."
   The sound of a trumpet beginning the piece. _____

2. For the next item, consider ordering something over the phone.
   The voice of an elderly clerk assisting you. _____

3. For the next item, consider being at the beach.
   The sound of the waves crashing against nearby rocks. _____

4. For the next item, consider going to a dentist appointment.
   The loud sound of the dentist's drill. _____

5. For the next item, consider being present at a jazz club.
   The sound of a saxophone solo. _____

6. For the next item, consider being at a live cricket game.
   The cheer of the crowd as a player hits the ball. _____

7. For the next item, consider attending a choir rehearsal.
   The sound of an all-children's choir singing the first verse of a song. _____

8. For the next item, consider attending an orchestral performance of Beethoven's Fifth.
   The sound of the ensemble playing. _____

9. For the next item, consider listening to a rain storm.
   The sound of gentle rain. _____

10. For the next item, consider attending classes.
     The slow-paced voice of your English teacher.   _____

11. For the next item, consider seeing a live opera performance.
     The voice of an opera singer in the middle of a verse.   _____

12. For the next item, consider attending a new tap-dance performance.
     The sound of tap-shoes on the stage.   _____

13. For the next item, consider a kindergarten class.
     The voice of the teacher reading a story to the children.   _____

14. For the next item, consider driving in a car.
   The sound of an upbeat rock song on the radio.   _____

Bucknell Auditory Imagery Control Scale

The following scale is designed to measure auditory imagery, or the way in which you "think about sounds in your head." For the following pairs of items you are asked to do the following:

Read the first item (marked "a") and consider whether you think of an image of the described sound in your head. Then read the second item (marked "b") and consider how easily you could change your image of the first sound to that of the second sound and hold this image. Rate how easily you could make this change using the "Ease of Change Rating Scale." If no images are generated, give a rating of 1.

Please read "a" first and "b" second for each pair. It may be necessary to cover up "b" so that you focus first on "a" for each pair.

Please feel free to use all of the levels in each scale when selecting your ratings.

<u>Ease of Change Rating Scale</u>

1	2	3	4	5	6	7

No Image
Present At All

Could Change
The Image,
But With Effort

Extremely Easy
To Change
The Image

Change
<u>Rating</u>

1. For the first pair, consider attending a choir rehearsal.
    a. The sound of an all-children's choir singing the first verse of a song.
    b. An all-adults' choir now sings the second verse of the song.   _____

2. For the next pair, consider being present at a jazz club.
    a. The sound of a saxophone solo.
    b. The saxophone is now accompanied by a piano.   _____

3. For the next pair, consider listening to a rain storm.
    a. The sound of gentle rain.
    b. The gentle rain turns into a violent thunderstorm.   _____

4. For the next pair, consider driving in a car.
    a. The sound of an upbeat rock song on the radio.
    b. The song is now masked by the sound of the car coming to a screeching halt.   _____

5. For the next pair, consider ordering something over the phone.
    a. The voice of an elderly clerk assisting you.
    b. The elderly clerk leaves and the voice of a younger clerk is now on the line.   _____

6. For the next pair, consider seeing a live opera performance.
    a. The voice of an opera singer in the middle of a verse.
    b. The opera singer now reaches the end of the piece and holds the final note.   _____

7. For the next pair, consider going to a dentist appointment.
    a. The loud sound of the dentist's drill.
    b. The drill stops and you can now hear the soothing voice of the receptionist.   _____

8. For the next pair, consider the beginning of the song "Happy Birthday."
    a. The sound of a trumpet beginning the piece.
    b. The trumpet stops and a violin continues the piece.   _____

9. For the next pair, consider attending an orchestral performance of Beethoven's Fifth.
    a. The sound of the ensemble playing.
    b. The ensemble stops but the sound of a piano solo is present.   _____

10. For the next pair, consider attending a new tap-dance performance.
    a. The sound of tap-shoes on the stage.
    b. The sound of the shoes speeds up and gets louder.   _____

11. For the next pair, consider being at a live cricket game.
    a. The cheer of the crowd as a player hits the ball.
    b. Now the crowd boos as the fielder catches the ball.   _____

12. For the next pair, consider a kindergarten class.
    a. The voice of the teacher reading a story to the children.
    b. The teacher stops reading for a minute to talk to another teacher.   _____

13. For the next pair, consider attending classes.
    a. The slow-paced voice of your English teacher.
    b. The pace of the teacher's voice gets faster at the end of class.   _____

14. For the next pair, consider being at the beach.
    a. The sound of the waves crashing against nearby rocks.
    b. The waves are now drowned out by the loud sound of a boat's horn out at sea.   _____

Appendix 4

Diary Booklet Questions for Study 4

1) Did you press the button directly after tapping?   YES   NO   (if NO, please do so now)

2) Date _____   3)Time of earworm: _____ am/ pm (circle one)

4) Time diary completed: _____am/ pm (circle one)

5) Name of Tune: _____6) Performer of Tune: _____

7) Lyrics of the part of the tune in your head--- please list AT LEAST THE FIRST AND LAST part of the lyrics of your earworm:

_____

8) Each of the following word pairs describes a scale where each word represents the extremes of the scale. For each of the word pairs, tick the category that most closely describes the way you felt just before the earworm began:

	Very	Quite	Somewhat	Neither	Somewhat	Quite	Very	
Alert								Drowsy
Happy								Sad
Lonely								Connected
Energetic								Tired
Involved								Detached
Tense								Relaxed
Interested								Bored

9) Were you walking/running/cycling or doing another repetitive movement when you had the earworm?   YES   NO   If YES, what type of movement was it? _____

10) When was the last time you heard this song played aloud (via stereo, iPod, live performance, etc.)?
☐ less than one hour ago ☐ 1-3 hours ago ☐ today, but more than 3 hours ago
☐ within the past week, but not today ☐ more than one week ago

11) Please tick a box below to explain why you might have been experiencing this particular piece of music as an earworm (you can tick as many boxes as you feel are relevant).
☐ I have no idea why this tune came into my head.
☐ I recently heard or performed this song.
☐ A person, word, sound, or other occurrence in my environment reminded me of this song.
☐ The song reminds me of something from my past.
☐ I was thinking about a future event and related it to this song.
☐ Other (please explain if you think you know some other reason this earworm occurred):

_____