THE CONTRIBUTION OF EXPERTISE ON MUSICAL AND PHONOLOGICAL PROCESSING

By Kyra L. Bowe

The goal of the present study was twofold, it 1) sought to investigate whether musical information is processed by the phonological loop, a component of working memory, and 2) explored whether musical expertise affects the degree to which this component is involved in the processing of this information. In order to investigate this issue, an auditory interference task was employed using musicians and non-musicians. Specifically, participants heard an initial stimulus (chord or word), followed by intervening stimuli (chords, words, or silence), and were then asked to indicate whether a final test stimulus was the same or different from that of the initial stimulus. It was predicted that as one becomes an expert in music, neural resources that are used to process phonological information would begin to process musical information to a larger degree. This would result in more interference for musicians in conditions in which cross modal interference (chords interfere with to-be-remembered words and vice versa) occurs as compared to non-musicians. Contrary to the expertise hypothesis, results indicated that musicians, as compared to non-musicians, had a larger relative difference between matched (i.e., music – music; words – words) and mismatched (i.e., music – words, words – music) conditions. This suggests that for non-musicians, the phonological loop may coopt the processing of both phonological and musical information. On the other hand, with expertise in music there seems to be distinct processing of the two different modalities as opposed to the greater recruitment of the processing of musical information within working memory.

THE CONTRIBUTION OF EXPERTISE ON MUSICAL AND PHONOLOGICAL PROCESSING

by

Kyra L. Bowe

A Thesis Submitted
In Partial Fulfillment of the Requirements
For the Degree of

Master of Science-Psychology
Cognitive & Affective Science

at

The University of Wisconsin Oshkosh
Oshkosh WI 54901-8621

July 2018

COMMITTEE APPROVAL

_____ Advisor

8/29/2018 Date Approved

_____ Member

8/29/2018 Date Approved

_____ Member

8/29/2018 Date Approved

PROVOST AND
VICE CHANCELLOR

_____

8/29/2018
Date Approved

FORMAT APPROVAL

_____

8/29/2018
Date Approved

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

LIST OF TABLES

Page

LIST OF FIGURES

Page

Introduction

Memory underlies our ability to learn and profit from experience, without it we could not function. One need not look any further than cases of severe memory impairment to begin to understand and appreciate the importance of our ability to retain information. For example, anterograde amnesia is characterized by a profound inability to retain new information after brain damage to cortical regions supporting memory. Individuals suffering from this type of amnesia are unable to consciously recollect information encountered after the brain damage, which severely affects their quality of life; it is as if they are constantly stuck in the past (Shallice & Warrington, 1974). Through the study of memory, it has become apparent that there are several different memory systems that are constantly interacting with one another as we go about our daily life. However, our ability to integrate information into meaningful experiences is accomplished through a cognitive architecture that supports both the temporary storage and active processing of information, commonly referred to as working memory.

The predecessor to working memory is referred to as the modal model of memory, which proposed that memory was comprised of three distinct components: sensory registers, short-term memory, and long-term memory (Atkinson & Shiffrin, 1968; see Figure 1). This model postulated that information from the environment is detected and briefly held by the sensory registers, which are specific to different sensory domains (e.g., visual, auditory, tactile, etc.). These brief traces of sensory information held by the sensory registrars are then available for potential entrance to short-term

memory (STM); otherwise, the information would be subject to decay. If the sensory information were attended to, it could then enter the STM system, which is where the temporary processing of information is handled. The longer the information resides in the STM store, the greater the likelihood it will be consolidated into long-term memory (LTM). The LTM system can be thought of as a vast repository that holds all of the experiences, knowledge, and skills that were accumulated throughout one's lifetime (Atkinson & Shiffrin, 1968).

Working Memory

Over time, the conception of STM has evolved into what is now referred to as the multicomponent working memory model, which emphasizes not only storage, but also the manipulation and processing of information in immediate memory (Baddeley, 2000; Baddeley & Hitch, 1974). As opposed to STM, which simply describes the temporary storage of information, working memory can be described as a limited capacity system that is involved in the active processing, manipulation, and storage of temporary information (Baddeley, 2004; Baddeley, 2012; Baddeley & Hitch, 1974). It has been proposed that information can enter working memory from either sensory sources or LTM and is believed to handle visuospatial and phonological information (Baddeley & Hitch, 1974). Visuospatial information refers to information pertaining to the visual modality such as color, shape, or spatial information. This information does not only include information that arrives via the visual sensory system, but also information from long-term memory. Phonological information, on the other hand, refers to spoken

information typical of language and language based written information (i.e., writing) once it has been recoded into a phonological code. Again, phonological information can stem from both incoming sensory information in addition to information stored in long-term memory (Baddeley & Hitch, 1974; Baddeley, 2012). Overall, the model is composed of several different processing components, each of which will be described in turn (Baddeley, 2000; Baddeley & Hitch, 1974; see Figure 2).

Central Executive. The central executive can be described as the coordinating mechanism of working memory. It controls and directs incoming sensory information to the appropriate subsystem, the visuospatial sketchpad or phonological loop (described in following sections), for processing. (Baddeley & Hitch, 1974; Baddeley, 1996). This component's function corresponds to executive attention (Baddeley & Hitch, 1974). When a particular task demands extensive involvement of the visuospatial sketchpad or phonological loop, the central executive deploys the necessary resources for efficient processing. The central executive is also thought to be responsible for higher-level thought processes involved in reasoning and language comprehension. Since the attentional capacity of the central executive is limited, tasks that are too demanding may exhaust its resources and can result in overload, leading to compromised or inefficient complex thinking. Overall, the central executive is in charge of all tasks thought to involve executive control. As more was discovered about the central executive, Baddeley (1996; 2001) proposed the fractionation of the central executive into functions that include selective attention, manipulation of information in working memory, task switching, and response inhibition.

Episodic Buffer. The episodic buffer is a limited capacity temporary storage system that holds information in a multimodal code, and can integrate information from the visuospatial sketchpad, phonological loop, and long-term memory creating a unitary episodic representation (Baddeley, 2000; Cowan, 2005). This component was added to account for anomalies found in phonological loop research. For example, subjects can typically only remember a sequence of five or six unrelated words, but if the words comprise a meaningful sentence then one is often able to remember up to 16 or more words (Baddeley, Vallar, & Wilson, 1987). The episodic buffer, in this case, acts as a gateway between working memory and long-term memory providing meaning to the to-be-remembered words, thus allowing an individual to remember more words than what is typically possible due to the limited capacity of working memory.

Visuospatial Sketchpad. The visuospatial sketchpad is a limited capacity system that holds and manipulates visual and spatial information (Baddeley, 2003a; Phillips, 1974; Phillips & Baddeley, 1971). Visuospatial memory became of interest during the 1960s when it was demonstrated that memory for a point on a line was retained for over a period ranging up to 30 seconds, but that it could be disrupted by a concurrent information-processing task (Posner & Konick, 1966). This suggested some form of active rehearsal for visuospatial information. This component seems to operate independently of the other working memory component (i.e., phonological loop), as it has been demonstrated that separate verbal and visual tasks do not tend to interfere with one another (Brooks, 1968). Furthermore, empirical work has also shown a dissociation between the processing of visual and spatial information within working memory,

whereby visual and spatial tasks do not interfere with each other (Della Sala, Gray, Baddeley, Allamano, & Wilson, 1999; Klauer & Zhao, 2004; Logie, 1986).

Phonological Loop. The phonological loop holds and manipulates speech-based information and is comprised of two subcomponents, the phonological store and the articulatory control process (Baddeley & Hitch, 1974; Baddeley, 1992; see Figure 2). The phonological store holds speech-based information and has a limited capacity of only a few seconds before information begins to decay. The articulatory control process has the ability to refresh information within the phonological store through the process of subvocalization (e.g., silently repeating information to yourself) to offset decay. This structure is also responsible for recoding visually presented stimuli (e.g., written words or pictures) into a phonological code that is then able to access the phonological store, whereas spoken words have direct access to the store (Baddeley, 1994; Baddeley, 2003a; Baddeley, 2012; Baddeley & Hitch, 1974).

Evidence that suggests the phonological loop is involved in the processing of linguistic information comes from a variety of studies using a wide range of methodology. One particular case study demonstrated that a patient with a pure phonological loop deficit in the absence of a compromised LTM was unable to acquire a second language (Baddeley, 1993; Baddeley, Papagno, & Vallar, 1988). This could suggest that because the new language information did not have access to the phonological loop, it could not be consolidated into LTM, thus compromising the ability of this patient to acquire the new language. Further, empirical studies have demonstrated that disrupting the phonological loop interferes with foreign language acquisition, but not

learning to associate pairs of unrelated native language words. This is thought to be the case because associating pairs of known words relies more on semantic coding as opposed to actually learning the vocabulary, which would depend to a high degree on the phonological loop (Papagno, Valentine, & Baddeley, 1991; Papagno & Vallar, 1992). Additional support for the phonological loop being language based comes from the observation that the ability of children and adults to learn a second language strongly relates to their phonological loop capacity (Service, 1992; Atkins & Baddeley, 1998).

Phonological Information in Working Memory

Research investigating how the phonological loop operates has revealed two different effects, the Word-Length Effect and the Phonological Similarity Effect, both of which help to elucidate how information is processed by this component.

Word-Length Effect. The word-length effect is characterized by a decrease in recall accuracy for longer as opposed to shorter words (Baddeley, Thompson & Buchanan, 1975). For example, a list of to be remembered five-syllable words (e.g., university, opportunity, international, constitutional, auditorium) will be recalled with less accuracy as compared to a list of one-syllable words (e.g., sum, pay, wit, bar, hop). The word-length effect is because longer words take longer to rehearse, and since the phonological loop has a limited capacity of only a few seconds, longer words are more susceptible to decay (Cowan et al., 1992). Additionally, the word-length effect demonstrates that rehearsal is important for the prevention of decay, as inhibiting the rehearsal process by repeating irrelevant sounds (i.e., subvocalization) such as the word

"the" can abolish the effect. This process of repeating irrelevant sounds impairs recall of both short and long words equally, because it blocks rehearsal of the to-be-remembered words in the articulatory control process and they therefore cannot enter the phonological store.

Phonological Similarity Effect. The phonological similarity effect demonstrates that letters and words that sound similar are more difficult to recall than dissimilar sounding letters and words. This can be demonstrated using an immediate serial recall task with the idea that the auditory characteristics of the to-be-remembered material would indicate the nature of the code on which the recall is based, in this case language. For example, people have a more difficult time recalling the letters "t, c, v, d, b, g" than "b, w, y, k, r, x" because the letters in the first sequence are more similar in term of their auditory characteristics (Conrad & Hull, 1964; Baddeley, 2003b; Baddeley, 2012). This effect sheds light on the nature of the code that affects subsequent recall and lends support to the idea that information within the phonological loop is based on a speech-based code specific to language.

Transfer of Language Based Visual Information to a Phonological Code. Although the phonological loop has been the subject of much research over the past 40 years, the question of whether non-phonological based information can access the phonological loop remains unanswered. It is theorized that auditory-based language information, but not visual based language information, has direct access to the phonological store (Baddeley & Hitch, 1974). Evidence supporting this notion comes from research utilizing articulatory suppression. Articulatory suppression is a technique

used to block the rehearsal process of the phonological loop. As mentioned previously, the articulatory control process of the phonological loop is responsible for recoding visual information into a phonological code that can be processed in the phonological store.

One study that utilized this technique to assess the recoding of visual stimuli had participants recall lists of short and long words presented through either the auditory or visual modality. During presentation of the to-be-remembered word lists participants were either instructed to count repeatedly from one to eight (i.e., articulatory suppression), or to simply remember the list of words (Baddeley & Hitch, 1975). Results revealed that the word length effect was present for the words presented through the auditory modality, meaning lists of shorter words were recalled more accurately than lists of longer words. Interestingly, the word length effect was completely abolished by suppression of the words presented through the visual modality, where recall was impaired for both short and long words. These results suggest that articulatory suppression, a process that taxes the articulatory control process, prevented the conversion of visually presented word lists into a phonological code that could then be processed within the phonological loop. Alternatively, the word length effect was present for the word lists presented through the auditory domain because words presented through this domain do not rely on recoding by the articulatory control process, as they are already in a phonological code and have direct access to the phonological store. Interestingly, articulatory suppression has also been shown to eliminate the effect of phonological similarity for the visual modality, but not for the auditory modality (Baddeley, Lewis, Eldridge, & Thomson, 1984; Murray, 1968). Overall, these results

support the notion that linguistic information presented through the auditory modality has direct access to the phonological store, whereas if it is presented through the visual modality, it does not.

It remains to be seen, however, whether other forms of auditory information have direct access to this same structure. Of particular interest to the current study is whether music, which shares some characteristics with linguistic information, can be actively processed by this system. Previous research has provided conflicting accounts regarding whether music has access to the phonological loop, however methodological issues (reviewed below) may explain these contradictory findings. Clarifying the nature of information that can directly access the phonological loop allows us to expand on what is known about working memory, and its subcomponents.

Music in Working Memory

Initial research investigating whether non-linguistic auditory information has access to the phonological loop relied on the use of tones for stimuli. Tones here refer to a single electronically produced pitch (i.e., frequency) presented outside of a musical context (i.e., chords, melody, etc.). Deutsch (1970) was one of the first researchers to examine non-linguistic auditory information processing within working memory. The purpose of this study was to explore whether a specific type of sensory information, pitch, would interfere with verbal recall. To do so, participants were presented with a series of trials in which a to-be-remembered initial tone was followed by interfering tones or words and ended with a comparison tone. Participants were instructed to ignore the

interfering stimuli and indicate whether the comparison tone was the same as or different from the initial to-be-remembered tone. Results revealed a specific interference effect for pitch in which participant's accuracy was worse when the initial to-be-remembered and comparison tones were separated by intervening tones as opposed to words. These results were interpreted to suggest that phonological and tonal information have separate storage areas, as the interfering tones caused a significantly greater amount of interference as compared to interfering words for the recall of the to-be-remembered tonal stimuli. This should not have been the case if shared processing in the phonological loop existed, as both types of stimuli should then have caused interference.

A study that sought to further investigate the processing of speech and potentially non-speech-based acoustic information in working memory was conducted by Semal, Demany, Ueda, and Halle (1996). Their two-experiment study examined the speech specificity hypothesis, which proposes that the pitch associated with language and the pitch associated with non-speech sounds are stored in separate systems. In Experiment 1 of this study, participants heard an initial tone followed by a set of intervening stimuli (words or tones), and finally a comparison tone. The participants' task was to indicate whether the comparison tone was the same pitch as the initial tone. Experiment 2 was similar to Experiment 1, with the exception that the initial and comparison stimuli were words instead of tones. Again, participants had to indicate whether the pitch of the comparison word was the same as the initial word. Results indicated that participant's recall for the pitch of tones and words was disrupted by both intervening tones and words. Overall, these results suggest, as opposed to the findings of Deutsch (1970), that

there may be shared storage of speech and non-speech pitch information in working memory.

Williamson, Baddeley, and Hitch (2010) expanded on the aforementioned research in a 3-experiment study that compared differences in the retention of auditory information using non-musicians and musicians. Studying both non-musicians and musicians is important because it is possible that auditory information is processed differently for these two populations. The aim of this study was to compare the serial recall of speech and music to assess for any similarities. The authors created a new measure to compare the phonological similarity effect accurately to a musical analogous-pitch-proximity effect. The pitch-proximity effect refers to the finding that tones closer together in pitch (pitch-proximal) are recalled less accurately than tones that are further apart in pitch (pitch-distal). The experiment required participants to listen to sequences of both similar and dissimilar sounding letters and tones, with the sequence length varying across trials (e.g., 4-8 items). The participants' task was to recall in order the letters and tones. Results revealed a pitch-proximity effect comparable to the phonological similarity effect for non-musician participants. Specifically, tones that were closer in pitch were more difficult to recall than tones that were further apart in pitch. This is comparable to the phonological similarity effect in which words that sound similar are more difficult to recall than words that sound dissimilar. The authors suggest that this result supports the notion that musical and phonological information have shared storage in non-musicians; however, the authors concede that it is also possible that there are separate structures for the processing of musical and linguistic information that

simply happen to operate in similar ways. Thus, this paradigm cannot successfully demonstrate shared storage, as it only assesses working memory independently for each stimulus. Surprisingly, musicians did not demonstrate a pitch-proximity effect, suggesting differences between musicians and non-musicians concerning cognitive processing. Specifically, it is possible that they have better working memory and auditory processing abilities (Chang, 2014; Gaser & Schlaug, 2003; Pallesen et al., 2010; Schlaug, Jancke, Huang, Staiger, & Steinmetz, 1995). Alternatively, it is possible that while processing skills are the same, musicians excel at rehearsal and encoding strategies, such as using multiple verbal, auditory, and tactile techniques to encode information (Roden, Grube, Bongard, & Kreutz, 2014; Williamson et al., 2010).

The aforementioned studies provide conflicting results with regard to shared processing of language and musical information, with Deutsch (1970) finding no shared storage and Semal et al. (1996) and Williamson et al. (2010) suggesting its possible existence. Unfortunately, the Semal et al. (1996) and Williamson et al. (2010) studies suffer from a methodological limitation related to how the interfering effect of musical stimuli on language processing was assessed. In Semal et al. (1996), the participants' task was to remember the pitch of the initial stimuli (tone or word), not the stimuli itself (i.e., the actual word). Thus, this experimental design may not accurately reflect any interfering effects, as memory for the semantic properties of the language stimuli was not assessed. The results of the Williamson et al. (2010) study should also be interpreted with caution, as the paradigm did not assess the potential effects of cross-modal interference (i.e., language interfering with music and vice versa). Consequently, these

results suggest that the processing of phonological and musical information are similar in appearance and cannot be used to conclude overlapping underlying mechanisms.

Atherton et al. (2018) set out to address these limitations by directly comparing the interfering effects of musical and language stimuli on one another. The experimental paradigm consisted of participants hearing a to-be-remembered target stimulus (word or chord), followed by intervening stimuli (words, chords, or silence), and ending with a comparison stimulus. Participants were instructed to ignore any intervening stimuli and judge whether the comparison stimulus was the same as or different from the target stimulus. The results revealed a similar pattern between the to-be-remembered language stimuli and the to-be-remembered music stimuli. Specifically, the silent conditions (e.g., language-silence, music-silence) resulted in the greatest accuracy, followed by the mismatched conditions (e.g., language-music, music-language), with the matched conditions (e.g., language-language, music-music) resulting in the least amount of accuracy. Since interference was observed in both mismatched conditions, it was suggested that musical and phonological information might share at least partial storage in working memory. Additionally, participants had higher accuracy in the language-music condition than the music-language condition, as well as higher accuracy in the language-language condition than the music-music condition. This was suggested to be because all participants were experts at processing language as opposed to music, thus allowing words direct access to the phonological loop. Music on the other hand may not be as readily accessible to the phonological loop because the participants were not musicians (i.e., experts in music) resulting in less interference, as was observed in the

music conditions (i.e., music-silent, music-language, music-music) with all conditions resulting in lower accuracy than the language conditions. If the participants were also experts in music as well as language, then one potential effect could be similar amounts of interference for both stimuli. Overall, the results support the notion that language and musical information share at least partial processing resources in working memory.

Expertise

Musical ability is dependent on exposure and practice. It is not an ability that everyone possesses and takes time to hone. Such skills subsequently result in certain advantages. For example, it is commonly accepted that the brain is malleable and can change due to experience and exposure. Chang (2014) provides a brief review of the structural and functional reorganization of the brain due to skill learning and expertise. Importantly, it has been demonstrated that an increase in skill and expertise not only results in changes to the motor cortex responsible for supporting the execution of that skill, but also in cognitive and perceptual domains; specifically, in auditory processing for musical expertise. In fact, research has demonstrated superior auditory processing of complex sounds for musicians compared to non-musicians (Kaganovich et al., 2013). With respect to the cognitive advantages of musical expertise, many studies have demonstrated that musicians may possess enhanced working memory for musical stimuli (Berti, Munzer, Schroger, & Pechmann, 2006; Cohen, Evans, Horowitz, & Wolfe, 2011; Pallesen et al., 2010; Schulze, Zysset, Mueller, Friederici, & Koelsch, 2011; Williamson

et al., 2010), as well as language stimuli (Chan, Ho, & Cheung, 1998; Roden et al., 2014; Schulze et al., 2011; Talamini, Carretti, & Grassi, 2016; Williamson et al., 2010).

Thus, in the context of previous research on musical processing, it is possible that most individuals have to recode musical information prior to its entry into the phonological loop, a structure thought to be devoted to the processing of linguistic information. However, with extensive practice this structure may become more involved in the active processing of this information (i.e., through repeated exposure and eventually expertise this structure begins to process this information). This could be analogous to findings in the face processing literature where it has been demonstrated that cortical regions that are prominently involved in processing faces (e.g., fusiform face area) have been found to also process other stimuli that individuals have a great deal of experience (i.e., expertise) with (e.g., birds, cars; Gauthier, Skudlarski, Gore, & Anderson, 2000). Moreover, increased activation of cortical regions responsible for face processing has also been observed when individuals become more familiar with processing classes of novel stimuli (Gauthier, Tarr, Anderson, Skudlarski, & Gore, 1999).

Schulze et al. (2011) has provided functional evidence of an expertise account, in which structures involved in processing verbal working memory are recruited for musical working memory, as one becomes an expert in music. In this study, musicians and non-musicians completed a verbal and tonal working memory task. Participants were first presented with either a visual cue, the letter V for verbal or T for tonal, depending on the subsequent stimuli to be rehearsed. This was followed by a stimulus sequence consisting of spoken syllables and tones presented simultaneously. Participants then silently

rehearsed either the syllables or tones, depending on the visual cue they received for the given trial. At the end of the trial a test stimulus was then presented, which was comprised of a simultaneously presented spoken syllable and tone. The participants' task was to indicate whether the test syllable or tone, depending on condition, was presented previously in the stimulus sequence. Behavioral results indicated that musicians were more accurate than the non-musicians in the tonal condition, but not the verbal condition. This might be expected due to the musicians' vast amount of exposure to musical stimuli as compared to non-musicians. Functional results revealed that for non-musicians, similar cortical regions were active during both verbal and tonal conditions, although these structures were more strongly activated during verbal as opposed to the tonal conditions. This differential activation was believed to be because non-musicians are only experts at processing language, and not music. Like non-musicians, it was observed that musicians had similar areas of activation for both verbal and tonal conditions, but there were several structures that were activated more strongly than non-musicians for both conditions. Furthermore, musicians had additional structures that were active for both verbal (i.e., the right insular cortex) and tonal (i.e., areas involved in sensorimotor encoding) conditions. The authors suggest that because of the activation of additional structures for the verbal and tonal conditions, this could indicate the existence of two WM systems in musicians: one for phonological information and one for musical information. Unfortunately, the activation of additional structures for verbal and tonal conditions does not conclusively reveal separate WM systems, as these additional areas that were activated could be due to processes not related to working memory per se.

Furthermore, the behavioral data that was collected should be interpreted with caution as the general increase in working memory (Berti et al., 2006; Chan et al., 1998; Cohen et al., 2011; Pallesen et al., 2010; Roden et al., 2014; Schulze et al., 2011; Talamini et al., 2016; Williamson et al., 2010), commonly observed in musicians was not controlled for by comparing performance in experimental conditions to that of a control condition. This would have allowed one to determine accurately how performance difference between musicians and non-musicians were related to either general differences in working memory capacity or differential systems being utilized. Employing a behavioral paradigm that includes cross-modal interference and attempts to account for overall working memory ability would be better suited to assess if expertise in music results in a larger overlap of shared processing in working memory. If the phonological loop began to process music-based information due to expertise with this type of information, it would be predicted that musicians, who are experts in music, would rely on the phonological loop more so for the processing of music as compared to non-musicians. Thus, in contrast to the results reported in Atherton et al. (2018), musicians would show a greater degree of similarity in levels of interference for both stimuli, as both stimuli would be competing for the same limited resource.

The Present Study

Research investigating whether shared storage mechanisms for linguistic (phonological) and non-linguistic (musical) information in the phonological loop have provided conflicting results. Some research has supported the notion of shared storage

(Atherton et al., 2018; Semal et al., 1996; Williamson et al., 2010), while other research suggests completely separate storage mechanisms (Deutsch, 1970; Schulze et al., 2011). The present study seeks to determine whether shared storage for musical and phonological information exists within the phonological loop and whether expertise for musical stimuli can lead to a greater recruitment of the phonological loop in the processing of this type of information. If linguistic and musical information are processed in separate structures, then the two should not interfere with one another on an interference task. However, if musical expertise results in the phonological loop recruiting musical information for processing, as it does with linguistic information, then more interference between linguistic and musical information should be observed for musicians as compared to non-musicians. To investigate this expertise hypothesis, the present study seeks to employ a paradigm similar to that used in Atherton et al. (2018) in which a target stimulus (e.g., word or chord) will be presented followed by intervening stimuli (e.g., silence, words, or chords) ending with a comparison stimulus (e.g., word or chord). The trials in the present study are different from Atherton et al. (2018), as six intervening stimuli were used instead of four. This is done in an attempt to make the task more difficult to decrease any possible ceiling effects that may occur for musicians, as they have been shown to possibly have enhanced working memory (Berti et al., 2006; Chan et al., 1998; Cohen et al., 2011; Pallesen et al., 2010; Roden et al., 2014; Schulze et al., 2011; Talamini et al., 2016; Williamson et al., 2010). For each trial, participants will be required to determine whether a target stimulus matches that of a comparison stimulus. This paradigm will allow for the assessment of cross-modal interference indicative of

possible shared storage, as well as the influence of expertise on shared storage. It should be noted that although there will be no-interference (i.e., silence condition), matched interference (i.e., language-language, music-music), and mismatched interference (i.e., language-music, music-language) conditions, difference scores (i.e., the accuracy scores for the appropriate silence condition will be subtracted from the appropriate experimental condition) will be used for statistical analysis. The language matched (LL) and language mismatched (LM) conditions will be subtracted from language silent condition, whereas the music matched (MM) and music mismatched (ML) conditions will be subtracted from the music silent condition. Difference scores will be used because it has been demonstrated that musicians may have superior working memory (Berti et al., 2006; Chan et al., 1998; Cohen et al., 2011; Pallesen et al., 2010; Roden et al., 2014; Schulze et al., 2011; Talamini et al., 2016; Williamson et al., 2010). Thus, difference scores will be better apt to account for relative increase for musicians in overall working memory capacity as the silent condition can capture the relative musician advantage and the difference scores will account for this increased working memory capacity. The resulting paradigm is thus a 2 (group: musician vs. non-musician) x 2 (target stimulus: language vs. music) x 2 (nature of overlap: matched [i.e., language-language & music-music] vs. mismatched [i.e., language-music & music-language]) mixed factorial design. Research question two and following hypotheses are based upon the difference scores in accuracy from the initial stimulus conditions to the silent condition.

Primary Research Questions and Hypotheses

RQ1. Is there shared storage of musical and phonological information in the phonological loop of working memory?

H1. Based upon previous research that demonstrates a degree of interference for both musical and phonological information, it is hypothesized that there will be evidence of shared storage for both musicians and non-musicians by interference occurring in both matched and mismatched conditions across both modalities.

RQ2. Does expertise in music result in the phonological loop being more heavily involved in the processing of musical information? If expertise in music results in the phonological loop were more heavily involved in processing musical stimuli, then one would expect to see similar degrees of interference between matched and mismatched domains. More specifically, because musical information is processed in a similar way to that of phonological information in musicians, a similar degree of interference should be observed between matched and mismatched conditions as compared to non-musicians.

H2a. Based upon previous research that has demonstrated that matched conditions (music-music; language-language) resulted in more interference as compared to mismatched conditions in non-musicians, it is hypothesized that a main effect of nature of overlap (i.e., matched vs. mismatched conditions) will be observed, where matched conditions will result in more interference as compared to mismatched conditions.

H2b. Based upon the expertise hypothesis, which would predict that expertise in music will result in the phonological loop being more heavily involved in the processing of this type of stimuli, it is predicted that an interaction of group by nature of overlap will

be observed. More specifically, it is predicted that a larger difference of interference between the matched as opposed to the mismatched conditions will be observed for the non-musicians as compared to the musicians. This is because, in musicians, musical information will be processed more heavily by the phonological loop and thus will be treated more like linguistic stimuli. This will result in more interference by this type of stimuli for both musical and linguistic information.

Method

Participants

A statistical power analysis was performed using GPower 3.1 (Faul, Erdfelder, Lang, & Buchner, 2007) for sample size estimation, based on data from Atherton et al. (2018). The effect size (ES) in this study was .50, which is large using Cohen's (1988) criteria. With an alpha = .05 and power = 0.80, the projected sample size needed for this effect size was approximately 20 participants per group, which was obtained.

Two hundred twenty-seven students from the University of Wisconsin, Oshkosh (94 males and 133 females) participated in the study for course credit and/or the opportunity to be entered into a drawing for one of eight $25.00 gift-cards. Participants' average age was 19.78 years, and 72.7% identified as Caucasian, 9.7% as African American, 6.2% as Hispanic American, 4.4% as Asian American, and 7% identified as other. Of the 227 subjects that participated in this study, data from 103 participants were not used for data analysis for the following reasons: 37 did not meet the 80 percent accuracy rate criterion for practice trials, 13 were fluent in more than one language, 29 encountered computer errors during participation, and 24 failed to follow instructions with respect to reporting their responses on the response sheet (i.e., participants did not indicate their response for all trials). Participants who were fluent in more than one language were excluded due to research suggesting that individuals who speak multiple languages may have enhanced working memory capabilities (Biedron & Szczepaniak, 2012; Schroeder, Marian, Shook, & Bartolotti, 2016). The data for the remaining 124

participants were used for statistical analysis. Of these participants, 104 were considered non-musicians with the majority (*N* = 56) having no musical training and a portion (*N* = 48) having less than 4 years of training. The remainder of the participants (*N* = 20) were considered musicians, with the majority (*N* = 16) having 5-9 years of musical training and a small portion (*N* = 4) having 10-19 years of musical training. A participant was classified as a musician if they had at least 5 years of musical training, which has been used as the criterion for defining one as a musician in multiple other studies (Kaganovich et al., 2013, Weijkamp & Sadakata, 2017, Zhang, Peng, & Hu, 2015).

Materials

Language stimuli consisted of monosyllabic words recorded and edited using the Audacity© software program (Audacity Team, 2012). Spoken words and musical chords were matched in pitch to control for any acoustical confounds. This was accomplished by specifying the desired frequency for word stimuli in Audacity© to match word frequency with those used for chords (see Table 2 for a list of the chords that were used). Music stimuli consisted of major chords recorded in the Reason© software. The duration of the stimuli was edited using the Izotope Ozone 5© software program to ensure that each chord and word were played for 750ms.

Language Stimuli. Language stimuli were adopted from Atherton et al. (2018) for use in this study. These stimuli were selected primarily from a list of phonologically similar and dissimilar words used by Coltheart (1993). Target stimuli consisted of nine 3-letter words. Intervening language stimuli consisted of six pseudo-randomly presented 3-

letter dissimilar sounding words per target word. The comparison language stimuli were randomly presented and consisted of the same word either as the target word or as one of two 3-letter similar sounding words. These similar sounding comparison words had either the first letter or last letter of the target word changed. For example, if the target word was "mad," the similar sounding comparison word was either "man" or "had." A list of all target and intervening words is provided in Table 1.

Music Stimuli. Musical stimuli were adopted from Atherton et al. (2018) for use in the study. Music stimuli consisted of nine randomly presented major chords located on the Circle of Fifths. Intervening chords consisted of six chords presented in a randomized order. These chords were chosen based on their distance from the target chord on the Circle of Fifths, in order to create chords that were dissimilar from the target chord. The Circle of Fifths is a circular visual representation that shows the harmonic relationships between the major and minor chords of the chromatic scale. Dissimilar chords were three, four, and five turns clockwise and counterclockwise from the target chord. Comparison chords were either the same chord or a similar sounding chord located on the Circle of Fifths. Similar chords were two turns clockwise or counterclockwise from the target chord. A list of all target and intervening chords is provided in Table 2.

Questionnaires. Demographic information was collected using a write-in questionnaire that asked participants to identify their gender, age, year in school, ethnicity, primary language, any languages that they are fluent in and for how many years, and any musical experience, such as musical instruments played and for how many

years. This questionnaire was completed at the end of the experiment. A list of all the demographic questions is provided in the Appendix.

Procedure

Overview. Experimental sessions were run in a classroom setting in groups of roughly 30-35 participants. Two separate classrooms were used for scheduling purposes. One classroom had windows the other did not. An attempt was made to control for external noise from individuals outside of the classrooms by posting signs requesting silence. As the participants entered the classroom, they were instructed to sit at a spot with a folder and not open the folder until instructed. Further, they were told that once instructed to open the folder, only view the specific document that they were told to view. Participants were also instructed to silence any electronic devices and set them aside for the remainder of the experiment. Participants received a folder containing: two consent forms, instructions, answer sheets, and a demographic questionnaire. Once the experiment was ready to begin participants were told to open their folder and retrieve the two informed consent documents. Participants were then told what the informed consent document was and that they were to read it in its entirety and sign one of the two documents if they agreed to participate, the other one was for them to keep if they wanted. If a participant did not wish to participate after reading the consent form, they were told they could either leave at that time or stay the duration of the experiment and leave at the end. The experimenter then told the participants to take the instruction sheet out of their folder and read along with the experimenter as the instructions were read

aloud (by the experimenter). It was emphasized that participants were to ignore any intervening sounds or silence, and that the primary task was to judge whether or not an initially presented word or musical chord was the same as or different from a target word or musical chord. Next, participants were instructed to take out the answer sheets, complete 12 practice trials, and ask questions afterwards if needed. This was followed by the main experiment. Trials were played over a loudspeaker at a medium volume level using iTunes©. After the main experiment was completed, participants were instructed to take the demographic form out of the folder and complete it. Finally, the debriefing information was read aloud to the participants and any questions were answered, completing the experiment.

Practice Trials. The twelve practice trials were administered to familiarize participants with the experimental procedure. Practice trials were presented randomly, which included two trials for each of the six different experimental conditions. A complete description of the trials is described in the next section.

Experimental Trials. Each trial began with an initial click sound lasting 750 ms; this signified the start of each trial. This click was followed by 250 ms of silence. Next, the target stimulus (word or chord) was played for 750 ms, also followed by 250 ms of silence. The intervening stimulus phase of the trial lasted eight s in total. The intervening stimuli consisted of six separate chords or words (depending upon the condition), each lasting 750 ms, followed by 250 ms of silence before the next stimulus. The intervening stimulus phase of the silent condition was eight s of silence. The comparison phase was the last phase of the trial. A comparison word or chord was play

for 750 ms, followed by 250 ms of silence. Finally, five s of silence was provided for participants to circle their response ("same" or "different") on their answer sheet. Thus, there were six different types of experimental conditions, three of which presented a target and comparison word, and another three that presented a target & comparison chord. For both the target & comparison, word and chord trial types there was a condition in which there was matched intervening stimuli, mismatched intervening stimuli, and intervening silence. For the matched intervening stimuli, word targets were followed by intervening word stimuli (word-word) and chord targets were followed by intervening chord stimuli (chord-chord) For the mismatched intervening stimuli, word targets were followed by intervening chord stimuli (word-chord) and chord targets were followed by intervening word stimuli (chord-word). Finally, for the intervening silence, word targets were followed by intervening silence (word-silence) and chord targets were followed by intervening silence (chord-silence). Each participant completed 162 trials (27 trials per condition). The experiment lasted roughly 60 minutes in total.

Results

Initial Working Memory Analysis

Previous research has demonstrated that musicians may possess enhanced working memory capabilities compared to their non-musician counterparts (Berti et al., 2006; Cohen et al., 2011; Pallesen et al., 2010; Schulze et al., 2011; Williamson et al., 2010). To determine whether musicians had superior working memory as compared to non-musicians in the present study, two different independent samples *t*-tests were used to examine whether a difference between groups (Musician vs. Non-Musician) existed for the two different silent conditions (i.e., language-silent, music-silent). The analysis for the language-silent condition revealed that musicians were more accurate (M = 99.81) as compared to the non-musicians (M = 97.76), $t(121.99) = 4.34$, $p < .001$. The same was true for the music-silent condition, where the musicians again were more accurate (M = 95.56) as compared to the non-musicians (M = 87.43), $t(42.82) = 4.03$, $p < .001$ (see Figure 3 and Table 3). The superior working memory of musicians as compared to non-musicians found in the present study is in accord with previous research (Berti et al., 2006; Chan et al., 1998; Cohen et al., 2011; Pallesen et al., 2010; Roden et al., 2014; Schulze et al., 2011; Talamini et al., 2016; Williamson et al., 2010). This provides support for the necessity of using difference scores in the ensuing analysis to control for the relative differences in working memory between groups. To create the difference scores for the subsequent analyses, each participants proportion of correct responses for both the word and chord interference conditions were subtracted from the proportion of

correct responses in the silent condition for the same to-be-remembered stimulus (i.e., word or chord) respectively (note that larger difference scores indicate a greater degree of interference). By using the silent condition as a baseline for working memory performance, these difference scores were meant to account for the relatively enhanced working memory of musicians.

Main Research Questions and Hypotheses

To test the first research question and hypothesis, which was designed to determine whether evidence of shared storage exists by demonstrating that interference was present for both matched and mismatched conditions; a one-way ANOVA was conducted using accuracy scores as the dependent measure. Accuracy scores, as opposed to difference scores, were used in this analysis, as it was necessary to compare both matched and mismatched conditions to that of the silent condition in order to determine whether interference was present. This analysis revealed a main effect of Condition (i.e., silent, matched, mismatched), $F(2, 246) = 237.56$, $p < .001$, $\eta_p^2 = .66$ (see Figure 4 and Table 4). Pairwise comparisons revealed that there was lower accuracy for both the matched ($M = 80.47$) and mismatched ($M = 84.93$) conditions as compared to the silent ($M = 93.41$) condition, and that the matched ($M = 80.47$) condition resulted in poorer performance as compared to the mismatched ($M = 84.93$) condition, $ps < .001$. This demonstrates that there was interference for both matched and mismatched conditions, and that the matched condition was more difficult as compared to the mismatched condition overall. This result is in accord with that of Atherton et al. (2018), where

interference both within and across modalities was found, and supports hypothesis one by demonstrating a degree of interference in both the matched and mismatched conditions.

The second research question was designed to determine whether expertise in music resulted in the phonological loop being more heavily involved in the processing of musical information. To test the two hypotheses derived from this research question a 2 (To-Be-Remembered Stimulus: Word vs. Chord) X 2 (Nature of Overlap: Matched vs. Mismatched) X 2 (Group: Musician vs. Non-Musician) Mixed Factorial ANOVA was conducted using the difference scores as the dependent measure. The first hypothesis for the second research question predicted that more interference would be present for matched as compared to mismatched conditions, whereas the second hypothesis predicted that a larger difference of interference between the matched and mismatched conditions would be present for the non-musicians as compared to the musicians. This analysis revealed a main effect of To-Be-Remembered Stimulus in which performance for to-be-remembered words was superior (M = 1.94) as compared to that of to-be-remembered chords (M = 18.69), $F(1, 122) = 117.70$, $p < .001$, $\eta_p^2 = .49$ (see Figure 5 and Table 5). Further, a main effect of Nature of Overlap was also observed in which performance for the mismatched conditions (M = 7.19) was superior to that of the matched conditions (M = 13.44), $F(1, 122) = 58.60$, $p < .001$, $\eta_p^2 = .32$ (see Figure 6 and Table 6), thus supporting the first hypothesis for the second research question. Both of these effects are consistent with what was observed in Atherton et al. (2018). This suggests that language stimuli are more easily remembered as compared to musical stimuli and/or that interference degrades its retention to a lesser degree, and that interference from the same

modality results in more severe costs in terms of recall accuracy. Additionally, an interaction of To-Be-Remembered Stimulus by Nature of Overlap was observed, $F(1, 122) = 16.00$, $p < .001$, $\eta_p^2 = .12$ (see Figure 7 and Table 7). This interaction was driven by a larger relative difference between to-be-remembered chords in the matched ($M = 23.57$) and mismatched ($M = 12.82$) conditions as compared to the to-be-remembered words in the matched ($M = 3.31$) and mismatched ($M = 0.56$) conditions. This was mainly due to near-ceiling performance for the to-be-remembered word condition. This will be discussed in more detail in the ensuing discussion section.

In terms of differences between those classified as musicians and non-musicians, a three-way interaction of To-Be-Remembered Stimulus by Nature of Overlap by Group was observed, $F(1, 122) = 5.53$, $p = .020$, $\eta_p^2 = .04$ (see Figure 8 and Table 8). This interaction was likely due in part to the previously mentioned near ceiling performance for the to-be-remembered word condition, which decreased the relative difference between the matched and mismatched conditions for to-be-remembered words as compared to to-be-remembered chords. Additionally, there was an interaction of Nature of Overlap by Group that was driven in a manner that was opposed to what was hypothesized for the second hypothesis. This second hypothesis was derived from the second research question. More specifically, this interaction was driven by a larger relative difference for musicians between the matched ($M = 14.17$) and mismatched condition ($M = 5.28$) as compared to non-musicians ($M = 12.71$, $M = 9.10$, respectively for the matched and mismatched conditions), $F(1, 122) = 10.43$, $p = .002$, $\eta_p^2 = .08$ (see Figure 9 and Table 9). This suggests that whereas the phonological loop in non-

musicians may coopt the processing of both phonological and musical information, with expertise in music there seems to be distinct processing of the two different modalities as opposed to the greater recruitment of the processing of musical information within working memory.

Discussion

The primary aim of the current study was to explore whether linguistic and musical information share storage in working memory, and whether expertise in music affects the nature of the recruitment of musical information being processed within working memory. These questions were examined by exploring the impact of intervening linguistic and musical information on musician and non-musician memory for linguistic and musical information. As expected, the initial analysis in which general working memory integrity was assessed revealed that musicians might have superior working memory as compared to non-musicians. This is consistent with previous research (Berti et al., 2006; Chan et al., 1998; Cohen et al., 2011; Pallesen et al., 2010; Roden et al., 2014; Schulze et al., 2011; Talamini et al., 2016; Williamson et al., 2010). Furthermore, this difference suggests that expertise in processing musical information can affect the way in which information is actively processed and/or temporarily stored and that the paradigm used in this study is sensitive to these changes.

The result in which interference was seen in both the matched and mismatched conditions as compared to the silent condition seems to suggest the presence of at least partially shared storage of phonological and musical information within working memory, supporting hypothesis One. The finding that intervening music can affect the recall of phonological information and vice versa suggests that that these types of information may have access to the same limited-capacity working memory structure. If these types of information did not share storage within working memory, then it would be

expected that no interference would be occur in the mismatched conditions, and thus fewer errors would have been made.

The results in which words were recalled with greater accuracy as compared to chords, and matched conditions resulted in more interference as compared to mismatched conditions support the observations of Atherton et al. (2018). The greater accuracy for words as compared to chords is likely because more participants are experts in language and are exposed to language more than chords. Additionally, the matched conditions likely resulted in more interference as compared to the mismatched conditions due to the intervening stimuli being of the same modality, which made it more difficult to discriminate the target stimulus from the comparison stimulus.

Hypothesis Two A was supported in that it was anticipated based upon previous research that more interference would be observed in matched as compared to mismatched conditions. However, the hypothesis in which a larger difference between matched and mismatched conditions for non-musician as compared to musicians would be exhibited was the opposite of what was observed. Instead, a larger difference between matched and mismatched conditions was observed for musicians as compared to non-musicians (see Figure 9 and Table 9). Closer examination of this result reveals an interesting pattern of data. Not only is there a larger difference between matched and mismatched conditions for musicians as compared to non-musicians, but the least amount of interference was observed for the mismatched condition in musicians. This runs counter to the prediction that musical expertise would result in the greater recruitment of musical information within working memory, wherein phonological and musical

information would be treated in a more similar manner as compared to non-musicians based upon the expertise hypothesis. This similarity in processing should result in both types of information interfering more so with each other as compared to non-musicians, which would in turn be reflected in greater error rates for musicians as compared to non-musicians. Further, the degree of interference between the matched and mismatched conditions should have been more similar in musicians as compared to non-musicians, which is not the case. Thus, the observed results are more representative of musical expertise resulting in the processing of musical information through a separate system, one that operates in a similar manner to that of what is generally construed as working memory for phonological information.

Several studies have investigated musical and phonological processing in working memory and have indicated the possibility of separate storage for musical information within working memory for musicians (Pechmann & Mohr, 1992; Schulze et al., 2011, Williamson, Mitchell, Hitch, & Baddeley, 2010a). For example, Pechmann and Mohr (1992) expanded on the study by Deutsch (1970) by investigating the interfering effects of tonal, phonological, and visual stimuli on the retention of tonal information in musicians and non-musicians. Results suggested that, for musicians, only tones interfered with the retention of a target tone. Non-musicians on the other hand showed less accuracy for the target tone across all interfering conditions. These results suggest that for musicians interfering tonal information can disrupt the processing of to-be-recalled tonal information, but this is not the case for interfering phonological or visual stimuli. Thus, it appears as if there is a processing system for musicians that selectively

handles the active retention of musical information, as only excessive musical information (i.e., interference), but not phonological or visual information, interferes with the retention of to-be-remembered musical information.

Using a different paradigm, Williamson et al. (2010a) came to similar a conclusion. They investigated musicians' verbal and musical working memory using a similar interference paradigm as the current study, but instead of participants remembering one target stimulus, the participants in this study had to remember a sequence of stimuli. Results indicated that irrelevant tones significantly disrupted memory for sequences of tones, whilst only irrelevant speech significantly disrupted memory for sequences of letters. These results again support the notion of a separate storage system for musical information in musicians. Similarly, in the current study an interaction of the group by nature of overlap was observed. This interaction was driven by the relative difference between the matched and mismatched conditions for musician and non-musicians. Most importantly, there was a larger difference between matched and mismatched conditions for musicians as compared to non-musicians. Upon closer examination of this interaction, it was found that musicians had less interference in the music-mismatched condition. If expertise in music results in recruitment of the phonological loop for the processing of musical information, then it would be expected that there would be more interference occurring in this condition. This result again might indicate that as one becomes an expert in music, a separate structure devoted to processing music develops. This separate structure would operate similarly to that of the

phonological loop, where this new structure would store and process musical information in working memory.

Expanding on this idea, and as a way to visualize the brain structures devoted to processing verbal and tonal information, Schulze et al. (2011) conducted a study utilizing functional imaging to assess the effects of verbal and tonal interference on the retention of verbal and tonal information between musicians and non-musicians. The results revealed that musicians and non-musicians had similar areas of activation for both verbal and tonal conditions. However, several structures were activated more strongly in musicians. Furthermore, musicians had additional structures that were active for both verbal and tonal conditions. For tonal information these were areas related to sensorimotor encoding, which makes sense if musicians visualize the motions involved in playing their instrument. For verbal information, this was the right insular cortex, which is thought to be involved in processing emotion. Since there were additional structures that were activated separately for the verbal and tonal conditions, the authors suggested this finding could be an indicator of a separate storage system they called the "tonal loop." This would operate similarly to that of the phonological loop, in which it would process and store musical information in a separate working memory system.

As not much research has investigated the "tonal loop," how it functions can only be speculated. The "tonal loop" may have the same components of the phonological loop (i.e., a store and articulatory control process), which may be used to rehearse musical information and recode written musical information into a tonal code. As the "tonal loop" may be considered a part of working memory, it would most likely be able to hold

musical information for only a few seconds and have a limited capacity, like the phonological loop. In fact, Williamson et al. (2010) demonstrated a serial recall effect for tonal information, which may be indicative of working memory span, in which longer chord sequences were more difficult to recall than shorter chord sequences for both musicians and non-musicians. It is also important to note that the authors found a large difference in serial recall performance over two very different sequence lengths (four tones and seven letters). This may suggest that the capacity of the "tonal loop" and phonological loop are different. Furthermore, as it was demonstrated in the current study that musicians might have two separate structures that process phonological and musical information, it could be speculated that information possessing phonological qualities, such as singing, which can be considered a form of music, may not need to be processed by the "tonal loop." To test some of these speculations, it would be necessary to perform some of the basic tests used in working memory research. For example, serial recall tasks could examine capacity, articulatory suppression tasks could detail what information needs to be recoded in order to enter the "tonal loop," and interference tasks could test if singing interferes with the recall of chords. Overall, the aforementioned studies help to explain the possibility of a separate "tonal loop" in working memory for musicians, although more research will need to be conducted to understand further, how the "tonal loop" operates within working memory.

Regarding some limitations of the current study, within the to-be-remembered language conditions there was very little interference occurring for both musicians and non-musicians. This ceiling effect could be because the task was not difficult enough.

Although an attempt was made to increase the difficulty of the task by adding two additional interfering stimuli in comparison to previous work (Atherton et al., 2018), the fact that the intervening words were matched to the intervening chords, thus altering their sound (e.g., higher or lower pitch), while the target and comparison words were not altered might have made it easier to discriminate the target words. Additionally, the recording of participants' answers could have been made easier by either allowing more time for participants to turn the answer sheets over between the trials, or by providing participants with the trial number verbally. This could have reduced the amount of distractions occurring between those specific trials and allowed participants to focus more, which could have resulted in fewer errors. In addition, the fact that participants were in a classroom setting could have certain implications. For example, being surrounded by many other participants could have been distracting and could have led to more errors. Furthermore, since the classrooms were not sound proof, this could have allowed for confounding noises, which could have also led to more errors.

This study investigated the shared storage of musical and phonological information in working memory between musicians and non-musicians. Although the expertise hypothesis was not supported, the data does support the notion of a separate working memory system (i.e., tonal loop) for musicians. Future studies should investigate how the degree of musical expertise (e.g., amount of years playing an instrument) affects the nature of separate storage mechanisms. The degree of separate storage could increase as one increases their musical expertise. Future studies could also investigate the types of materials used in a similar paradigm, such as not matching the

pitch of words to the chords. This change in stimuli might produce results not at ceiling and make the task more difficult for participants. This change in materials could add clarification to how musicians and non-musicians process musical and phonological information, as the materials would accurately represent the modalities being tested. Additionally, future studies could better control for confounds (e.g., noise, group distractions, acoustics) by running their experiment in smaller groups or individually in a sound-controlled room. This could provide the participants with a distraction free environment and allow them to hear the stimuli clearly, which could provide a better depiction of results.

APPENDIX A

Demographics

Please answer the following demographic questions.

1. What is your gender?
   - ☐ Male
   - ☐ Female
   - ☐ Other
2. What is your age? _____

3. How many semesters of college have you taken? _____

4. What is your ethnicity?

____ African-American

____ Asian American

____ Hispanic American

____ Native American

____ White/Caucasian

____ Other

5. What is your primary language? _____

6. What other languages are you fluent at?

_____
_____

7. How long have you been fluent in these languages?

_____

8. What percent of time have you or are you exposed to these other languages?

_____
_____

9. Do you play a musical instrument? If so, which one(s)?

_____
_____

10. How many years have you been playing?

_____

APPENDIX B

Tables

Table 1

Stimuli for the Language Conditions of the Present Study

	Intervening Words						Comparison Words		
Target Word	Dissimilar - e	Dissimilar - i	Dissimilar - o	Dissimilar - u	Dissimilar **	Dissimilar **	Same	Different +	Different -
cap	few	tic	joy	hub	den	bit	cap	cab	gap
mad	gel	rib	mop	dux	rob	fun	mad	man	had
sat	hem	dim	fog	sup	fig	job	sat	sax	bat
cad	vet	mix	don	cur	wig	set	cad	can	fad
mat	web	zip	cox	rum	bed	cop	mat	*mal	hat
sap	beg	win	hot	*hug	dot	bug	sap	*sac	lap
cat	sew	pit	cow	bus	rip	hut	cat	*cal	rat
map	fed	jig	god	mud	set	lug	map	*maw	tap
sad	pen	his	low	nut	wed	met	sad	*san	pad

Note. Every word is three letters long. Each target word consists of the same middle letter ("a"). Also, each beginning letter ("c", "m", and "s") and ending letter ("p", "d", and "t") is used 3 times without repeating a word. Comparison words can be the same or similar by either the first or last letter of the target word. * refers to a word added by Atherton et al. (2018). While not from the list of Coltheart (1993) they are phonologically similar and should not present any added ease or difficulty in a discrimination task. ** refers to words added by the author.

Table 2

Stimuli for the Music Conditions of the Present Study

	Intervening Chords						Comparison Chords		
Target Chord	3+	4+	5+	4-	5-	3-	Same	Different +	Different -
C	A	E	B	A♭	D♭	E♭	C	D	B♭
G	E	B	F♯	E♭	A♭	B♭	G	A	F
D	B	F♯	D♭	B♭	E♭	F	D	E	C
A	F♯	D♭	A♭	F	B♭	C	A	B	G
E	G♭	A♭	E♭	C	F	G	E	F♯	D
F	D	A	E	D♭	G♭	A♭	F	G	E♭
B♭	G	D	A	G♭	B	D♭	B♭	C	A♭
E♭	C	G	D	B	E	G♭	E♭	F	D♭
A♭	F	C	G	E	A	F♯	A♭	B♭	G♭

Note. The "Different+" column refers to a clockwise rotation on circle of fifths and the "Different-" column refers to a counter-clockwise rotation on circle of fifths. Additionally, "Different" refers to 2 turns on the circle of fifths in the indicated direction.

Table 3

Initial results of overall working memory

	Condition	
Group	Language-Silent	Music-Silent
	M (SD)	*M (SD)*
Non-Musicians	97.76 (4.42)	87.43 (12.20)
Musicians	99.81 (0.83)	95.56 (7.30)

Note. Scores are presented as proportions and reflect accuracy.

Table 4

Main effect of condition

Condition	*M (SD)*
Silence	93.41 (6.94)
Matched	80.47 (8.00)
Mismatched	84.93 (8.43)

Note. Scores are presented as proportions and reflect amount of interference.

Table 5

Main effect of to-be-remembered stimulus

To-be-remembered stimulus	*M (SD)*
Words	1.94 (8.95)
Chords	18.70 (22.20)

Note. Scores are presented as proportions and reflect amount of interference.

Table 6

Main effect of nature of overlap

Nature of overlap	*M (SD)*
Matched	13.44 (6.93)
Mismatched	7.20 (6.27)

Note. Scores are presented as proportions and reflect amount of interference.

Table 7

Interaction of to-be-remembered stimulus by nature of overlap

	Nature of overlap	
To-be-remembered stimulus	Matched	Mismatched
	M (SD)	*M (SD)*
Word	3.31 (6.39)	0.60 (3.76)
Chord	23.57 (13.15)	13.82 (12.64)

Note. Scores are presented as proportions and reflect amount of interference.

Table 8

3-way interaction of to-be-remembered stimulus by nature of overlap by group

	Condition			
Group	Word Match	Word Mismatch	Chord Match	Chord Mismatch
	M (SD)	*M (SD)*	*M (SD)*	*M (SD)*
Non-Musicians	2.92 (6.70)	0.75 (4.04)	22.51 (13.40)	17.45 (12.62)
Musicians	3.70 (4.50)	0.37 (1.66)	24.63 (12.00)	10.19 (11.14)

Note. Scores are presented as proportions and reflect amount of interference.

Table 9

Interaction of nature of overlap by group

	Nature of overlap	
Group	Matched	Mismatched
	M (SD)	*M (SD)*
Non-musicians	12.71 (7.11)	9.10 (6.26)
Musicians	14.17 (5.90)	5.28 (5.42)

Note. Scores are presented as proportions and reflect amount of interference.

APPENDIX C

Figures

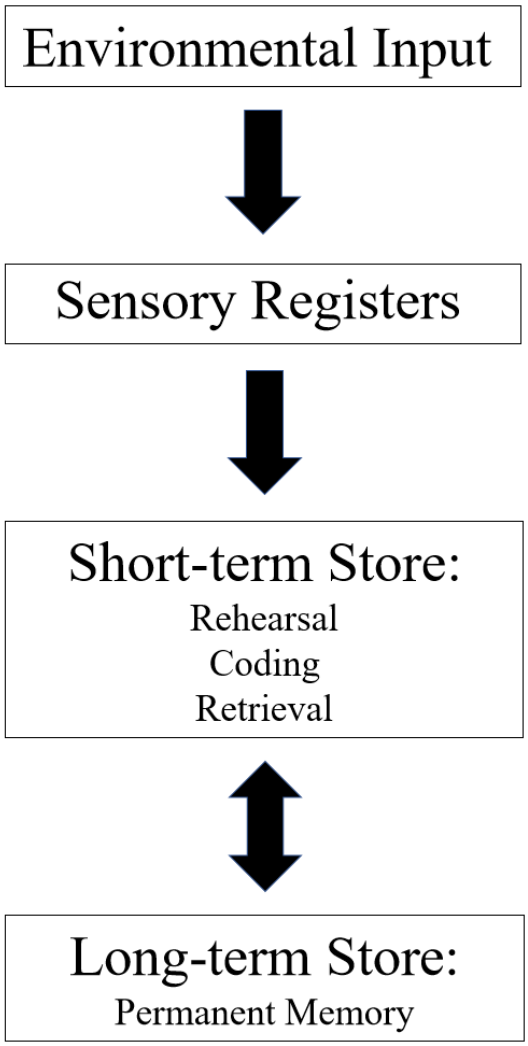

Figure 1. Representation of the Atkinson & Shiffrin (1968) modal model of short-term memory. In this model, environmental sensory input enters the appropriate sensory registers briefly before being processed in the short-term memory store. In this store, rehearsal, coding, and retrieval of information occurs. Information can flow back and forth between the short-term and long-term stores.

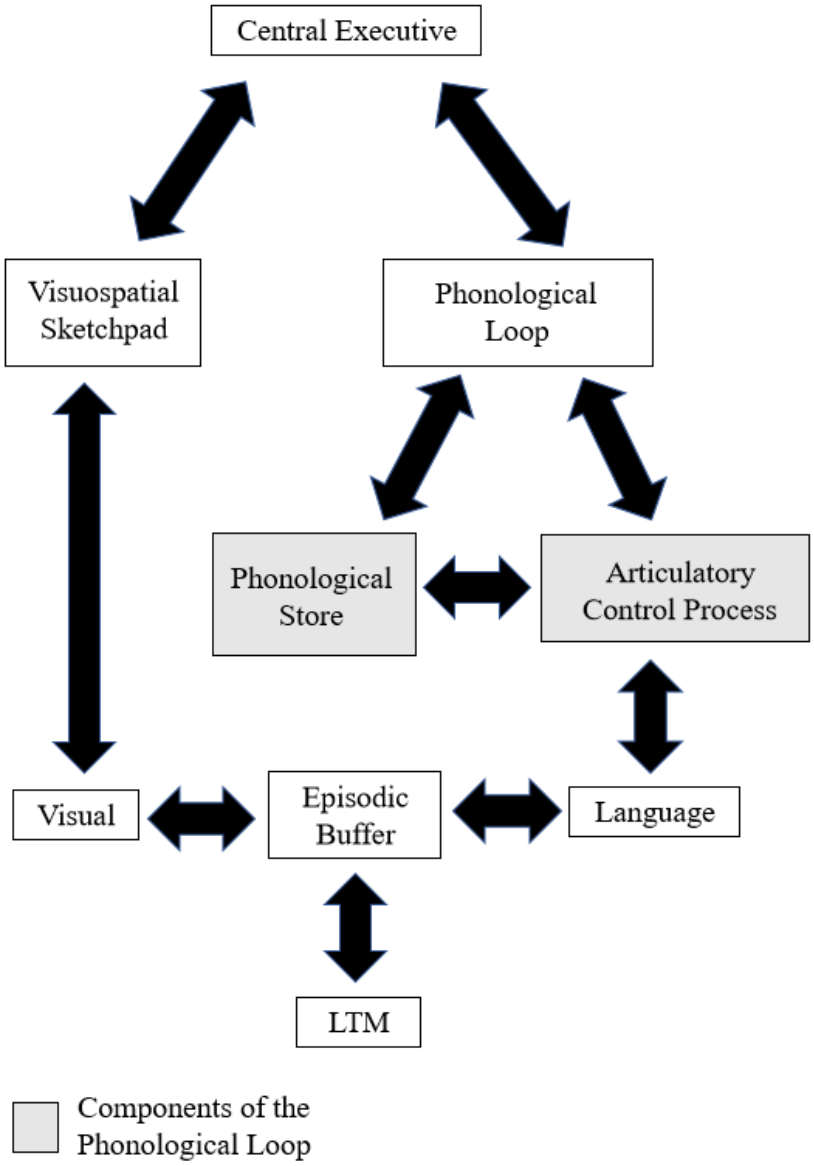

Figure 2. Representation of the Baddeley & Hitch (1974) Multicomponent Working Memory Model with their additional component (Baddeley, 2000). The central executive coordinates information to either the visuospatial sketchpad or the phonological loop. The visuospatial sketchpad processes visual and spatial information, whereas the phonological loop processes language-based information. The phonological loop consists of the phonological store, which briefly holds information, and the articulatory control process, which rehearses information within the store and recodes visual information into a phonological code. The episodic buffer connects information from the visuospatial sketchpad, phonological loop, and LTM.

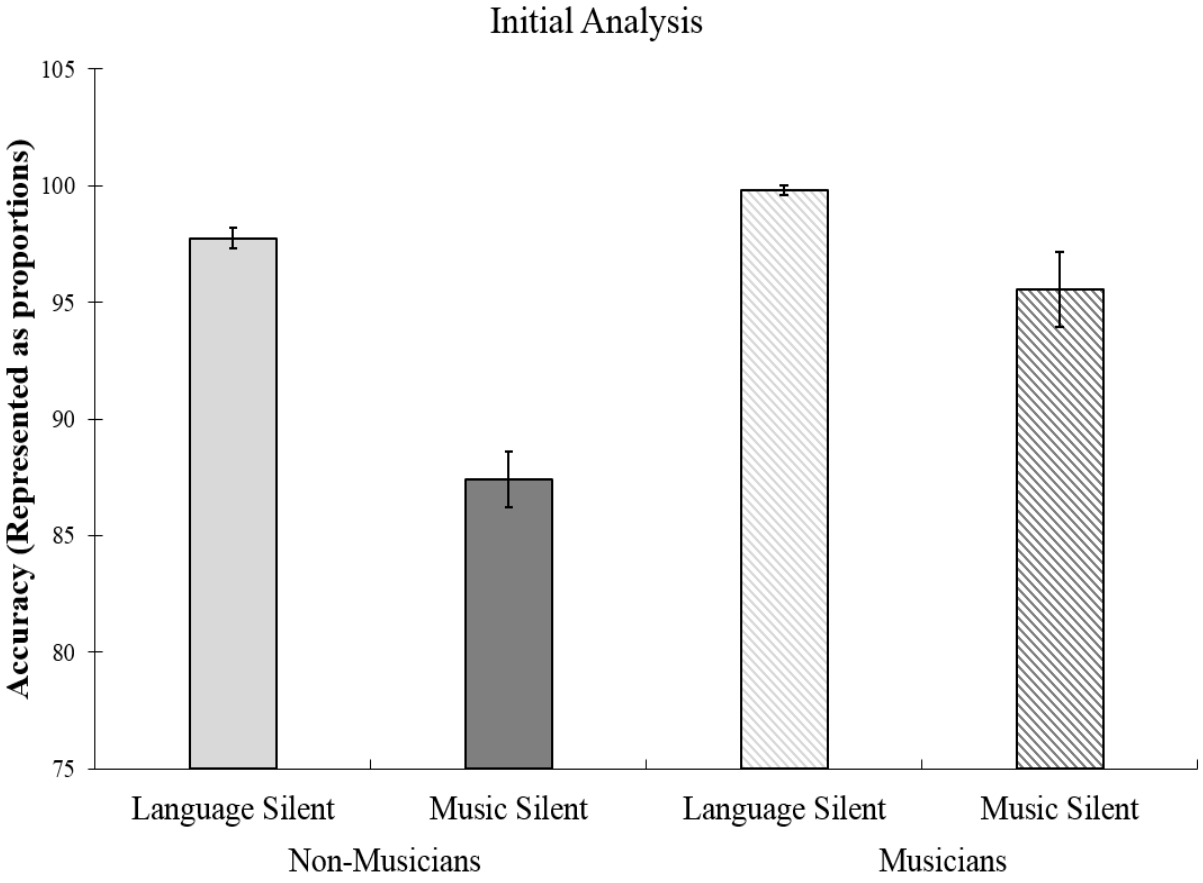

Figure 3. Representation of the accuracy for each silent condition for non-musicians and musicians. Both groups show similar profiles, but musicians are significantly more accurate in both the language silent and music silent condition compared to the non-musicians.

Figure 4. Representation of the accuracy for the silent, matched, and mismatched conditions. The silent condition had the highest accuracy, followed by the mismatched and matched conditions.

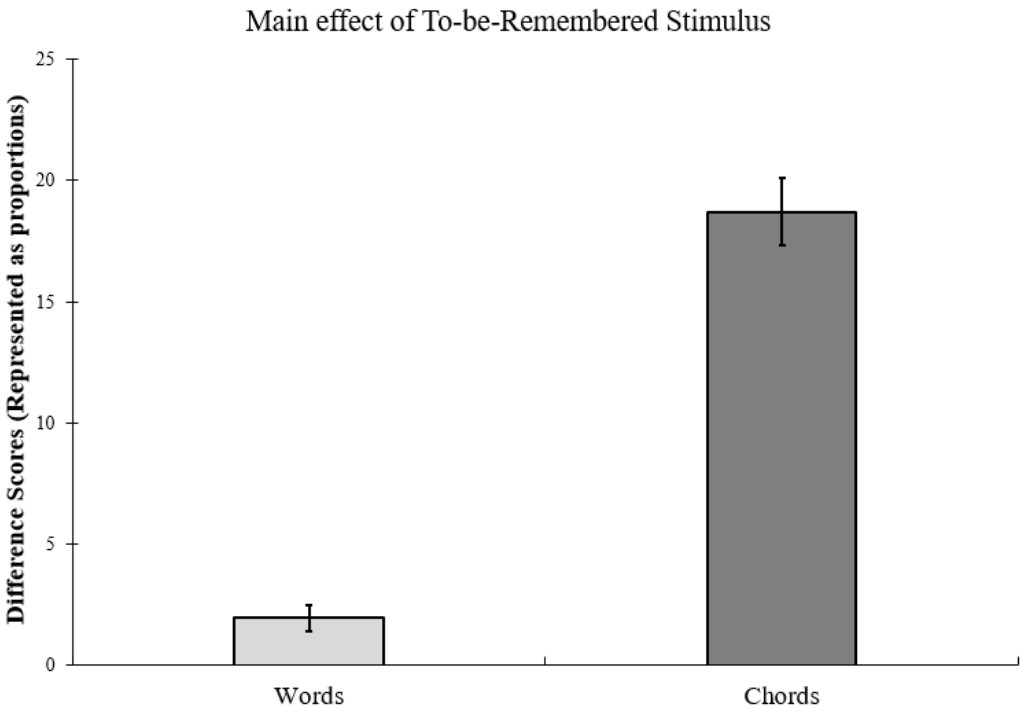

Figure 5. Representation of the difference scores for the word and chord conditions. Higher scores indicate more interference, thus less accuracy. The chord condition had significantly less accuracy than the word condition.

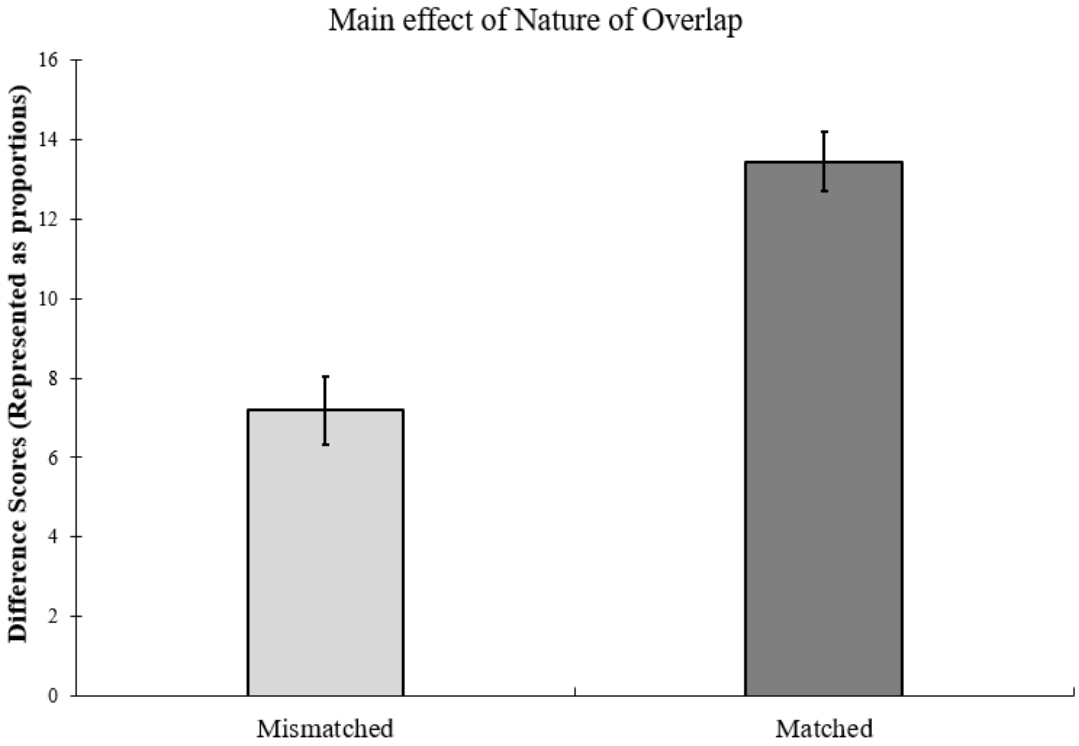

Figure 6. Representation of the difference scores for the matched and mismatched conditions. Higher scores indicate more interference, thus less accuracy. The matched condition had significantly less accuracy than the mismatched condition.

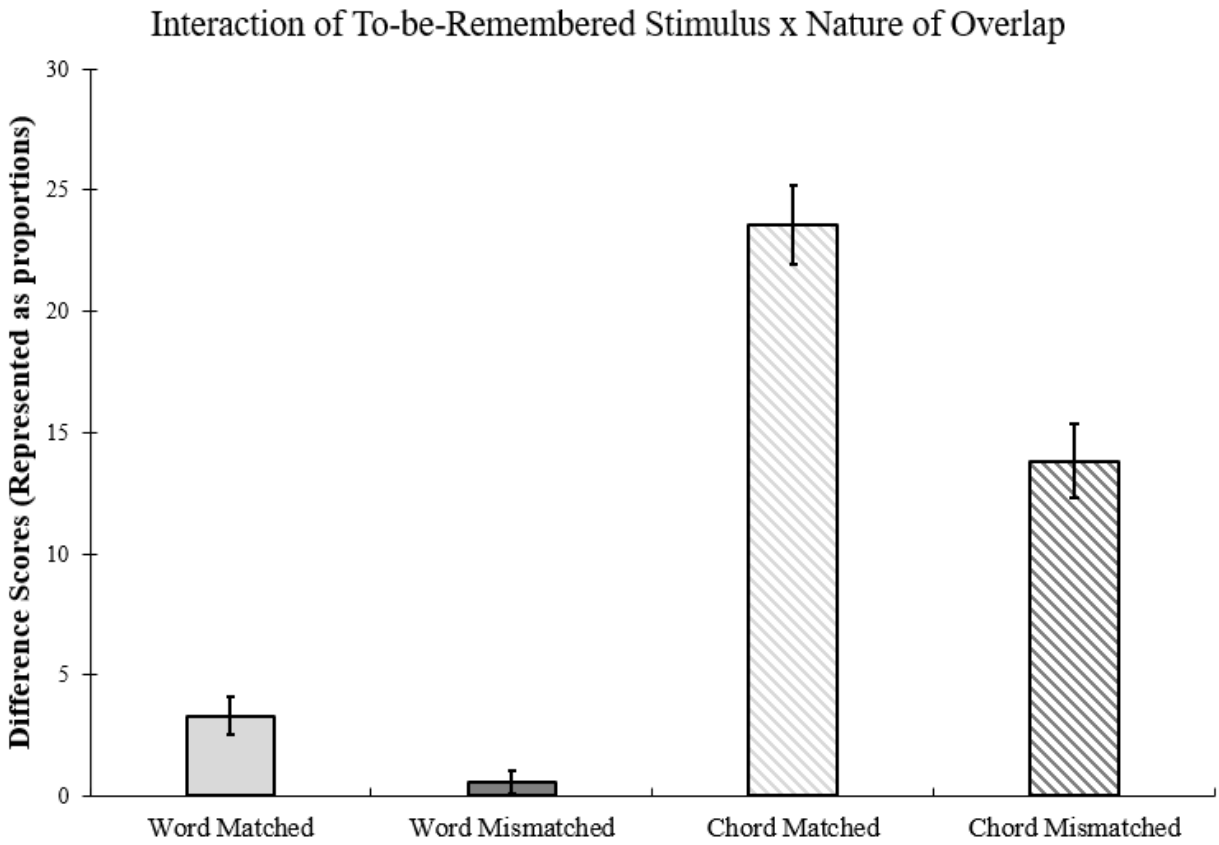

Figure 7. Representation of the difference scores for the word and chord matched and mismatched conditions. Higher scores indicate more interference, thus less accuracy. There is a larger relative difference between to-be-remembered chords in the matched and mismatched conditions as compared to to-be-remembered words in the matched and mismatched conditions.

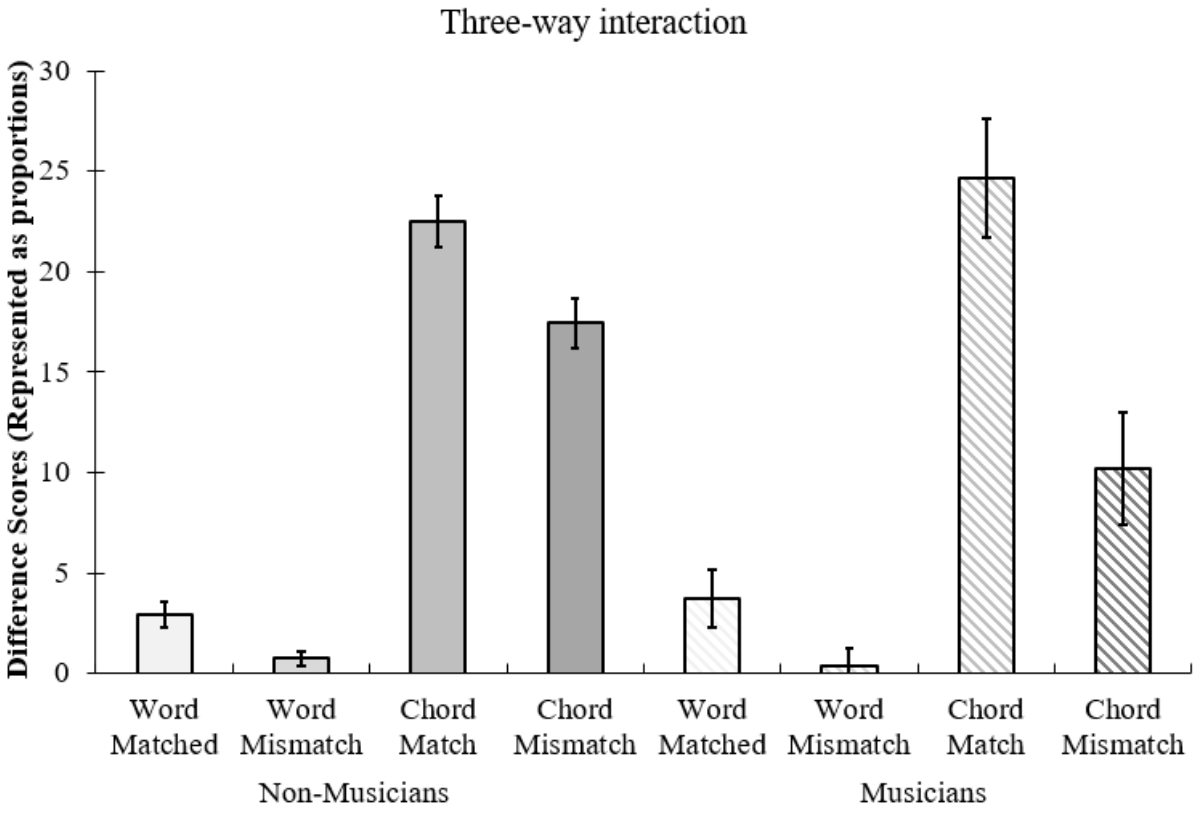

Figure 8. Representation of the difference scores for all conditions across non-musicians and musicians. Higher scores indicate more interference, thus less accuracy. There appears to be a ceiling effect for the to-be-remembered word condition, which decreased the relative difference between the matched and mismatched conditions for the to-be-remembered words as compared to the to-be-remembered chords.

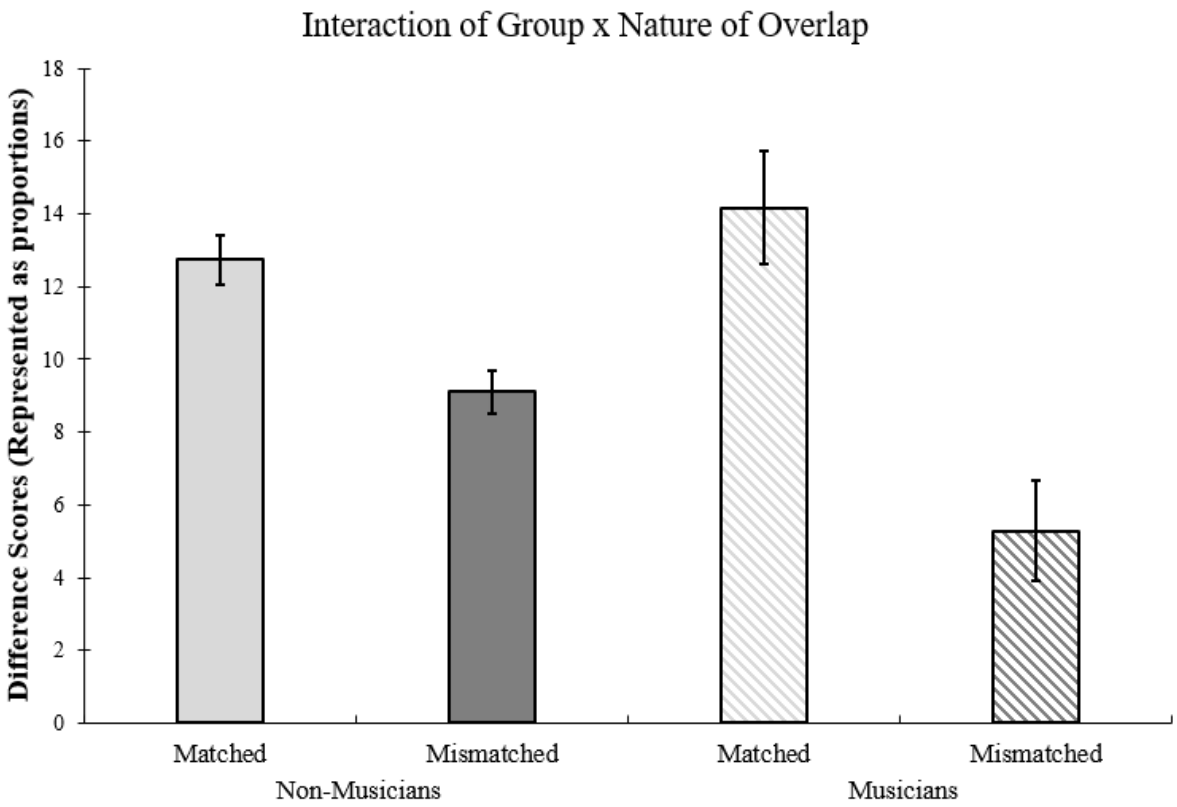

Figure 9. Representation of the difference scores for matched and mismatched conditions for non-musicians and musicians. Higher scores indicate more interference, thus less accuracy. There is a larger relative difference for musicians between the matched and mismatched conditions as compared to the non-musicians.

References

Atherton, R. P., Chrobak, Q. M., Rauscher, F. H., Karst, A. T., Hanson, M. D., Steinert, S. W., & Bowe, K. L. (2018), Shared processing of language and music: Evidence from a cross- modal interference paradigm, *Experimental Psychology, 65*(1), 40-48. doi:10.1027/1618-3169/a000388

Atkins, P. W. B., & Baddeley, A. D. (1998). Working memory and distributed vocabulary learning. *Applied Psycholinguistics, 19*(4), 537-552. doi:10.1017/S0142716400010353

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory,* 42–775. New York, NY: Academic Press.

Baddeley, A. D. (1992). Working memory. *Science, 255,* 556-559. doi:10.1126/science.1736359

Baddeley, A. D. (1993). Short-term phonological memory and long-term learning: A single case study. *European Journal of Cognitive Psychology, 5*, 129-148. doi:10.1080/09541449308520112

Baddeley, A. D. (1996). Exploring the central executive. *Quarterly Journal of Experimental Psychology*. *49A*(1), 5-28. doi:10.1080/027249896392784

Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences, 4*(11), 417-423. doi:10.1016/S1364-6613(00)01538-2

Baddeley, A. D. (2001). Is working memory still working? *European Psychologist, 7*, 85-97. doi:10.1027//1016-9040.7.2.85

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

Baddeley, A. D. (2003b). Looking back and looking forward. *Nature Reviews Neuroscience, 4*, 829-839. doi:10.1038/nrn1201

Baddeley, A. D. (2004). The psychology of memory. In A. Baddeley, M. Kopelman, & A.Wilson (Eds.) *The Essential Handbook of Memory Disorders for Clinicians*, (pp. 1-13). York, UK: John Wiley & Sons, Ltd.

Baddeley, A. D. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology, 63*, 1-29. doi:10.1146/annurev-psych-120710-100422

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory*, 8, (pp.47-89). New York: Academic Press.

Baddeley, A. D., & Hitch, G. J. (1994). Developments in the concept of working memory. *Neuropsychology, 8*(4), 485-493. doi:10.1037/0894-4105.8.4.485

Baddeley, A. D., Lewis, V., Eldridge, M., & Thomson, N. (1984). Attention and retrieval from long-term memory. *Journal of Experimental Psychology: General, 113*(4), 518-540. doi:10.1037/0096-3445.113.4.518

Baddeley, A. D., Papagno, C., & Vallar, G. (1988). When long-term learning depends on short-term storage. *Journal of Memory and Language, 27*, 586-595. doi:10.1016/0749-596X(88)90028-9

Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior, 14*(6), 575-589. doi:10.1016/S0022-5371(75)80045-4

Baddeley, A. D., Vallar, G., & Wilson, B. (1987). Sentence comprehension and phonological memory: Some neuropsychological evidence. In M. Coltheart (Ed.), Attention and Performance XII (pp. 509–529). London: Lawrence Erlbaum Associates.

Berti, S., Munzer, S., Schroger, E., & Pechmann, T. (2006). Different interference effects in musicians and a control group. *Experimental Psychology, 53*(2), 111-116. doi:10.1027/1618-3169.53.2.111

Biedron, A., & Szczepaniak, A. (2012). Working memory and short-term memory abilities in accomplished multilinguals. *The Modern Language Journal, 96*(2), 290-306. doi:10.1111/j.1540-4781.2012.01332.x

Brooks, L. R. (1968). Spatial and verbal components of the act of recall. *Canadian Journal of Psychology, 22,* 349-366. doi:10.1037/h0082775

Chan, A. S., Ho, Y., & Cheung, M. (1998). Music training improves verbal memory. *Nature, 396,* 128. doi:10.1038/24075

Chang, Y. (2014). Reorganization and plastic changes of the human brain associated with skill learning and expertise. *Frontiers in Human Neuroscience, 8*(35*)*, 1-7. doi:10.3389/fnhum.2014.00035

Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences. New York, NY: Routledge Academic.

Cohen, M. A., Evans, K. K., Horowitz, T. S., & Wolfe, J. M. (2011). Auditory and visual memory in musicians and nonmusicians. *Psychonomic Bulletin Review, 18*(3), 586-591. doi:10.3758/s13423-011-0074-0

Coltheart, V. (1993). Effects of phonological similarity and concurrent irrelevant articulation on short-term-memory recall of repeated and novel word lists. *Memory & Cognition, 21*(4), 539-545. doi:10.3758/BF03197185

Conrad, R., & Hull, A. J. (1964). Information, acoustic confusion and memory span. *British Journal of Psychology*, 55, 429–432. doi:10.1111/j.20448295.1964.tb00928.x

Cowan, N. (2005). Working Memory Capacity. Hove, UK: Psychology Press.

Cowan, N., Day, L., Saults, J. S., Keller, T. A., Johnson, T., & Flores, L. (1992). The role of verbal output time in the effects of word length on immediate memory. *Journal of Memory and Language, 31,* 1-17. doi:10.1016/0749-596X(92)90002-F

Della Sala, S., Gray, C., Baddeley, A., Allamano. N., & Wilson L. (1999). Pattern span: A tool for unwelding visuo-spatial memory. *Neuropsychologia, 37*, 1189-99. doi:10.1016/S0028-3932(98)00159-6

Deutsch, D. (1970). Tones and numbers: Specificity of interference in immediate memory. *Science, 168,* 1604-1605. doi:10.1126/science.168.3939.1604

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39,* 175-191. doi:10.3758/BF03193146

Gaser, C., & Schlaug, G. (2003). Brain structures differ between musicians and non-musicians. *The Journal of Neuroscience, 23*(27), 9240-9245.

Gauthier, I., Tarr, M. J., Anderson, A. W., Skudlarski, P., & Gore, J. C. (1999). Activation of the middle fusiform "face area" increases with expertise in recognizing novel objects. *Nature Neuroscience, 2,* 568–573. doi:10.1038/9224

Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience, 3,* 191–197. doi:10.1038/72140

Kaganovich, N., Kim, J., Herring, C., Schumaker, J., MacPherson, M., & Weber-Fox, C. (2013). Musicians show general enhancement of complex sound encoding and better inhibition of irrelevant auditory change in music: An ERP study. *European Journal of Neuroscience, 37,* 1295-1307. doi:10.1111/ejn.12110

Klauer, K. C., & Zhao, Z. (2004). Double dissociations in visual and spatial short-term memory. *Journal of Experimental Psychology: General, 133*(3), 355-81. doi:10.1037/0096-3445.133.3.355

Logie, R. H. (1986). Visuo-spatial processing in working memory. *Quarterly Journal of Experimental Psychology, 38A*, 229-247. doi:10.1080/14640748608401596

Murray, D. J. (1968). Articulation and acoustic confusability in short-term memory. *Journal of Experimental Psychology, 78*, 679-684. doi:10.1037/h0026641

Pallesen, K. J., Brattico, E., Bailey, C.J., Korvenoja, A., Koivisto, J., Gjedde, A., & Carlson, S. (2010). Cognitive control in auditory working memory is enhanced in musicians. *PLoS ONE 5*(6): e11120. doi:10.1371/journal.pone.0011120

Papagno, C., Valentine, T., & Baddeley, A. D. (1991). Phonological short-term memory and foreign language vocabulary learning. *Journal of Memory and Language, 30*, 331-347. doi:10.1016/0749-596X(91)90040-Q

Papagno, C., & Vallar, G. (1992). Phonological short-term memory and the learning of novel words: The effect of phonological similarity and item length. *Quarterly Journal of Experimental Psychology, 44A*(1), 47-67. doi:10.1080/14640749208401283

Pechmann, T., & Mohr, G. (1992). Interference in memory for tonal pitch: Implications for a working-memory model. *Memory & Cognition, 20*(3), 314–320. doi:10.3758/BF03199668

Phillips, W. A. (1974). On the distinction between sensory storage and short-term visual memory. *Perception and Psychophysics, 16*(2), 283–90. doi:10.3758/BF03203943

Phillips, W. A., & Baddeley, A. D. (1971). Reaction time and short-term visual memory. *Psychonomic Science, 22*(2), 73–74. doi:10.3758/BF03332500

Posner, M. I., & Konick, A.F. (1966). Short-term retention of visual and kinesthetic information. *Organizational Behavior and Human Performance, 1*, 71–86. doi:10.1016/0030-5073(66)90006-7

Roden, I., Grube, D., Bongard, S., & Kreutz, G. (2014). Does music training enhance working memory performance? Findings from a quasi-experimental longitudinal study. *Psychology of Music, 42*(2), 284-298. doi:10.1177/0305735612471239

Schlaug, G., Jancke, L., Huang, Y., Staiger, J. F., & Steinmetz, H. (1995). Increased corpus callosum size in musicians. *Neuropsychologia, 33*(8), 1047-1055. doi:10.1016/0028-3932(95)00045-5

Schroeder, S., Marian, V., Shook, A., & Bartolotti, J. (2016). Bilingualism and musicianship enhance cognitive control. *Neural Plasticity,* 1-11. doi:10.2266/2016/4058620

Schulze, K., Zysset, S., Mueller, K., Friederici, A. D., & Koelsch, S. (2011). Neuroarchitecture of verbal and tonal working memory in nonmusicians and musicians. *Human Brain Mapping, 32,* 771-783. doi:10.1002/hbm.21060

Semal, C., Demany, L., Ueda, K., & Halle, P. (1996). Speech versus nonspeech in pitch memory. *Acoustical Society of America, 100*(2), 1132-1140. doi:10.1121/1.416298

Service, E. (1992). Phonology, working memory and foreign-language learning. *Quarterly Journal of Experimental Psychology, 45A*, 21-50. doi:10.1080/14640749208401314

Shallice, T., & Warrington, E. (1974). The dissociation between the short-term retention of meaningful sounds and verbal material. *Neuropsychologia, 12*(4), 553-555. doi:10.1016/0028-3932(74)90087-6

Talamini, F., Carretti, B., & Grassi, M. (2016). The working memory of musicians and nonmusicians. *Music Perception, 34(2),* 183-191. doi:10.1525/2016.34.2.183

Weijkamp, J., & Sadakata, M. (2017). Attention to affective audio-visual information: Comparison between musicians and non-musicians. *Psychology of Music, 45*(2), 204-215. doi:10.1177/03057356166542156

Williamson, V. J., Baddeley, A. D., & Hitch, G. J. (2010). Musicians' and nonmusicians' short-term memory for verbal and musical sequences: Comparing phonological similarity and pitch proximity. *Memory & Cognition, 38(2),* 163-175. doi:10.3758/MC.38.2.163

Williamson, V. J., Mitchell, T., Hitch, G. J., & Baddeley, A. D. (2010a). Musicians' memory for verbal and tonal materials under conditions of irrelevant sound. *Psychology of Music, 38(3),* 331-350. doi:10.1177/0305735609351918

Zhang, L., Weiwei, P., & Hu, L. (2015). Electrophysiological evidences demonstrating differences in brain functions between nonmusicians and musicians. *Scientific Reports, 5,* 1-14. doi:10.1038/srep13796