NONADJACENT KEY RELATIONSHIPS

THE EFFECTS OF TIME AND EVENTS ON THE PERCEPTION OF NONADJACENT KEY RELATIONSHIPS

By JOANNA SPYRA, B.Mus, M.Mus, B.A.

A Thesis Submitted to the School of Graduate Studies in Partial Fulfillment of the Requirements for the Degree Master of Science

McMaster University MASTER OF SCIENCE (2017) Hamilton, Ontario (PNB)

TITLE: The Effects of Time and Events on the Perception of Nonadjacent Key Relationships

AUTHOR: Joanna Spyra, B.Mus (McMaster University), M.Mus (University of Western Ontario), B.A. (McMaster University)

Supervisor: Professor M. Woolhouse

NUMBER OF PAGES: vii, 35

Lay Abstract

We have a very short memory span for musical keys. Humans can only remember a key for approximately 20 seconds after it's changed. However, it is unclear whether it is time spent away from the key that makes us forget or if it is the number of chords we heard in a second, interrupting key. The current study tests this distinction using nonadjacent key relationships—in which a key is established, a different key interrupts for a random amount of time and number of chords before returning to the original key for two target chords. Results confirmed previous findings as there was a significant effect of time on memory. However, there was no effect of number of chords, suggesting that it is the length of time spent away from a key, not number of chords that has an effect on memory.

Abstract

A musical key can only be retained in memory for approximately 20 seconds. However, other factors may be influencing the strength of this retention. The current study tests the influence of time and number of events (chords) in an intervening key on the deterioration in memory of a nonadjacent key. Stimuli first established a major key using traditional harmonic rules, then modulated to an intervening key that was either 6 or 9 seconds in duration and formed from either 4 or 6 chords. Stimuli then returned to the original key in a probe cadence. Participants were asked to rate this cadence in terms of its sense of closure. It was revealed that there is a significant negative effect of time on the probe cadence though no effect of number of events was found. This suggests that spending more time in an intervening key, and not the number of intervening chords, diminishes the memory of the original key. However, it is unclear from this study where in memory the nonadjacent key relationship is processed. Relevant literature is examined to form a working hypothesis with the goal of strengthening future studies with a capable foundation in memory research and theories.

Acknowledgements

I would like to thank Dr Matthew Woolhouse and the members of my graduate committee, Drs Dan Goldreich, Ranil Sonnadara, and Scott Watter, for their overwhelming support and guidance throughout the year. I have learned so many invaluable lessons from them all and have been given the confidence to ask my own questions and shape this thesis in my own way. I would also like to thank Matthew Stodolak and Kyle Gauder for their supreme editing efforts and continued support. Matthew has played a large role in this project as it was his thesis originally and he made a tremendous effort in the previous year traversing an unfamiliar field. And of course, I'd like to acknowledge the incredible support of family and friends, without whom my mental state would have been suboptimal.

Table of Contents

List of Figures and Tables

Figures

Tables

Introduction

What happens when we change musical keys? Do we maintain a clear memory of them or do they soon disappear? The current study aims to answer this and similar questions in a behavioural experiment using nonadjacent key relationships. This nonadjacency refers to the movement away from, and back again to, a harmonically established musical key. Imagine a trending pop song with its recurring choruses, or a romantic aria with its wonderous overarching structure; nonadjacent relationships are clear and ubiquitous in music. However, studies such as Cook (1987) demonstrate how inefficient the human mind is in processing such large scale tonal structures. How brief, then, does musical structure need to be for optimal perception and what other factors influence this ability?

A typical—nonadjacent—stimulus would be created from three parts: a nonadjacent key, intervening section and probe section. The nonadjacent key would first be established by using Western harmonic rules. It then moves to a different, intervening key (through harmonic modulation) which could be modified by a variety of factors such as time, rhythm, or timbre. Afterwards, the intervening section modulates back to the nonadjacent key for a probe cadence—the harmonic equivalent of a period or comma in language. Modulating back allows us to measure how novel the participant finds this key after experiencing the modifications in the intervening section. For example, if we wanted to know how well a participant remembered the nonadjacent key after a particular period of time, we could modify the intervening section to last varying periods of time within a predetermined range (i.e. between 2-20 seconds). The participant would then be asked to make a judgement on the probe cadence and based on their responses, we could track the novelty of the probe cadence as time increases (or decreases) in the intervening section.

A similar example to the nonadjacent key relationship could be presented in the sentence "Nora went to the store, passing by the huge dog, to buy some groceries". In this sentence, Nora's excursion to the store is nonadjacent to her task of buying groceries. This is interrupted by intervening information about the huge dog. Continuing our previous example, depending on whether we described the largeness of this dog for two seconds or for two hours, Nora's task of buying groceries would become increasingly difficult to connect with her trek to the store.

Previous research has used this paradigm to test the strength with which a representation of a key remains after a modified amount of time in an intervening key (as per our examples; Woolhouse, Cross & Horton, 2016; Farbood, 2016). However, it remains unclear how reliable the internal validity of these studies was as they extended time by adding chords to the intervening section. This addition could have introduced a possible confounding factor; that it was number of chords (events) that was affecting the representation of the nonadjacent key, not the time spent in a modulation (movement away from a pre-established key). The current study tests the effects of time as well as number of events on key representation and has found significant effects of time but not of number of events, providing support for studies like Woolhouse et al. (2016) and Farbood (2016). But what these studies lack is a description of the cognitive mechanisms that are responsible for the phenomenon of nonadjacent key relationships. Woolhouse et al. (2016) describe it in comparison to similar language constructs and though Farbood (2016) discusses memory in the context of the experiment, she does not describe how it is involved.

It would be useful to explore nonadjacent key relationships through memory research as there is a logical connection between the maintenance, and/or interpretation, of a key center and memory processing. Future experiments on the topic could be modified to test nonadjacency with a firm foundation in memory research. This would expand the types of questions that could be explored in the field and inform researchers of possible confounding factors that could be involved.

1.1 An Introduction to Memory

Memory can be classified into three systems: sensory, short-term and long-term memory (Craik & Lockhart, 1972). Sensory memory is fleeting and a keeper of all incoming information, whether attentional or not (Craik & Lockhart, 1972). Information is then transferred into short-term memory (STM) where it can be processed and encoded, usually to a certain time and storage capacity, before it is discarded or consolidated into long-term memory (LTM). LTM is a presumably limit-less store of information learned (or experienced) throughout life (Craik & Lockhart, 1972).

There is a problem when juxtaposing a nonadjacent relationship to memory models as there are a number of factors involved in its design. Because of its A1-B-A2 structure, responses to section A2 may be influenced by both memory for section A1 (global influences) and its harmonic relationship to section B (local influences). According to Leman (2000), however, both local and global effects are a part of STM, and thus, musical perception and retention.

Leman's psychoacoustic model for music cognition observes neuronal firings in the peripheral auditory system and brain stem to describe an image of musical processing at any given moment (Leman, 2000). The original signal is transformed many times (for roughness, rhythm etc.) before arriving at the Echoic Memory Module (EMM; Leman, Lesaffre & Tanghe, 2001). EMM calculates the correlation between an image of the music with a very short decay rate (0.1 sec) to one that has a longer decay rate (1.5 sec for example)—or in other words, a local image to a global one (Leman, Lesaffre & Tanghe, 2001). The higher the correlation, the better the two images fit together and the more fitting the harmonies remain over time. This allows us to make a more accurate judgement of key.

There remains, however, an issue with this model as it only works well with natural timbres (natural sounding instruments, for example). This model fails to find correlations between global and local images for pure tones (tones that are stripped of harmonic overtones and thus sound electric and unnatural) thus failing to represent the full capabilities of the human acoustic perceptual system (Marmel, Tillman & Delbe, 2010).

The most commonly used memory model in music cognition research is Baddeley's working memory (WM) model (Schulze & Koelsch, 2012), which divides STM into three components: a temporary verbal-acoustic process called the phonological loop, a parallel visual system (the visuospatial sketchpad), and a central executive that controls these systems through attentional regulation (Baddeley, 2003). The phonological loop is a verbal-acoustic storage system that is necessary, for example, to maintain a string of numbers—like a phone number—for a short period of time (Baddeley, 2003). The visuospatial sketchpad is a parallel system for the manipulation and storage of visual information. Lastly, the central executive controls these systems through a limited capacity attentional system (Baddeley, 2003).

The phonological loop is further divided into two subcomponents: articulation and storage (Baddeley, 2003). The storage subcomponent temporarily stores information in memory within a limited capacity. Though studies seldom agree on the limit of this capacity, it is estimated to range between 2-9 unique items or up to as many as 20 related items (Craik & Lockhart, 1972, Baddeley, 2003). The articulation component of the phonological loop is a rehearsal system that is responsible for maintaining information within the storage system—rehearsing a phone number, for example (Baddeley, 2003). Together, the subcomponents of the phonological loop are capable of maintaining and processing information for a short period of time, and have the capability to transfer information into LTM. Baddeley further appended an additional component to this system called the episodic buffer when it was evident that the limited capacity of WM was at times exceeded for STM tasks (Baddeley, 2003). This buffer allows us to temporarily store more information than the limited capacity systems (phonological loop and visuospatial sketchpad) can hold, creating the possibility of more robust processing (Baddeley, 2003). It is yet unclear, however, where music falls into this system. Does it share resources with language in the phonological loop or does it have a loop of its own?

1.2 Exploring the Phonological Loop

To explore music's possible connection to the phonological loop, studies juxtapose music to language, especially in terms of syntax. Syntax is a set of rules by which music theory and grammar combines elements into cohesive wholes (Patel, 2003). Studies in neuroscience suggest that brain areas, such as Broca and Wernicke's areas, may be shared between music and language syntax processing (Patel, 2003). However, it is unclear whether syntax processing for language uses the same cognitive systems as that for musical syntax despite shared neural resources and, indeed, whether these processes are part of the phonological loop or something else altogether. This distinction was tested using a 2X3 within subjects design by Fiveash and Pammer (2012). 61 participants were visually presented with either a list of words or a complex sentence. These language-based presentations were accompanied by either an unmodified musical stimulus (that followed harmonic rules), one that included a syntactic error, or a musical stimulus that included timbral manipulation. The dependant variable of this experiment was accuracy in a word/sentence recall.

The unmodified musical stimulus was used as a baseline for comparison to the other conditions. The stimulus with syntactic manipulation included one out-of-key chord that was at least three places away on the circle of fifths; this would sound like an error in the context of the musical stimulus. Fiveash and Pammer (2012) hypothesized that, in combination with the complex sentence condition, this syntactical manipulation would interfere with the syntactical processing needed for language and decrease accuracy of sentence recall. They also hypothesized that since simple word lists do not use syntactical processing, this musical manipulation would have no affect on word recall. To control for the possibility that an error in the music would cause decreases in recall accuracy due to an attentional distraction, the last musical manipulation was included. In this manipulation, one chord was played in flute timbre instead of the usual guitar timbre, potentially causing an attentional distraction without interfering with syntax processing. Fiveash and Pammer (2012) hypothesized that if attention has no effect on recall, there should be no statistically significant difference between results for this timbral condition and the baseline—no manipulation—condition. During the experiment, participants were asked to verbally recall the word list or sentence to the best of their abilities after every stimulus. Responses were recorded and recall accuracy was manually scored on a scale from 1-5 (1 = no recall, 2 = two or less correct, 3 = three or more correct, 4 = all correct, different order, 5 = all correct and in the right order).

Both stimuli type (word list, sentence) and music interference (normal, syntax, timbre) were statistically significant, as was the interaction between them. Unfortunately, the article does not mention whether or not simple main effects were calculated to account for the significant interaction found in the ANOVA. Failing to adjust for the interaction can inflate experiment-wise Type I error rate and cause the researchers to incorrectly reject the null hypothesis. Therefore, these results, though significant, must be interpreted with caution. T-tests were then performed between musical conditions for word lists and sentences separately. As hypothesized, syntactical manipulation was significantly different—lower—from any other in the sentence condition. However, syntax did not have the same effect in the word list condition. Also, no noteworthy significance was found between baseline and timbral manipulation in either word list or sentence trials, suggesting that there was no effect of distraction on recall.

Syntactical processing for music seems to interfere with that of language, suggesting there are shared resources between the processes. As the syntax manipulation condition had no significant effect on recall accuracy for word lists, the mechanism for processing syntax—at least musically—seems to be separate from the phonological loop (Fiveash & Pammer, 2012). This was a cleverly designed experiment as successful results could easily separate processing of syntax between the mediums as well as syntax from the phonological loop. However, the choice of scoring results on a scale instead of mathematically calculating percent accuracy, potentially not accounting for statistical significance in the interaction and then using at least 6 t-tests to compare conditions leaves some room for doubt in the results. A more rigorous experiment would have calculated more precise accuracy rates, would have used simple main effects when the interaction proved significant, and could have used a Tukey's HSD instead of t-tests to compare between conditions.

1.3 Behavioural Evidence

Further supporting the separation of music and language, many review articles (Berz, 1995; Schulze & Koelsch, 2012) have argued in favor of a tonal loop that is separate from the phonological loop. Berz (1995) reviewed behavioural data in an attempt to compare tonal to verbal WM and suggests an additional—separate—tonal loop to traditional WM models. He explored many of the arguments used in Baddeley studies to distinguish WM from LTM such as capacity limits, unattended stimulus effects, chunking and rehearsal (Berz, 1995). Many of these effects provide results that are similar in tonal processing as they are in verbal processing. For example, studies agree that there is a capacity limit to tonal processing. Though there are many studies that explore this limit in tonal WM, none of them seem to agree on a comparative range. An approximation based on the various studies estimates tonal recall to range from 7-15 notes (Berz, 1995), slightly larger than its language counterpart of 2-9 items (Craik & Lockhart, 1972).

There are also unattended stimulus and recency effects in tonal WM that are also similar to ones found in verbal studies. Interestingly, tonal processing seems to be disrupted by tones but not by spoken words, suggesting these processes rely on separate cognitive resources (Berz, 1995). Berz further argues that if music was part of the phonological loop, then unattended music would cause disruptions of verbal processing—as unattended speech does. However, there is no such interference, suggesting a separation between the tonal and phonological loops. Unfortunately, Berz himself points out that there is much contradicting information in this field of research and there seem to be many factors, especially attentional ones, that can have a high effect on the results.

The limitations of WM capacity can be greatly improved with the significant influence of LTM in forming cognitive strategies (Berz, 1995). Musically trained individuals, for example, are more efficient at chunking information than others; presumably through their superior knowledge of harmony. Chunking refers to a technique of grouping elements together into a unit. There are stronger similarities between objects in a unit then there are between units (Schulze & Koelsch, 2012). Chunking items on a list into units allows us to surpass the limited capacity for recall in WM by remembering the unit as a single item instead of many separate ones. Though this does propose that there is a connection between WM and LTM, this process is similar to all factions of the WM model.

Based on arguments such as these, Berz proposes to add a separate tonal loop to more traditional WM models. However, many of these arguments are weakly, if at all, tied to a separation between loops and Berz does not explain how they are connected in his proposal. In the end, he seems to admit defeat by stating there is not enough evidence to suggest tonal information is processed together with verbal information in the phonological loop, therefore, there must be a separation between them.

Though this is a convincing article in terms of providing support for the use of WM in musical processing, it does not discuss experiments in great detail and does not often discuss their credibility. Therefore, readers are provided with much anecdotal evidence by the end of the article but no empirical evidence to support it. Berz (1995) also does not explain the ways in which many of the arguments favour the proposed separation of the tonal from the phonological loop, leaving the reader to speculate the reasons on their own.

1.4 Neurological Evidence

Schulze and Koelsch (2012) review both behavioural and neurological studies. They argue that WM for music and language share some physical resources but, ultimately, musical processing is a separate system to language. This article provides some support to Fiveash and Pammer (2012) and Berz (1995) who have also argued for the existence of two separate systems (though Fiveash and Pammer argue that loops specific to syntax exist and those are separate from the phonological loop). This article explores similar behavioural evidence to Berz (1995): intervening tones interfere with tonal WM tasks more than verbal stimuli does and also that vocal music interferes more with verbal WM than instrumental music, suggesting that music and language processing does not solely rely on shared cognitive resources. However, neuroscience studies show very similar activations for verbal and tonal WM tasks, mainly left-lateralized frontal regions including Broca's area, the premotor cortex and the planum temporale (also similar activations in the cerebellum; Schulze & Koelsch, 2012). This is contrary to behavioural data as it clearly suggests shared resources.

However, there is an interesting difference in neural activations between musicians and nonmusicians. Whereas nonmusicians' activation for tonal WM mimics that of the phonological loop, musicians recruit many additional structures that are exclusively used for either verbal or tonal WM tasks (Schulze & Koelsch, 2012). This provides evidence for the existence of two separate WM loops for phonological and tonal information as even though both systems show considerable overlap, they also rely on different subcomponents (Schulze & Koelsch, 2012).

This idea was explored further in an event-related brain potential (ERP) study by Koelsch, Gunter, Schröger and Friederici (2003). They observed brain responses to harmonic violations and compared these reactions to syntactical violations in language. Brain potentials were juxtaposed between stimuli that did not modulate to those that modulated to distantly related keys. These distant modulations were hypothesized to be experienced by the brain as violations of harmonic expectancies. Brain responses to modulated stimuli elicited an early right anterior negativity (ERAN) in the inferior fronto-lateral cortex; an area corresponding to Broca's area but in the right hemisphere. The ERAN supports the authors' hypothesis as it is a signal that is typically elicited by harmonic violations. Interestingly, it is reminiscent of the ELAN (early left anterior negativity) which is elicited during syntax processing in language (Koelsch et al., 2003). Though there are similarities in responses to syntactical violations in music and language, these responses specific to their hemispheres.

Another response elicited by harmonic violation was a frontal negativity that was maximal at approximately 500 msec (N5). The N5 has a higher amplitude whenever music moves to less appropriate keys (modulates to a more distantly related key; Koelsch et al., 2003). This effect was, indeed, seen in the study as the amplitude of the N5 greatly differed for modulating stimuli than from those that did not. This can be interpreted as the result of a more strenuous process of harmonic integration as moving to a distantly related key requires more cognitive effort to process than music that follows expectation. The N5 is also reminiscent of a similar signal—the N400—that is elicited in language processing (Koelsch et al., 2003). The N400 similarly has a higher amplitude as a consequence of contextually inappropriate words than for appropriate words (Koelsch et al., 2003). However, the N5 and N400 are not indicative of exactly the same processes though the authors hypothesize that the N5 may contribute to the N400, or vice versa.

There was also a slow-going negativity found that was maximal at approximately 500-1500 msec and was predominant in the frontal regions of the right hemisphere. This signal is particularly interesting as it is characteristic of WM processes that have also been observed during language processing (Koelsch et al., 2003). This provides support to studies that have suggested some overlap between music and language WM processes.

This study provides interesting support to previous findings; there seem to be similar responses to music syntax violations as there are to those in language and there are certain overlaps in WM processes as well. However, these responses occurred in separate hemispheres, thus failing to support previous findings of shared neural resources. There may also be reason to question the statistical validity of the study as many of the ANOVAs resulted in significant interactions. The authors did not address how, if at all, they accounted for this significance.

1.5 Discussion

Though many studies have contradicting, and often confusing, results and conclusions, there are some commonalities that can be derived from them. Syntactical violations have no effect on word processing, suggesting there are separate WM functions for syntax and phonological information (Fiveash & Pammer, 2012). Within processing for syntax however, there seem to be shared cognitive resources as syntactical violations in music have a detrimental effect on sentence recall (Fiveash & Pammer, 2012; Patel, 2003). Nonetheless, unattended stimulus effects suggest a divide between the phonological loop, which processes verbal information such as word lists, and a tonal loop for musical analysis (Berz, 1995; Schulze & Koelsch, 2012). In summary, language and music processing seems to be domain specific yet syntactical processing is shared amongst modalities.

Neuroscience studies are no less confusing. There are significant overlapping neural resources between music and phonological systems (Schulze & Koelsch, 2012). Nonetheless, the two systems also have domain specific subcomponents, suggesting that despite considerable overlap, music and language WM functions are unique (Schulze & Koelsch, 2012). This evidence supports behavioural studies which argued for the separation of tonality from the phonological loop. ERP studies further the divide between modalities. Despite remarkable similarities between brain responses to music and language, these responses are specific to opposing hemispheres (Koelsch et al., 2003).

1.6 Conclusion

A common theme between behavioural and neuroscience studies has been to divide tonality from the phonological loop within WM. Nonadjacent key relationships are most likely processed within WM as an established key can only maintain its influence for a short period of time (Woolhouse, Cross & Horton, 2016; Farbood, 2016). However, since modulation was highly significant in this study, it is probable that nonadjacent key relationships involve both WM and semantic memory found in LTM (Firmino, Bueno & Bigand, 2009). But where in WM does nonadjacency processing occur?

Based on findings in behavioural and neuroscience studies, I hypothesize that nonadjacency is processed within the tonal loop in WM, and at times making use of the episodic buffer as WM approaches its capacity limits. It is also possible, due to the effect of modulation, that syntactical processing is also occurring, thus making use of the syntactical loop proposed by Fiveash and Pammer (2012) and corresponding links to LTM. WM limits in terms of number of items (chords/events) can be tested in the future as it is possible the current study did not reach tonal loop capacity limits. This would be the first step in exploring the connection between nonadjacent key relationships and WM and thus, provide valuable information for future studies in this direction.

The Effects of Time and Events on the Perception of Nonadjacent Key Relationships

2.1 Abstract

Increasing the duration of an intervening key has a negative effect on memory for the original, nonadjacent key. Evidence suggests the recollection of a key only remains for 20 seconds after the beginning of a new key section. But factors other than time might influence the cognitive processing of nonadjacent key relationships. By using a probe-cadence paradigm, this study sought to determine whether time or the number of musical events (chords) determined the deterioration in memory of nonadjacent key perception. Stimuli were constructed in three parts: (1) a major key was established through a standard chord progression; (2) an intervening section, either 6 or 9 seconds in duration formed from either 4 or 6 chords, was introduced in 12 possible keys; and (3) a short pause was followed by the probe cadence in the original key, i.e. the key at (1). 51 participants were asked to estimate the amount of harmonic closure they perceived at the probe cadence. Results confirmed previous findings on the significant negative effects of time on the influence of the nonadjacent key. However, there were no significant effects of number of events. This provides evidence that it is the length of time, not the number of musical events, in an intervening section of music that determines the recollection of the original key.

2.2 Introduction

It is common for music to be heralded as a golden thread that unifies humanity; to be seen as a universal, connecting language. And there certainly are homogeneous properties—structural and behavioural commonalities—amongst the musics of different cultures (Campbell, 1997; Patel, 2003). But music, as an abstract experience, may not be perceived as music theory dictates (Cook, 1987; Cuddy & Thompson, 1992). Our perception of current musical harmonies, for example, is deeply influenced, and often revised, by preceding events (Temperley, 2007, p. 89). Furthermore, the establishment of a 'home' key has a profound impact on our perception of the following modulation—movement from one key to another.

In traditional music theory, keys are organized into a circle of fifths; a geometric representation of key relationships in which keys are arranged into a circle, each separated by a distance of a fifth (see Figure 1). Neighbouring keys are thus constructed of all the same notes except for one. Perceptually, this is the closest and most consonant interval other than the octave.

These keys may be spaced evenly in either direction in theory, but our perception of the modulation is asymmetrical; movement in a clockwise direction is judged as significantly closer in distance than counter-clockwise modulations (Cuddy & Thompson, 1992). This may be due to familiarity—clockwise modulations are used more often in music than counter-clockwise movements (Cuddy & Thompson, 1992). The higher probability of modulation in this direction may be causing us to judge the keys as more related and thus, the distance traversed as smaller. The same patterns are not seen in single voice sequences (Thompson & Cuddy, 1992), suggesting there are specific factors in play that impact perception.

Though increasing the amount of harmonic information directly influences our perception, it remains unclear from these studies which other factors may have an effect and how long the home key maintains its influence after modulation. One of the most influential studies on structural perception, Cook (1987) explored the perception of large-scale tonal structures in music. Composers of typical Western music would often write larger structures into their works but it was unclear how well, if at all, we were able to perceive these structures. To this end, participants were presented two versions of six excerpts from the classical era (for solo piano). These versions included an original, nonmodified excerpt that remained in one key and one that was modified by transposing the last few measures to a different key. Participants were asked to compare these two versions and choose one based on various characteristics (expressiveness, coherence, pleasure, and sense of completion). Results were surprising as only the two shortest excerpts prompted any preferences from the participants, suggesting that listeners do not perceive large-scale tonal structures past approximately one minute (Cook, 1987). However, this study only measured the duration of the excerpt and not the modified section itself, providing little distinction between local and global effects. This, and similar works, have sparked research into nonadjacent key relationships.

Nonadjacency is a relationship between three key sections in music: the establishment of an original, nonadjacent key; the movement to a novel, interfering key section; and an eventual return to the original key. Such relationships are common in Western music though may vary in terms of length—they can span the length of a phrase, movement or even the entirety of a piece. Depending on factors involved, this second modulation is either perceived as a definite return to the 'home' key, or can be perceived as a new key altogether.

Woolhouse, Cross and Horton (2016) measured the effect of a nonadjacent key on the closure of a musical passage after varying lengths of time. Stimuli were based on the nonadjacent key template: a nonadjacent key section was first established through traditional Western harmony, the intervening section modulated to one of 12 major keys (as found on the circle of fifths in Figure 1), after which the stimulus ended on a probe cadence—a pair of chords with the harmonic equivalence of a period or comma—in the original key. Participants were asked to rate this cadence on its sense of completion using a Likert-type scale. The intervening key was modified between trials to be of varying lengths, lasting approximately 2-9 seconds in duration.

Participants rated shorter trials as having significantly more closure than longer trials, suggesting the effects of the nonadjacent key deteriorate as time increases. As the sense of closure declines, the return to the original key for the closing cadence seems more novel to participants, thus suggesting they have spent enough time in the intervening section to override—to various degrees—the influence of the pre-established original key. In post hoc analysis, Woolhouse et al. (2016) predicted that this effect lasts approximately 11 seconds. In a similar study, Farbood (2016) extended the duration of the intervening key past the predicted value in Woolhouse et al. (2016). This time, participants were instructed to rate, on a sliding scale, the amount of tension felt (based on Lerdahl's tonal tension model; Lerdahl & Krumhansl, 2007) as the stimulus progressed. Tension would spike when participants came across a novelty in the stimulus—like a modulation to a different key. The more novel the event was, the more extreme the spike in tension. Thus, participants would experience a significant spike at the first modulation (from nonadjacent to intervening key), followed by a second spike as the stimulus modulated back to the original key. If tension decreased rapidly after the second spike, the second modulation was perceived as less novel. As such, Farbood theorized the participants must be recognizing the original key to some degree despite shifting to an intervening key along the way. Results confirm this pattern; as the duration of the intervening key increased, so did the amplitude of tension felt at the second modulation. These tension profiles extended the predicted duration of the effect of nonadjacency from Woolhouse et al. (2016) to approximately 20 seconds (Farbood, 2016).

However, both studies modified the duration of the intervening key by adding or subtracting chords (Woolhouse et al., 2016; Farbood, 2016). A duration of 5 seconds, for example, was always comprised of 6 chords while a duration of 2.5 seconds only had 2 chords in it (Woolhouse et al., 2016). This creates a possible impeding factor: it may be the number of chords, and not time, that is influencing the effect of the nonadjacent key.

Historically, various memory models have suggested a capacity limit to short term memory (or later known as working memory; a part of short term memory that is responsible for processing of incoming sensory information). Models have suggested that working memory has a limited capacity depending on the type of stimulus (Craik & Lockhart, 1972); it can span anywhere from 2-9 items if the items are perceived as separate events and up to as many as 20 items if they are related (such as words that grammatically relate to form a sentence). After this limit, items stored in a working memory buffer are replaced with newer items. It is possible that the number of chords used in these studies is effecting the representation of the original key in working memory by replacing the chords stored with newer chords presented in the intervening section.

2.3 Present Study

The aim of this study was to test the perception and retention of global key structures despite intervening auditory stimuli. To this end, participants were presented with a musical phrase consisting of a nonadjacent section, an intervening section in a different key, and a two-chord cadence (V-I) to complete the phrase and served as a probe stimulus (as adapted from Woolhouse et al., 2016). The intervening section was modified to test the effects of time and events—the number of chords—separately. It consisted of either 4 or 6 chords and lasted either 6 or 9 seconds in total. Table 1 shows a summary of all possible variations of the intervening key. In comparing these modifications, we can judge the effect of either time or number of events on the sense of closure and thus, the participants' ability to retain a key over time despite the presence of distractors (in the form of a variable intervening key).

This experimental design poses inherent challenges. First, the nonadjacent key must be as salient as possible in order to reliably establish a 'home' key through all stimuli equally. Secondly, the effects of key must be limited to the global effect of the nonadjacent key on the probe cadence rather than that of the more local intervening key. In order to correct for the limitation of key salience, the nonadjacent section was always constructed from a I-I^6-ii^{65}-V-I chord progression as this progression ensures the nonadjacent key is well established according to the practices of music theory (Piston, 1978).

Woolhouse et al. (2016) demonstrated an effective means of separating global from local effects; they compared the ratings for a typical stimulus (keys A-B-A) to a non-modulating stimulus (with keys B-B-A). If the ratings were significantly different, the typical stimulus was presumably effected by both global and local structures, while the non-modulating sequence was only effected locally. By subtracting them from each other, the only factor left is the global effect. This provided evidence that participants could, in fact, match a probe cadence to the original key despite intervening sections of varying lengths.

There were two main areas of interest: the length of time spent in the intervening key and the number of events (chords) it contained. This led to six possible hypotheses as illustrated in Table 2. The first hypothesis was that increasing the amount of time spent in the intervening key would have a negative effect on the sense of completion of the probe cadence (hereon called T-). This means that spending more time in a novel key decreases the global effect of the nonadjacent key, thus lowering ratings for completion of the probe cadence. In contrast, time could have a *positive* effect (T+) or no effect whatsoever (T). The same three hypotheses could be applied to number of events: increasing the number of events in the intervening section could have a negative, positive, or no effect on the capacity of the probe cadence to complete the musical phrase (E-, E+, E respectively).

2.4 Methods

Participants. 51 undergraduate students (17-21 years of age; 46 female, 5 male) participated in the experiment. 21 participants had 5 or more years of formal musical training (average: 6.86 years, standard deviation: 1.71). Participants were drawn from a variety of departments at McMaster University and were rewarded for their participation with a course credit under the McMaster Psychology, Neuroscience and Behaviour research system.

Apparatus. Stimuli were generated in *MuseScore2* and programmed into an online web module using HTML5 and JavaScipt. Participants listened to the stimuli through headphones and were able to adjust the volume freely. Responses were entered by adjusting a slider programmed into the web module.

Stimuli & Procedure. 48 stimuli were created for this experiment; there were 4 conditions with 12 key modulations each. Modulations would span from no modulation to an 11 semitone modulation from the nonadjacent to the intervening key. Each of the key progressions were used for all four of the experimental conditions. Conditions were determined using two independent variables: number of events and duration (in seconds) of the intervening key. The events in consideration were defined as the number of chords used in the intervening key progression (Figure 2). Two of the experimental conditions had 4 events and two had 6 events in the intervening key. Out of these, intervening keys could last either 6 or 9 seconds.

The nonadjacent key was established using a I-I^{6}-ii^{65}-V-I progression that lasted approximately 3.75 seconds. The intervening adjacent key was wrap-constructed so that it always began and ended with the same chords (V-I), minimizing any ambiguity as to the key it was in. Each stimulus ended on a probe cadence in the nonadjacent key. This was made up of two chords (the tonic and dominant) that are typically used in music to indicate the end of a piece. A two-beat rest separated the intervening key from the probe cadence to bring attention to the probe. The two-beat rest and probe cadence lasted 1.5 seconds each. The tempo was set at 80 bpm and only major keys were used in this experiment. The stimuli were made up of chords as they offer less ambiguity in key finding (Thompson & Cuddy, 1992; Cuddy & Thompson, 1992). Participants were asked to rate the probe cadence on its sense of tonal closure using a Likert-type scale of 1-7 (from not at all to strong sense of closure). The experiment lasted for approximately 25 minutes.

2.5 Results

Using normalized scores, a 2X2 factorial within-subjects ANOVA was used to explore the effects of time (6 seconds vs 9 seconds) spent in the intervening key and number of events (4 or 6 events) as well as the interaction between them. The effects of key distances between nonadjacent and intervening keys were accounted for as a cofactor. The assumption of sphericity was not violated as the degrees of freedom were no larger than 1. A Tukey HSD analysis was then performed comparing all possible conditions and their individual significances.

There was no significant interaction between time and number of events ($F_{1, 50} = 0.0127$, $p = 0.9106$). The main effect of number of events also did not reach significance ($F_{1, 50} = 1.6339$, $p = 0.2071$). However, the main effect of time was highly significant ($F_{1, 50} = 22.0617$, $p < 0.0001$). With regards to hypotheses in Table 2, finding no significant effect of events provides evidence for E. However, there was a significant *negative* effect of spending more time in the intervening section on the sense of completion (providing evidence for T-). This means that the more time spent with a distractor made the probe cadence less structurally viable.

Mean Residuals. Results can further be compared by subtracting the means of conditions, as seen in Table 3. For example, by subtracting the mean of trials (across participants) in the condition comprised of 9 seconds and 6 events from the mean of the condition comprised of 6 seconds and 6 events, a mean residual of 0.21 remains. Note that this mean reflects the difference between conditions with the same number of events but different lengths of time. Since 0.21 is a positive difference, as seen in Figure 3, it suggests that participants rated the trials with a 6 second intervening key higher on average than trials with longer—9 second—intervening sections.

All trials comparing number of events had mean residuals close to zero, further illustrating that there were no significant effects of events on closure. Mean residuals of time, however, were always significant and pointed to a negative effect of increasing time in the intervening key.

Two of the comparisons in Table and Figure 3 compare conditions between which both time and events were different. Both are highly significant ($p < 0.00005$ and $p < 0.05$, respectively); however, they support opposing hypotheses for number of events. Logically, events are not the influencing factor in these comparisons, whereas the effect of time is highly supported by these results.

2.6 Discussion

Increasing the duration of an intervening key significantly deteriorates the global effects of the nonadjacent key. However, global effects are not disturbed by increasing, or decreasing, the number of events in the second key. These results provide validity to previous studies that increased the duration of the intervening key by also increasing the number of events. This is a universal trend as there was no significant difference between musicians and nonmusicians in this study.

There are a number of possible explanations for why we saw no effect of the number of events in this experiment. One hypothesis is that the nonadjacent key phenomenon is not an issue of memory at all. As memory research suggests various capacity limits (Berz, 1995; Craik & Lockhart, 1972), it would be expected that as memory load is increased to approach these limits, there would be greater decay of the nonadjacent key retention and thus, a greater effect size. Though the current study found no effect of events, the fact that there was a highly significant decay of key retention over time strongly supports the hypothesis that memory is, indeed, part of the nonadjacent key effect.

Therefore, a second hypothesis is that memory is being tested but either the conceptualization or the operationalization of ‘number of events’ was not quite correct. In our design, we considered each chord to be experienced as its own, separate event. However, the definition of an event in music is subjective—it can be defined as one chord, one phrase, or even an entire movement. It is possible that participants were not experiencing each chord as its own event; each event followed traditional musical harmony and were all related by being in the same key. This could lead to chunking (Berz, 1995) which would cause the entirety of the intervening section to be experienced as one event.

Conversely, the difference between 4 and 6 events could have been too small to effect memory in a meaningful way. According to Craik and Lockhart (1972), working memory (WM) has a limited capacity. It has been estimated to only hold between 2-9 items unless these items are connected somehow (such as words that form a sentence). Connected items increase WM capacity to approximately 20 items (Craik & Lockhart, 1972). Studies on tonal recall extended this capacity limit to approximately 7-15 notes without chunking effects (Berz, 1995). In the present study, it is possible that the number of events fall short of wholly replacing the representation of the original key. As the chords in the intervening key are related harmonically, they may be perceived as connected—similar to when words form a sentence—and thereby increase memory capacity to ~20 items. Future studies should expand the number of events in this paradigm to thoroughly test whether a higher number of events has a negative impact on WM capacity or, indeed, if a different conceptualization of an event in music is necessary to uncover these types of capacity limits.

Current results paired with future exploration of this missing capacity limit could provide support for memory research in music cognition that divorces tonality from the phonological loop in WM models (Berz, 1995; Schulze & Koelsch, 2012). By providing a strong foundation in WM research, we could explore the tonal loop from a cognitive standpoint.

2.7 The effects of modulation

A post hoc analysis of the data indicates some perplexing trends in terms of modulation. Most modulations were rated negatively on average except for the non-modulating sequences (keys A-A-A), a modulation by a semitone, and a modulation up the fourth from the nonadjacent key. This is contrary to Krumhansl and Kessler's (1982) multidimensional map of key relatedness in which keys based on the fourth and fifth scale degree are rated as higher—and therefore more pleasant together—than the semitone. In our study, the modulation to the fifth is rated almost as poorly as a modulation to the tritone or seventh scale degree (see Figure 4), both of which are extremely dissonant and generally unpleasing to the ear.

Furthermore, Cuddy and Thompson (1992) found that modulations to the fifth scale degree are perceived as being closer to the original key than modulations to the fourth. We hypothesize that the reason for the poor ratings of the fifth are due to the relationships in the chord progressions between the intervening key and the probe cadence. In isolation, the probe cadence should be perceived as V-I in the nonadjacent key. However, due to the intervening key being in the dominant key itself, the first chord of the cadence may be interpreted as being the tonic of the dominant key and not the dominant of the tonic key. This may sound odd to the participant as the phrase then ends by moving away from the established key. This could cause confusion and unwarranted interference. Future studies may need to avoid certain modulations (such as the fifth), as they may artificially increase local effects.

Negative ratings for the majority of modulations involved may also be due to awkward voice leading. Stimuli in this study included relatively large leaps in the bass which could sound unnatural to participants. Careful reorganization of the harmonic structure in the stimuli may influence the effects of modulation in future studies.

2.8 Conclusion

The significant effect of time and the insignificance of number of events provides support for such studies as Woolhouse et al. (2016) or Farbood (2016). In their exploration of the duration of musical memory, time spent in the intervening key was prolonged by increasing the number of chords (or ‘events’) in that section. This addition provided a possible conflicting factor: that the number of events had a negative effect on key retention. The present study extended this research by investigating both time and number of events. Based on current results, we can be confident that, to the extent to which this study tests the factor, including more chords in a sequence did not interfere with the intended tests.

However, there remain some limitations to the current study that can be addressed in future research. Chiefly, there was a large amount of unexplained variance in the data. This may be due to unknown factors affecting closure in this type of paradigm. Conversely, the variance could also be due to a flaw in the experimental design. In the future, it would be worth exploring which factors affect the variance and run an experiment that uses a more robust methodological approach.

As discussed in the previous section, the number of events may not have been extensive enough to interfere with working memory. It should be noted, however, that despite rather small differences in both time and number of events, the manipulation of time was enough to result in highly significant differences. This provides compelling support for previous studies even if the possibility yet remains for events to be of some interference on a larger scale. Future experiments should expand on this factor and fully test events as a possible intervening factor with the goal of decreasing unexplained variance in the present data.

As with many psychological studies of music, the current stimuli are not particularly representative of music; in our quest to control the factors involved, musical stimuli are limited in their representation of music heard in regular environments. As such, many other exciting factors can be explored that have the potential to significantly alter results.

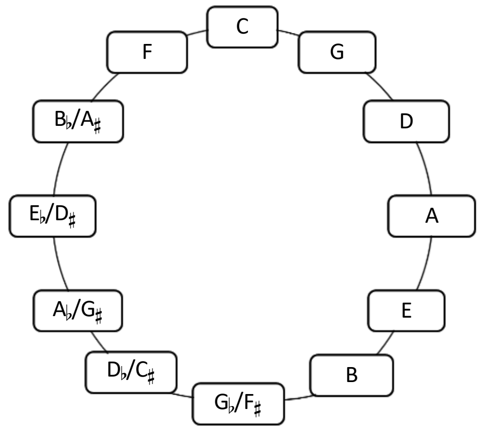

Fig. 1. The circle of fifths for 12 major keys.

Condition	Time (sec)	Events
6t/6e	6	6
9t/6e	9	6
6t/4e	6	4
9t/4e	9	4

Table 1: Summary of the possible modifications to the intervening key

Hypothesis (time, events)	Description
T-	Negative effect on closure of more time
T+	Positive effect on closure of more time
T	No effect of time
E-	Negative effect on closure of more events
E+	Positive effect on closure of more events
E	No effect of number of events

Table 2: Summary of proposed hypotheses

Figure 2: Sample of stimuli demonstrating 4 modulations and all possible modifications to the intervening key section. Also shown is the wrap construction of the intervening section. Stimuli shown all begin in C major but it should be noted that stimuli were designed in all key combinations for the experiment.

Target Factor	Mean Residual	Relationship $T_1E_1 – T_2E_2$	No Difference	Positive	Negative
Time	0.21	66 - 96	T	T-	T+
Events	0.06	66 - 64	E	E+	E-
Both Change	0.28	66 - 94	T and E	T- / E+	T+ / E-
Both Change	-0.15	96 - 64	T and E	T+ / E+	T- / E-
Events	0.07	96 - 94	E	E+	E-
Time	0.22	64 - 94	T	T-	T+

Table 3: Summary of the hypotheses supported by residual valence.
Highlighted boxes correspond to findings above.

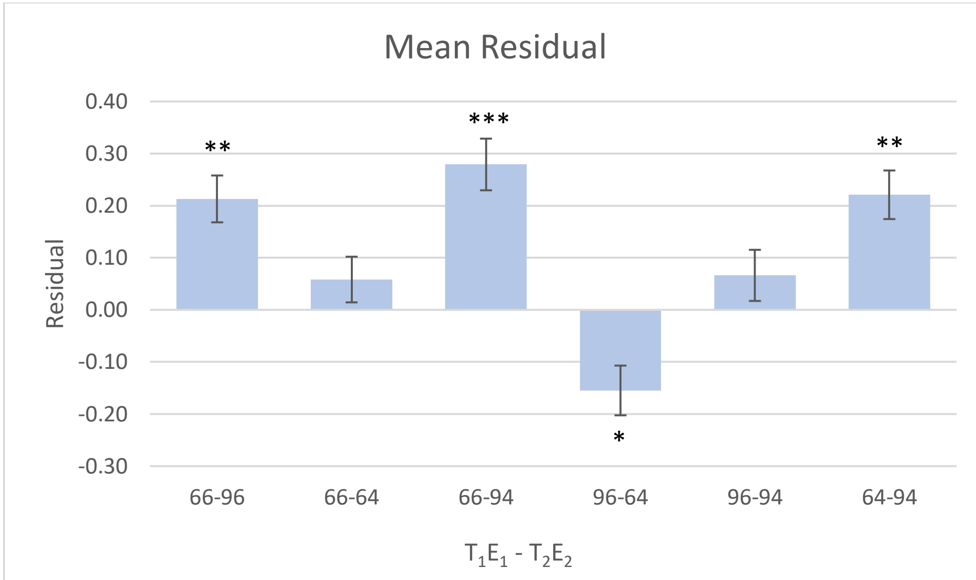

Figure 3: Mean residual values between conditions. X-axis describes condition; 66-96 indicates a 9 second/6 event condition was subtracted from a 6 second/6 event condition. * $p < 0.05$, ** $p \leq 0.0005$, *** $p < 0.00005$

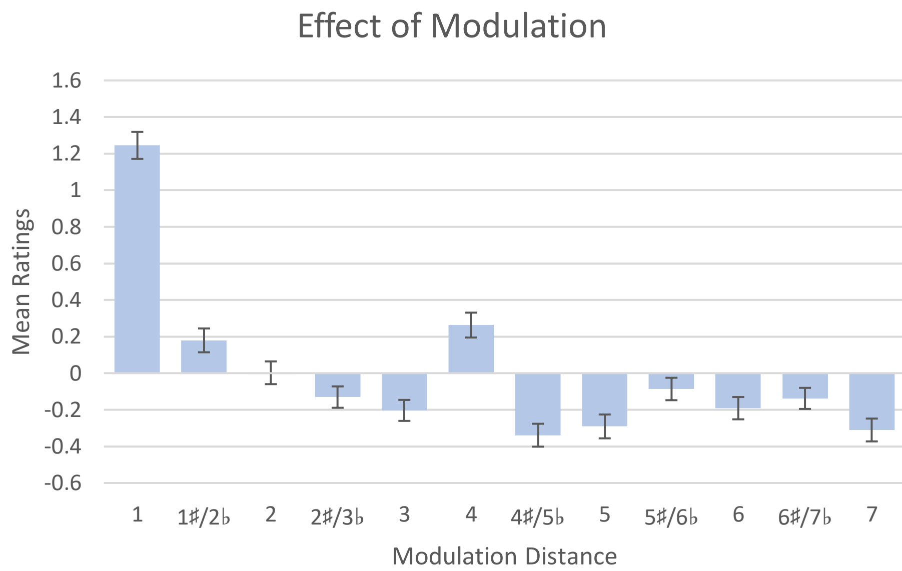

Figure 4: a breakdown of the average ratings for modulations to each scale degree.

References

Baddeley, A. (2003). Working memory and language: An overview. *Journal of communication disorders*, *36*(3), 189-208. doi:10.1016/S0021-9924(03)00019-4

Berz, W.L. (1995). Working memory in music: A theoretical model. *Music Perception: An Interdisciplinary Journal*, *12*(3), 353-364. doi:10.2307/40286188

Campbell, P. S. (1997). Music, the universal language: Fact or fallacy?. *International Journal of Music Education,* (1), 32-39. doi:10.1177/025576149702900105

Cook, N. (1987). The perception of large-scale tonal closure. *Music Perception, 5*(2), 197-206. doi:10.2307/40285392

Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*(6), 971-684. doi:10.1016/S0022-5371(72)80001-X

Cuddy, L. L. & Thompson, W. F. (1992). Asymmetry of perceived key movement in chorale sequences: Converging evidence from a probe-tone analysis. *Psychological Research, 54*(2), 51-59. doi:10.1007/BF00937133

Farbood, M. (2016). Memory of a tonal center after modulation. *Music Perception, 34*(1), 71-93. doi:10.1525/mp.2016.34.1.71

Firmino, É. A., Bueno, J. L. O., & Bigand, E. (2009). Travelling through pitch space speeds up musical time. *Music Perception: An Interdisciplinary Journal*, *26*(3), 205-209. doi:10.1525/mp.2009.26.3.205

Fiveash, A. & Pammer, K. (2012). Music and language: Do they draw on similar syntactic working memory resources? *Psychology of Music*, *42*(2), 190-209. doi:10.1177/0305735612463949

Koelsch, S., Gunter, T., Schröger, E., & Friederici, A.D. (2003). Processing tonal modulations: An ERP study. *Journal of Cognitive Neuroscience*, *15*(8), 1149-1159. doi:10.1162/089892903322598111

Krumhansl, C. L., & Kessler, E. J. (1982). Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review*, *89*(4), 334. doi:10.1037/0033-295X.89.4.334

Leman, M. (2000). An auditory model of the role of short-term memory in probe-tone ratings. *Music Perception*, 481-509. doi:10.2307/40285830

Leman, M., Lesaffre, M., & Tanghe, K. (2001). Introduction to the IPEM toolbox for perception based music analysis. *Mikropolyphonie-The Online Contemporary Music Journal*, *7*.

Lerdahl, F. & Krumhansl, C. L. (2007). Modeling tonal tension. *Music Perception*, *24*, 329-366. doi:10.1525/mp.2007.24.4.329

Marmel, F., Tillmann, B., & Delbé, C. (2010). Priming in melody perception: Tracking down the strength of cognitive expectations. *Journal of Experimental Psychology: Human Perception and Performance, 36*(4), 1016. doi:10.1037/a0018735

Patel, A. (2003). Language, music, syntax and the brain. *Nature Neuroscience, 6*(7), 674-681. doi:0.1038/nn1082

Piston, W. (1978). *Harmony* (4th ed.). New York, NY: Norton.

Schulze, K. & Koelsch, S. (2012). Working memory for speech and music. *Annals of the New York Academy of Sciences*, *1252*(1), 229-236. doi:10.1111/j.1749-6632.2012.06447.x

Temperley, D. (2007). *Music and probability.* Cambridge, MA: MIT Press.

Thompson, W. F. & Cuddy, L. L. (1992). Perceived key movement in four-voice harmony and single voices. *Music Perception, 9*(4), 427-438. doi:10.2307/40285563

Woolhouse, M., Cross, I., & Horton, T. (2016). Perception of nonadjacent tonic-key relationships. *Psychology of Music, 44*(4), 802-815. doi:10.1177/0305735615593409