WORKING MEMORY FOR MUSICAL AND VERBAL MATERIAL UNDER CONDITIONS OF IRRELEVANT SOUND

by

Kristi M. Von Handorf

A thesis submitted in partial fulfillment of the requirements
for graduation with Honors in Psychology.

Whitman College
2014

Certificate of Approval

This is to certify that the accompanying thesis by Kristi M. Von Handorf has been accepted in partial fulfillment of the requirements for graduation with Honors in Psychology.

____

Matthew Prull

Whitman College
May 14, 2014

Abstract

Baddeley and Hitch's (1974) working memory model concerns the storage and processing of information in the short term. The present research suggests possible changes to the model because the model does not account for the storage and processing of music. Previous studies have found evidence that musical memory should not be considered part of the phonological loop, which stores language information, and that it may require a separate loop altogether. This assertion has been tested by examining the size of the irrelevant sound effect across modalities through the new visual-auditory recognition method. Previous research has found that irrelevant sound in the form of tones only disrupts memory for tones, whereas irrelevant sound in the form of speech only disrupts memory for letters. This modality-specific interference effect suggests that processing of musical and verbal material occurs in separate stores. Although previous studies have introduced the irrelevant sound for the duration of the trial, the present study separated the placement of the irrelevant sound to occur 1) simultaneously with the visual sequence and 2) during the retention interval only, in order to rule out encoding or masking effects that might have confounded previous findings. Though none of the results was statistically significant, patterns in memory scores indicated a modality-specific effect in the letters condition and a general distraction effect in the tones condition. Further work is needed to make more definite conclusions about the nature of memory for musical material.

Working Memory for Verbal and Musical Material Under Conditions of Retroactive Irrelevant Sound

Music is ubiquitous, present in all cultures, and existing as far back as known history reaches (Wallin, Merker, & Brown, 2000). The study of music from a psychological point of view has been a new and exciting development over the past few decades. Psychologists study such topics as music perception, emotion, performance, and memory. Memory is an important area of study, as it forms the basis for people's experiences with music, from any person remembering popular tunes and singing along with the radio to professional musicians memorizing entire concertos. This paper primarily concerns working memory for music. Working memory is distinct from short-term memory in that working memory implies a combination of storage and processing, whereas short-term memory is conceptualized as the simple temporary storage of information. Understanding how musical memory works may inform under what conditions music should be learned and practiced for best retention. This knowledge could lead to greater efficiency in both practice and performance. Additionally, not much research exists comparing short-term processing of both verbal and tonal material in parallel.

Most research on working memory has focused on verbal and visual material. Baddeley (2012) proposed a widely supported model of working memory, in which a central executive coordinates attention between relevant and irrelevant information. Two components feed into the central executive that are responsible for both storage and processing of information in memory: the phonological loop and the visuospatial sketchpad. The phonological loop stores language information and keeps it in memory via articulatory rehearsal (inner speech). Baddeley cited several well-established effects as evidence for the existence of the phonological loop, including articulatory suppression (Salamé & Baddeley, 1982; Schendel & Palmer, 2007) and irrelevant sound effects (Colle & Welsh, 1976). The irrelevant sound effect arises when participants are asked to remember a set of digits, but they hear irrelevant speech for several seconds during the presentation of the digits until they are asked to recall those digits. Results from experiments using this method indicate that irrelevant sound significantly reduces recall of the digits relative to silence. The effect occurs even when participants are instructed to ignore any intervening sound and when the irrelevant speech is in a different language (Colle & Welsh, 1976). Attention is divided between the digits and competing information, which suggests that there is some kind of automatic processing of the information held in short term memory. If one is unable to engage in articulatory rehearsal, then the information is lost. Articulatory suppression works in a similar way. When participants are required to continuously speak a single word such as "the", they are not able to convert written to-be-remembered material into inner speech, and memory is disrupted.

By comparison, little attention has been paid to the workings of short-term memory for music specifically. There have been some preliminary conclusions about the size and nature of short-term memory for music. Melodic memory capacity in short-term memory is approximately 7-11 notes, depending on such factors as contour, tonality, and range (Pembrook, 1987), although some researchers argue that peak capacity is between 11 and 15 pitches (Long, 1977). As with language, music held in short-term memory is subject to disruption. Some of the earliest research in the study of musical short-term memory capacity examined various factors that influenced recall of a given tone after intervening material was presented (Bull & Cuddy, 1972; Elliott, 1970; Wickelgren, 1966, 1969). Deutsch (1970) presented a tone aurally and then presented either six extra tones or six spoken numbers before memory for the original tone was tested by asking whether a second tone was either the same or different from the first tone. When the distracting material consisted of spoken numbers, participants made tone-recognition errors only 3% of the time on average. However, when the distracting materials were extra tones, participants made errors 32% of time. The fact that participants made far more errors remembering tones accurately when the distracting sounds were also musical in nature suggests that there is a separate storage loop for musical material. These results led Berz (1995) to develop a theoretical model of working memory, in which a music memory loop is connected to the central executive, separate from the phonological loop.

However, some researchers assert that music is processed in the same way as language and should not be separate from the phonological loop. Jones and Macken (1993) argued that Deutsch's (1970) results occurred not because tones and speech are stored separately, but because of the way the materials were grouped perceptually. The initial tone was likely to be grouped perceptually with, and hence be difficult to differentiate from, the intervening tones. However, when the intervening material consisted of spoken numbers, the initial and final tones would have been easier to differentiate from the distracting material.

Salamé and Baddeley (1982) researched the irrelevant sound effect as it related to both visual and verbal material. They suggested that visually-presented sequences of verbal material are remembered via the articulatory loop system. This process involves visual items being recoded through subvocalization, which stores the material phonologically. Therefore, the irrelevant sound effect is assumed to occur because spoken material also gains access to that phonological store and interferes with its ability to retain previous material. Salamé and Baddeley suggested that some kind of filter governs which sounds gain access to the phonological store. They distinguished between "noise" and "speech," which raised the question: what makes speech speechlike? In attempting to answer this question, Salamé and Baddeley (1989) believed that the highly structured and patterned nature of speech was crucial. In this case, music, which is also highly structured, would gain access to the phonological store and disrupt performance. An alternative perspective would be that some aspect of the human voice is the crucial element in making speech speechlike, in which case neither noise nor music would disrupt performance.

To test these possibilities, the researchers asked participants to complete a digit memory task while they heard different types of background material: silence, vocal music, or instrumental music. Participants saw a sequence of nine digits, were instructed to ignore any music, and wrote down the digits after 13 seconds. Salamé and Baddeley found that irrelevant speech caused significantly greater disruption to recall of verbal material than vocal music, which in turn caused greater disruption than instrumental music. They proposed that, because vocal music has more acoustic features in common with speech than instrumental music, vocal music created more interference in memory for the digits. If there were a single acoustic store for both musical and verbal material, then irrelevant instrumental music would cause the same degree of disruption as irrelevant speech or vocal music, but that was not the case. Although the results might seem to support a separate store for musical material, Salamé and Baddeley explained the presence of a filter that governs access to the phonological store based on certain characteristics, such as being speech-like. In that case, musical and verbal material could be part of the same acoustic store, and the difference in disruption between vocal and instrumental music would occur because vocal music has more acoustic features in common with subvocal speech than does instrumental music.

Schendel and Palmer (2007) weighed in on the argument by developing a new recognition procedure involving a cross in presentation modalities from visual stimuli to auditory stimuli. In the visual-auditory recognition method, participants first *see* a sequence of notes on a staff and are later asked to indicate whether a sequence that they *hear* is the same or different. This design encourages participants to convert the visual image of music into an auditory one for rehearsal, the same way that the process would work if they were to rehearse a given sequence of letters. In three experiments, adult musicians with at least six years of experience engaged in articulatory suppression while trying to remember either tone or digit sequences. The musical suppression condition included singing the syllable “la” after the onset of the to-be-remembered sequence until the presentation of the comparison sequence. The verbal suppression condition included speaking the syllable “the” for the same amount of time. Of the three experiments, only the last experiment consisted of a change in presentation modalities (i.e., participants were presented with visual sequences and asked if an auditory sequence was the same or different). In their third experiment, Schendel and Palmer found a modality-specific interference effect, in that memory for tones was disrupted only when the participants engaged in musical suppression, and memory for digits was disrupted only when participants engaged in verbal suppression. The modality-specific effect suggests that there are separate stores for musical and verbal processing, because, if there were a singular store, musical suppression should have the same effects on verbal memory as would verbal suppression. However, Schendel and Palmer (2007) argued a more specific account that the same mechanisms are responsible for the storage and rehearsal of verbal material and auditorily encoded music, given sufficient experience with both. They suggested that the modality-specific effect occurred within the visual-auditory recognition paradigm because of the unique integration of visual and auditory cues in a single representation, but researchers do not yet know how integration occurs across sensory modalities.

Wiliamson, Mitchell, Hitch, and Baddeley (2010) expanded upon Schendel and Palmer's (2007) experiments by using the new visual-auditory recognition method to examine the irrelevant sound effect. Instead of participants singing or speaking syllables during the retention interval, they were exposed to various conditions of irrelevant sound similar to the ones that Deutsch (1970) used. Thirty-two amateur and professional musicians took part in the study, all of whom had at least eight years of training on an instrument or voice. On average, the participants reported 16 years of training. In Williamson et al.'s (2010) experiment, half the participants were asked to remember sequences of tones, and the other half were asked to remember sequences of letters. Participants studied the visual sequence of tones or letters for eight seconds, after which the screen went blank for 10 seconds. During each trial, participants heard either irrelevant tones, irrelevant speech, white noise, or no irrelevant sound. Then, an auditory sequence played that was either the same as or different than the visual sequence. Like

Schendel and Palmer (2007), Williamson et al. (2010) found a modality-specific interference effect. The results corroborate Berz's assertion that there are separate stores for musical and verbal processing.

The focus of the present study was to add more evidence for the existence of a modality-specific effect and to determine further what processes are responsible for the effect. In Williamson et al.'s (2010) study, it is difficult to determine why or how the modality specific effect occurred because the irrelevant sound was presented for the duration of the trial. Thus, the findings could be explained by an encoding effect, in which greater interference occurs when similar materials overlap in memory. In other words, trying to encode a visual tone sequence in memory would be disrupted by immediately hearing irrelevant tones, whereas encoding would not be interrupted if the participant were to immediately hear irrelevant digits.

In the current experiment, the original 13-second irrelevant sound is split into two conditions, with each irrelevant sound sequence lasting for seven seconds. In the first, sound occurs simultaneously with the to-be-remembered visual stimulus and lasts until the stimulus disappears. In the second condition, sound occurs after the visual stimulus disappears and lasts until the comparison sequence is played. The split allows for a distinction between interference to the encoding process and interference to the storage process. Research on retroactive irrelevant sound effects (those presented after the stimulus disappears) is relatively uncommon, but the effects are still significant (Deutsch, 1970; Norris, Page, & Baddeley, 2004). The presence of retroactive irrelevant sound effects in these studies suggest that the locus of the irrelevant sound effect is during storage.

The first goal of the current experiment was to replicate Williamson et al.’s (2010) findings. If musical and verbal sounds exist in a singular acoustic store, then each should be disrupted equally by the irrelevant sounds, whether musical or verbal. If a modality-specific effect occurs, it would provide further support for Berz’s (1995) model that there are separate acoustical stores for musical and verbal sounds. The second goal was to determine whether the results change when the irrelevant sound is presented retroactively. A final goal was to consider a wider range of musical expertise, because previous studies (Schendel & Palmer, 2007; Williamson et. al, 2010) used musicians with at least several years of formal experience.

Method

Participants

The sample consisted of 40 students, staff, and faculty from Whitman College as well as members of the greater Walla Walla community. All participants had music-reading ability. The sample was 57.5% male and 42.5% female. Participants' ages ranged from 19 to 60 (M = 28.00, SD = 12.93) and years of formal study in music ranged from one to 25 (M = 10.45, SD = 5.90). The names of all participants were entered into a raffle in which one winner received a $40 gift card to Starbucks. A convenience sample was used because of limited access to a large sample and limited funds to compensate participants.

Design

This experiment was a 2 (stimulus type: letters vs. tones) X 5 (irrelevant sound type and placement: simultaneous spoken digits, delayed spoken digits, simultaneous tones, delayed tones, or silence) mixed design. Stimulus type was a between-subjects independent variable, and irrelevant sound placement and irrelevant sound type were within-subjects variables. Each of the conditions was blocked, and the order of blocks was counterbalanced. Each participant completed two practice and 16 experimental trials in each of the five blocks, making a total of 90 trials. The dependent variable was the size of the irrelevant sound effect as measured by recognition accuracy in the each of the distraction conditions subtracted from recognition accuracy in the silent, control condition.

Materials

Tone memory task. Visual tone sequences consisted of four tones chosen from all possible combinations of the nine pitches in the C major scale (from C4 to D5). Sequences were generated using a true random number generator on the Internet (www.random.org), but with the following constraints: there were no successive repeated tones or melodic intervals greater than an octave, an equal number of sequences in each block began on the same pitch, and no sequence started and ended on the same tone. Once generated, sequences were presented as stem-less quarter notes on a treble-clef staff, with all four tones visible at once.

Auditory comparison sequences were different on half the trials. Each “different” trial was created by altering one of the four tones. Half the alterations in each block took place on the second note, and half took place on the third note. Half ascended in pitch by two whole steps (or five half steps if the original tone was B or E), and half descended in pitch by two whole steps (half if the original tone was C or F). These sequences were entered into a music notation program, Sibelius (Finn & Finn, 1993). The sequences were then exported to another music notation program, MuseScore (Schweer, 2009) to be converted into digital audio files. MuseScore was used because of its cleaner audio export function, in which notes do not overlap with each other. Audio files were four seconds each, with each tone lasting 600 ms with a 400 ms silence before each subsequent tone onset.

Letter memory task. Visual letter sequences consisted of seven letters chosen from nine phonologically dissimilar consonants in the English alphabet (B, F, H, J, K, L, M, Q, and R). Seven letters were chosen because a pilot experiment by Williamson et al. (2010) indicated that participants exhibited levels of recall for seven letters that were comparable to the levels observed for four tones. The same constraints used for the tone sequences were applied to the letter sequences. On "different" trials, the alterations were distributed across the middle five letters and consisted of a change of two alphabetical steps (given the restricted set of letters), either ascending or descending (e.g., a B became an H). To match the tone sequence presentation, the letters were presented together as a list. Letters were in 64-point font in the center of the screen. A woman recorded the letters of each same and different sequence in Audacity.

Irrelevant materials. The irrelevant sound sequences were always seven seconds long. Irrelevant sounds started either one second after the onset of the visual sequence or immediately after the visual sequence disappeared, lasting until the presentation of the auditory comparison sequence. Seven sequences were randomly generated from a pool of nine items. The irrelevant speech item pool consisted of three individuals each speaking the digits “one", "two", and "three”. The pitch register of each speaker also varied: a man spoke in low register, a woman spoke in mid-register, and a woman spoke in high register. In each sequence, there were no immediate repeats of either number or speaker.

The irrelevant tone item pool consisted of three instruments, each of which played one pitch chroma in the range of C3-B5 across three octaves. For example, one sequence might comprise C3, C4, C5, C4, C3, C4, C5. The instruments were the organ, guitar, and clarinet. These sound files were created in MuseScore and exported as digital audio files. On each trial, the pitch chroma that appeared in the irrelevant tone sequence was one that did not appear at any point during the visual or auditory comparison sequences. Like the irrelevant speech sequences, there were no immediate repeats of either octave or instrument. The manipulation of timbres was meant to match the manipulation of speaker identity in the irrelevant speech sequences, and the manipulation of octaves matched the manipulation of pitch height in the different speaker's voices.

Procedure

Participants provided written informed consent by signing a form. After providing consent, participants were directed to a computer program which first displayed a set of instructions for the task. All stimuli were presented on a Mac running OSX 10.9 and equipped with PsyScope (Cohen, MacWhinney, Flatt, & Provost, 1993). Each participant was randomly assigned into a tone or letter condition, with 20 participants in each group. Participants were instructed to ignore any sounds they heard and concentrate on remembering the sequence that they saw. A visual cue (+) appeared on screen for 2 s. For participants in the tone condition, the first tone of the visual sequence was played over headphones at the same time as the cue over headphones to orient the participant to the correct absolute starting pitch level. After the cue disappeared and 2 s of silence, the visual sequence (of either tones or letters) was displayed for 8 s. After the visual sequence disappeared, the screen was blank for 10 s. Another visual cue then appeared on screen for 1 s before the comparison sequence played. At the end of the auditory sequence, participants indicated by a key press whether they thought the auditory sequence was the same (s) or different (d) as the visual sequence. A diagram of the second-by-second presentation of material is presented in Figure 1.

During the simultaneous irrelevant sound condition, irrelevant sound sequences began one second after the onset of the to-be-remembered visual sequence and lasted for seven seconds until the visual sequence disappeared. The one second gap was included to allow the participants to orient themselves to the correct pitch level after seeing the sequence before they heard any irrelevant sounds. During the delayed irrelevant sound condition, the irrelevant sound sequences began when the visual sequence disappeared and lasted for seven seconds throughout the retention interval. There were then three seconds of silence before the visual cue appeared to signal the presentation of the auditory sequence. The entire experiment lasted about an hour, including short breaks between blocks and debriefing at the end.

Results

Each participant's memory scores were calculated using the guess correction formula: *Hits – False Alarms*. A response was considered a hit when the comparison auditory stimulus was the same as the visual stimulus and the participant correctly identified it as "same." A response was considered a false alarm when the comparison stimulus was in fact different from the visual stimulus, but the participant incorrectly identified it as "same." The average guess-corrected scores, across participants, are shown in Table 1 and Table 2. Tables 3 and 4 indicate the mean irrelevant sound effect by distraction condition for the tone stimuli and letter stimuli, respectively. Mean irrelevant sound effects were calculated by subtracting the participants' scores in each of the distraction conditions from the silence (control) condition.

A 2 (stimulus type: tones or letters) X 5 (irrelevant sound type and placement: silence, digits-delayed, digits-simultaneous, tones-delayed, tones-simultaneous) mixed ANOVA was conducted on the guess corrected scores. An alpha level of .05 was used for all statistical tests. The type of distraction did not significantly influence recall accuracy, $F(4, 152) = 0.86$, $MSE = 1.68$, $p = .49$. The type of to-be-remembered material did not significantly influence recall accuracy, $F(1, 38) = 0.01$, $MSE = 17.02$, $p = .90$. There was no significant interaction between irrelevant sound condition and stimulus type on memory scores, $F(4, 152) = 1.86$, $MSE = 3.12$, $p = .12$.

Further analyses addressed the memory scores separately by stimulus type. Table 1 displays mean scores by distraction condition for participants who recalled tones, and Table 2 displays scores for participants who recalled letters. The analysis results of the means in Table 1 are inconclusive because participants performed the worst on average in the silence condition, resulting in negative irrelevant sound effects. However, Table 2 shows a clearer pattern. Though not statistically significant, the pattern is suggested numerically in that the silence condition resulted in the best performance overall on average ($M = 5.40$, $SE = .50$), whereas the speech-delayed condition disrupted performance the most ($M = 4.40$, $SE = .46$). The speech-simultaneous condition also disrupted performance relative to silence ($M = 4.80$, $SE = .45$), though not as much as the speech-delayed condition ($M = 4.40$, $SE = .46$).

Performance in the letters condition in the silence, tones-delayed, tones-simultaneous, digits-simultaneous, and digits-delayed conditions, respectively, did not significantly correlate with instrument type (vocalist or non-vocalist), $r(18) = .04$, $p = .88$, $r(18) = -.02$, $p = .94$, $r(18) = -.38$, $p = .10$, $r(18) = -.38$, $p = .10$, and $r(18) = .14$, $p = .55$. Performance in the letters condition in the silence, tones-delayed, tones-simultaneous, digits-simultaneous, and digits-delayed conditions, respectively, did not significantly correlate with number of years of formal study, $r(18) = .32$, $p = .16$, $r(18) = .29$, $p = .21$, $r(18) = .13$, $p = .58$, $r(18) = .08$, $p = .73$, and $r(18) = -.05$, $p = .82$.

Performance in the tones condition in the silence, tones-delayed, tones-simultaneous, digits-simultaneous, and digits-delayed conditions, respectively, did not significantly correlate with instrument type (vocalist or non-vocalist), $r(18) = -.32$, $p = .17$, $r(18) = -.19$, $p = .42$, $r(18) = -.08$, $p = .74$, $r(18) = -.18$, $p = .45$, $r(18) = -.16$, $p = .50$. Performance in the tones condition in the digits-delayed condition did correlate significantly positively with number of years of formal study, $r(18) = .50$, $p = .02$.

Discussion

The hypothesis that recall for tones would be more disrupted by irrelevant tones than speech and that recall for letters would be more disrupted by irrelevant speech than tones was not supported. Considering each stimulus type separately, the letters results indicate at least somewhat of a modality-specific effect, in that irrelevant speech was more disruptive to memory on average than irrelevant tones. Interestingly, the delayed speech condition caused more disruption than the simultaneous speech condition, which suggests that the irrelevant sound effect was not occurring because of disruption to encoding, but more because of disruption to rehearsal. Because none of the relationships were statistically significant, however, it is difficult to make any conclusions about how the irrelevant sound disrupted recall in this particular situation. The lack of significance might have occurred simply because, while in Williamson et al.'s (2010) study the irrelevant sound lasted for 13 seconds throughout the duration of the trial, in the present study the irrelevant sound lasted only 7 seconds no matter where it was placed. Thus, participants had an extended period of silent time within each condition during which they were able to focus on and rehearse the stimulus, a time that was not present in Williamson et al.'s (2010) study.

The tones condition presents difficulties for analysis. Again, though not significant, the pattern of results showed that participants performed the worst on average in the silence condition (i.e., no irrelevant sound for the duration of the trial). This pattern might have occurred because the stimuli in the silence condition just happened to be more difficult than the other conditions, regardless of any irrelevant sound attached to it. I was not able to counterbalance each tone stimulus with each of the distraction conditions because that would have resulted in too many possible combinations, requiring more time and participants than were available. If participants performed the worst on average in the silence condition due to chance only, then it would be beneficial to consider what pattern the results would show if participants were to perform most accurately on the silence condition. If the average recall score on the silence condition were to be a six, then there would be a slight irrelevant sound effect for simultaneous tones and a larger effect for delayed tones, which mirrors the pattern in the letters condition. However, there would also be an irrelevant sound effect, a larger one in fact, for simultaneous irrelevant speech as well as delayed irrelevant speech. The finding of an irrelevant sound effect no matter what the distraction would not match Williamson et al's (2010) finding of a modality-specific interference effect and might suggest that tonal and verbal materials are stored in the same loop. However, only the tones condition would have exhibited a general distraction effect, while the letters condition exhibited a modality-specific distraction effect. The idea that memory for tones and memory for letters exhibit such different patterns suggests that different mechanisms are responsible for both.

Conjecture or not, the results in the tones condition are curious and suggest another issue with comparing tonal and verbal stimuli that had not been raised in the literature previously: tonality. It is difficult to truly randomize tonal stimuli in the same way that verbal stimuli can be randomized because of the sense of "key" or "home base" inherent in music that does not apply to language. Tone sequences are perceived to be in a certain key when they outline a certain chord or otherwise appear to center around a home pitch, and these sequences are easier to remember because they fit into a schematic expectation with which musicians are familiar. It may seem a simple fix to utilize only sequences that are outside of any perception of a key (e.g., sequences that use large intervals and do not appear to center around a home pitch). However, because all the stimuli in Williamson et al.'s (2010) study as well as the current study used diatonic pitches (also known as the key of C), debriefing feedback indicated that many participants began to hear each trial as being in C even when the sequence itself did not necessarily suggest that key. For example, a sequence with the notes D, B, G, and A with large intervallic skips in between notes might be difficult without any context, but to a participant who was already expecting each sequence to be in C, it would be easier to place those notes within the context of that key and thus easier to remember the sequence despite distracting noise. This phenomenon of key perception is problematic because of the lack of a comparable phenomenon in language and could at least be partially remedied by using, firstly, a more varied set of tones that are not merely diatonic, and secondly, a tone mask between each trial. A tone mask is a random sequence of tones that would scramble the participant's perception of any tonality such that the participant would not already be primed to hear the next sequence in a certain key, regardless of the actual tones.

Other than the main goal to replicate a modality-specific interference effect with this design, there was also a goal to consider a larger range of musical expertise than in Williamson et al.'s (2010) study. There were no significant correlations between memory scores in the letters condition and years of formal study. However, it is interesting to note that all of the correlations between the memory scores in the tone condition and years of formal study were either almost significant or significant (as in the digits-delayed condition). The correlations would tentatively suggest that there is a relationship between more musical experience and greater ability to perform notational audiation (hearing what written music would sound like in the inner ear). However, the numbers remain inconclusive and would require a larger sample and an improved measure of experience other than self-report, because some participants have different ideas of what "formal study" entails. Researchers have found that the ability to perform notational audiation varies widely in musicians, even across levels of skill (Brodsky et al., 2003, 2008). Considering the type of instrument played could factor into notational audiation ability because some musicians need to know how written music would sound in order to play or sing (e.g., vocal, trumpet, french horn), whereas others largely only need to know fingerings (e.g., piano, flute). However, the correlations in the current study indicate that both vocalists and non-vocalists performed similarly in each of the conditions. Analysis in the current study was limited to vocalists versus non-vocalists because those groups yielded the largest samples in which to find patterns, whereas other groups would have only had a few participants.

The lack of a significant stimulus type effect suggests that the letters and tones were similar in their level of difficulty for participants to remember, on average. This result differs from Williamson et al.'s (2010) finding that participants' recall was significantly lower for tones than for letters across each of the irrelevant sound conditions. The similar level of difficulty across stimulus type in the current study is encouraging and indicates that the controls on the tone stimulus generation were working as intended. Effective controls included "different" trials containing a note either two whole steps above or below the original, instead of just a half step, as well as the presence of an orientation pitch at the beginning of each trial so that participants heard the absolute pitch level of the first tone. Another strength of this study was the use of a wider variety of participants than has been used in past studies; participants' ages ranged from 19 to 60 and were sampled from the greater Walla Walla community in addition to Whitman students, staff, and faculty. Formal study experience also ranged from just one year to 25 years. Despite this diversity, scores across participants with varying levels of education, ages, and music experience were generally homogenous. This finding allows me to be more confident in generalizing results, rather than making the possibly erroneous assumption that Whitman College student musicians represent musicians in general (Sue, 1999).

This research would benefit from several directions of future study. Piloting sets of both tonal and verbal stimuli in the absence of any distraction would help researchers to be sure that differences are occurring solely because of the distraction conditions themselves. Currently, there are only a few studies that compare short-term memory for verbal and tonal materials using parallel tests (Schendel & Palmer, 2007; Williamson et al., 2010), and being able to directly compare the two is a valuable area of study. Reworking the tone stimuli in future research will help to shed more light on the question of whether or not musical information comprises a separate store in memory.

Another concern is that it is difficult to determine what strategies participants used to encode the tone stimuli, even with the use of the visual-auditory recognition method. For example, debriefing data indicated that some participants completed each trial by verbally encoding the names of the notes and then comparing what they knew those particular intervallic relations should have sounded like with the auditory comparison stimulus at the end of the trial. Other participants indicated that they automatically encoded the material in terms of tactile movement; in other words, instrumentalists could "feel" what the particular sequence of notes would be if they were to play it and could compare that sensation with what they heard in the auditory stimulus. An idea for a future study is to retain the design of the current experiment but to include conditions in which participants are specifically asked to encode the stimuli in different ways and determine whether the strategies produce in significant differences as far as the patterns of disruption to memory.

As a final note, it would be valuable to include conditions with stimuli that are more complex, as previous studies have sometimes utilized (e.g., Salamé & Baddeley, 1989). More complex stimuli would include greater harmonic variety, density of texture, and a greater range of pitches that would be present in real-life music and not merely tones under laboratory conditions. Using stimuli with this level of complexity would allow researchers to consider how the knowledge about musical storage might be applied to musicians' daily lives in terms of how they practice, learn, and perform music.

References

Baddeley, A. D. (1990). *Human memory: Theory and practice.* Boston, MA: Allyn and Bacon.

Baddeley, A. D. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology, 63,* 1-29.

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

Berz, W. L. (1995). Working memory in music: A theoretical model. *Music Perception, 12,* 353-364.

Brodsky, W., Henik, A., Rubinstein, B. S., & Zorman, M. (2003). Auditory imagery from musical notation in expert musicians. *Perception and Psychophysics, 65,* 602-612.

Brodsky, W., Kessler, Y., Rubinstein, B. S., Ginsborg, J., & Henik, A. (2008). The mental representation of music notation: Notational audiation. *Journal of Experimental Psychology: Human Perception and Performance, 34,* 427-445.

Bull, A. R., & Cuddy, L. L. (1972). Recognition memory for pitch of fixed and roving stimulus tones. *Perception and Psychophysics, 11,* 105-109.

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers, 25,* 257-271.

Colle, H. A., & Welsh, A. (1976). Acoustic masking in primary memory. *Journal of Verbal Learning and Verbal Behavior, 15,* 17-32.

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

Elliott, L. L. (1970). Pitch memory for short tones. *Perception and Psychophysics, 8,* 379-384.

Finn, B., & Finn, J. (1993). Sibelius (Version 7.1.3) [Computer software]. Avid Technology: Burlington, MA.

Jones, D. M., & Macken, W. J. (1993). Irrelevant tones produce an irrelevant speech effect: Implications for phonological coding in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 369-381.

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

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

Norris, D., Baddeley, A. D., & Page, M. P. A. (2004). Retroactive effects of irrelevant speech on serial recall from short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition 30,* 1093-1105.

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

Salamé, P., & Baddeley, A. D. (1982). Disruption of short-term memory by unattended speech: Implications for the structure of working memory. *Journal of Verbal Learning and Verbal Behavior, 21,* 150-164.

Salamé, P., & Baddeley, A. D. (1989). Effects of background music on phonological short-term memory. *Quarterly Journal of Experimental Psychology, 41,* 107-122.

Schendel, Z. A., & Palmer, C. (2007). Suppression effects on musical and verbal memory.

Memory and Cognition, 35(4), 640-650.

Schweer, W. (2008). MuseScore (Version1.3) [Computer software]. Retrieved February 28, 2013. Available from http://www.musescore.org.

Sue, S. (1999). Science, ethnicity, and bias: Where have we gone wrong? *American Psychologist, 54,* 1070-1077.

Wallin, M. L., Merker, B., & Brown, S. (2000). *The origins of music.* Cambridge: MIT Press.

Wickelgren, W. A. (1966). Consolidation and retroactive interference in short-term recognition memory for pitch. *Journal of Experimental Psychology, 72,* 250-259.

Wickelgren, W. A. (1969). Associative strength theory of recognition memory for pitch. *Journal of Mathematical Psychology, 6,* 13-61.

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

Table 1

Mean Recall Accuracy Scores By Distraction Condition for Tones Stimuli

Distraction condition	Mean	SD
Silence	4.60	0.50
Tones-simultaneous	5.45	0.47
Tones-delayed	4.95	0.56
Digits-simultaneous	5.00	0.52
Digits-delayed	5.05	0.49

Table 2

Mean Recall Accuracy Scores By Distraction Condition for Letters Stimuli

Distraction condition	Mean	SD
Silence	5.40	0.50
Tones-simultaneous	5.05	0.54
Tones-delayed	5.05	0.43
Digits-simultaneous	4.80	0.45
Digits-delayed	4.40	0.46

Table 3

Mean Irrelevant Sound Effect By Distraction Condition for Tones Stimuli

Distraction condition	Mean	SD
Tones-simultaneous	-0.84	1.75
Tones-delayed	-0.35	1.69
Digits-simultaneous	-0.40	1.79
Digits-delayed	-0.45	1.82

Table 4

Mean Irrelevant Sound Effect By Distraction Condition for Letters Stimuli

Distraction condition	Mean	SD
Tones-simultaneous	0.35	2.56
Tones-delayed	0.35	1.27
Digits-simultaneous	0.60	1.81
Digits-delayed	1.00	1.55

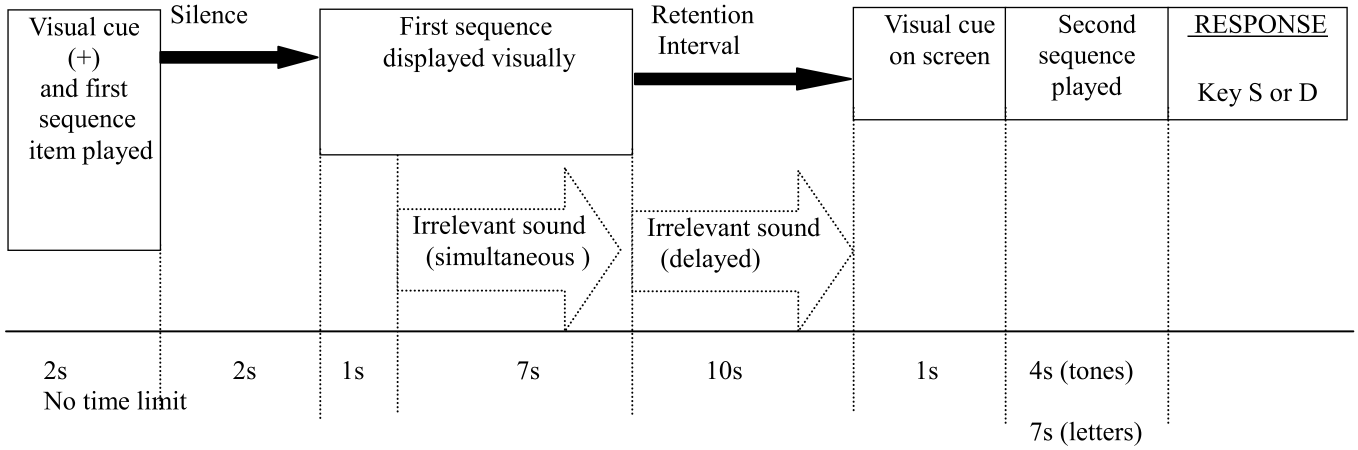

Figure 1. Diagram of presentation of materials during the experimental procedure.