Spring 2015

The Influence of Preferred Music on Vigilance and Mental Workload

Michelle H. Wang
Otterbein University, haonan.wang@otterbein.edu

The influence of preferred music on vigilance and mental workload

Michelle H. Wang

Psychology Department, Otterbein University, Westerville, OH, USA

The influence of preferred music on vigilance and mental workload

The benefit of music has been widely studied. The most well-known influence of music is the emotional impact causes by a favored piece. It is not surprising that music has a close relationship with emotions. Music can trigger a variety of emotions such as joy, fear and sadness (Sloboda & O'Neill, 2001). Emotional response of music is often measured by the number of "chills" or by examining the sensitivity of skin conductance (Panksepp, 1995; Sloboda, 1991; Rickard, 2004). "Chills" refer to self-reported strong emotional response to music, whereas the sensitivity of skin conductance can be measure by an instrument. Previous studies found that when participants listen to self-selected emotionally powerful music, it triggers higher level of arousal, or more intense emotional response, as compared to other types of music and emotionally powerful film scene (Rickard, 2004). Also, strong emotional responses to music are of the same nature as emotions produced by non-aesthetic, or real-world, stimuli (Krumhansl, 1997; Rickard, 2004). In fact, emotional impact and regulation was reported as the main reason for individuals to listen to music (Chanda & Levitin, 2013).

Other than the emotional effect of music, researchers are finding that music can improve performance both physically and cognitively. Music has been used to improve performance through many types of exercise, such as, maximal (reaches maximum heart rate), submaximal (reaches 85% of the maximum heart rate) and pre-competition preparations (Waterhouse, Hudson, & Edwards, 2010). One experiment exploring the influence of music on submaximal cycling found significant differences in speed between faster tempo and normal and slower tempo groups (Waterhouse, et al., 2010). Waterhouse and his colleagues suggested that participants not only worked harder, they were more motivated and enjoyed the music better at a faster tempo. In addition to the tempo aspect of music, research showed that the mere exposure to music can improve physical performance. A study of muscular endurance reported that participants held weights significantly longer when music was present as compared to white noise (Crust, 2004). The motivational quality of music is also apparent during physical exercise. While performances outcome may remain unchanged, individuals often report less stress and more enjoyable exercising experience when music is present (Wininger & Pargman, 2003; Crust, 2004; Tenenbaum, et al., 2004).

Inquiries of the facilitative effect of music are not limited to physical exercises. For decades, researches have explored the influence of music on learning and performance. Rauscher, Ky and Shaw (1993) tested the influence of music on spatial cognition. They found that the participants scored higher in the spatial reasoning category of the Stanford-Binet IQ test after listening to a piano sonata by Mozart. This was dubbed the “Mozart Effect”. Since its original publication, nearly 40 studies attempted to replicate the results over the next 16 years and most could not (Pietschnig, Voracek, & Formann, 2010). Criticisms leveled against these findings included small effect size, inadequate controls and publication bias (a bias towards reporting significant results (Chabris, 1999; Pietschnig, et al., 2010). Despite the equivocal findings, the role of music in cognition remains an active interest in the field.

One of the theories emerged from the Mozart Effect study is the arousal theory. Chabris’(1999) suggests that any effect of music observed in the Mozart Effect was induced by arousal and enjoyment of the music. A later study attempting to replicate the Mozart Effect recorded enjoyment, arousal and mood of the participants (Thompson, Schellenberg, & Gabriela, 2001). Participants were assigned to one of three groups, the sonata used in the original Mozart Effect experiment (fast tempo), another sonata with slow tempo and silence. Participants who listened to the Mozart sonata had higher scores on the test of spatial abilities as well as positive mood and arousal assessments. The findings suggest that music can affects cognitive abilities by modulating arousal and mood. Music's ability to regulate arousal has since been widely studied. A larger portion of the field of music therapy is dedication to stress and anxiety management (Dileo & Bradt, 2007). A long line of studies have shown that relaxing music significantly reduce stress and anxiety in patients undergoing medical procedures (Pittman & Kridli, 2011; Dileo, et al., 2007; Bradt & Dileo, 2009) as well as healthy participants (Dileo, et al., 2007; Koelsch & Stegemann, 2012).

One aspect of cognition, attention, has been frequently studied alongside music. Research has shown that background music is able to improve attention and concentration. Shih, Huang and Chiang (2012) examined background music in the work place and found that music does in fact have an impact on performance. Two groups of participants listened to background music while completing attention tests. Both groups heard the same songs, however, the lyrics were removed from the experimental group. Music with lyrics seemed to have a larger, negative impact on concentration and attention as compared to music with no lyrics. Even though music with no lyrics has a less significant effect, participants reported higher satisfaction of their work environment in this condition. This finding indicates that music without lyrics can enhance attention performances both directly and indirectly (by increasing workers' satisfaction of the environment). Similar study in visual neglect also found that preferred music enhanced patient's performance on a perceptual report test by improving patients' attention and vigilance (Soto et al., 2009). Patients showed enhance visual awareness when completing the task with preferred music as compared to non-preferred music and silence. Soto and colleagues suggested that music can decrease visual neglect by increasing attention resource. Not only music, background noise in general can affect attention, especially sustained attention. Varied noise has been shown to improve sustained attention performance when it is played softly in the background to a low demanding task (Lysaght, 1982).

The preference of music was the focus on the present investigation. Previous research (Johansson, Holmqyist, Mossberg, & Lindgren, 2012) examined the relationship between preferred/non-preferred music and reading comprehension. Johansson and colleagues allowed their participants to provide their own preferred music. They concluded that participants scored significantly worse on a reading comprehension test when they listened to non-preferred music. Importantly, they observed no significant differences between silence and preferred music. Johansson and colleagues advanced this area of research by allowing participants to listen to music that was truly preferred. Previous research (Daoussis & McKelvie, 1986) had participants provide a general preference of genre but the researchers provided the music. The present investigation adopted the same method as Johansson et al. by asking participants to provide their own music.

In addition to the impact music may have on performance, music may also effect workload. Workload refers to the relationship between the amount of labor and the task demand. It is an assessment of the psychological cost of a task. Laurie-Rose, Frey, Ennis and Zamary (2014) defined workload as a combination of task demand and the individual's inherent characteristics, mood, personality, capabilities and motivations. It is a multidimensional measurement – there are many different aspects that contribute to workload. Workload is often studied in the fields of human factors (Laurie-Rose, et al., 2014). Traditionally, workload is often studied under military settings due to its important role in decision-making process (Stone, 2008). Aviation psychology concerns itself with vital issues regarding safety decisions in complex environments with people as well as technology. In this area, workload of the pilot is of utmost importance (Helmreich, 2000). Perhaps due to similar work environment, workload of emergency room doctors also gained interests in the field in recent years (Levin et al., 2006). Outside of human factors, educational psychologists are also interested in the mental workload, especially in researches of multimedia learning environment and design (Wiebe, Roberts, & Behrend, 2010). A gold standard scale used to measure workload is the NASA Task Load Index (NASA-TLX) (Hart & Staveland, 1998). There are six subscales in the NASA-TLX, Mental Demand, Temporal Demand, Physical Demand, Frustration, Effort, and Performance; each measures a related yet distinct area of mental workload.

The Present Study

The current experiment examines performance and workload in a sustained attention task under quiet and preferred music conditions. Sustained attention refers to the ability to maintain focus on attention and to remain alert to specific stimuli over prolonged, unbroken periods of time (Warm, 1984). Sustained attention, or vigilance, can be related to many aspects of musical influence. Emotional impact, arousal and stress regulation, cognitive performance enhancement and motivation could all contribute to the vigilance outcome. Vigilance task functions as the experimental way to study sustained attention. Most, but not all, vigilance tasks display a gradual decline in performance, or the vigilance decrement (Davies & Parasuraman, 1982). In most cases, participants must detect specific changes in a stream of stimuli that signify the target (Warm, 1984).

The vigilance task in the current experiment consisted of a stream of digits in pairs, with the target be the pairs with the sum of nine. I predicted that the presence of preferred music will lower participant's workload as compared to silence. As a result, vigilance performance will also improve in terms of more correct detection, fewer false alarm and faster response latency.

Method

Participant

A total of 47 students (17 male, 30 female) ranged in age between 18 and 45 (*M* = 20.68, *SD* = 4.83) participated in the experiment. All students were recruited from the psychology department participant pool at Otterbein University. Research participation course credits were awarded in exchange for participation. The experiments were carried out in accordance with the guidelines of Otterbein University Institutional Review Board, and all participants gave informed written consent.

Task and Material

Vigilance Task. The experiment was programmed and executed using *Superlab* on 13" Macbook Pro computers. The vigilance task consisted of pairs of digits, 1-8 (Geneva font, size 26, color black, 50% transparency) separated by a space. The location of each stimulus pair was chosen randomly from nine different locations near the center of the screen. All nine locations were within a 4cm x 4cm square with the center of the screen as its center. The locations were of equal distance to the center and they were labeled as top left, top middle, top right, center left, center, center right, bottom left, bottom middle and bottom right (Figure 1). Eight designated pairs with the sum of 9 (1 8, 2 7, 3 6, 4 5, 5 4, 6 3, 7 2, 8 1) were chosen as the targets, whereas all other pairs of digits were distractors. There were 52 distinct pairs of distractors. Each stimulus was randomly chosen to be presented for 100 ms, the signal probability was 0.15 and the event rate is 45/minute. Each period comprised 360 stimuli, of which 54 were targets. Participants were instructed to press the spacebar when they detected the target. There were three consecutive 8-minute periods with a total of task duration of 24 minutes for one task.

NASA Task Load Index (NASA-TLX). A computer adapted version of the NASA-TLX (Hart & Staveland, 1988) was administered after each of the vigilance tasks. The NASA-TLX is a subjective measurement originally developed for aviation research. I selected the NASA-TLX as the workload assessment for the current study because it is considered a reliable and valid measurement (Moroney, Biers, & Eggemeier, 1995; Wierwille & Eggemeier, 1993), its multidimensional nature (Hart & Staveland, 1988) and its widespread utility. The NASA-TLX assesses six distinct and interconnected dimensions of workload, Mental Demand, Temporal Demand, Physical Demand, Performance, Effort and Frustration. Of the six dimensions, Mental, Temporal and Physical Demand focus on the participant's evaluation of the task; whereas Performance, Effort and Frustration focus the participant's evaluation of his/her response to the task (Warm et al., 1996).

Procedure

Participants completed the vigilance tasks in a laboratory room with up to five other participants. Each participant was seated separately in individual carrels. Participants were instructed to bring their own music playing device, headphones as well as a 30-minute playlist of music with lyrics. Participants were also instructed that they must be familiar with songs included in the playlist and the songs contain lyrics. The general genre of each participant's playlist was recorded.

Upon agreeing to participate in the study, participants were instructed to complete a brief practice vigilance task. Once the participant became familiar with the nature of the task, he/she continued to complete the vigilance tasks both with music and in quiet. The Order in which participants proceeded through the vigilance task was balanced. Following each vigilance task, participants rated their workload using the NASA-TLX. Participants were asked to rate their experience in all six subscales on a scale from 1 to 100 using multiples of 5. Participants were allowed an approximately 5-minutes break (self-paced) between vigilance tasks.

Results

Initial inspection of box plots revealed that two participants showed extreme scores and were consequently removed from all analyses. One of these exhibited extreme median reaction times (RT) and percentage hits scores whereas the other exhibited extreme percentage FA scores. I examined the following dependent measures for the vigilance task: percentage hits, percentage false alarms (FA's), A', B'', median reaction time (RT), and reaction time variability (RT^{SD}). Percentage hits are defined as percentage of targets of which the participants correctly detected. Similarly, percentage FA's reflect the percentage of distractors that the participants falsely reported as targets. A' is a non-parametric measure of perceptual sensitivity devised by Pollack and Norman (1964). A' is calculated using percentage hits and percentage FA's. A typical value of A' ranges from .5 (targets cannot be distinguished from distractors) to 1 (targets are always correctly detected) (Stanislaw & Todorov, 1999). However, the minimum possible value of A' can be 0. Any scores between 0 and 0.5 may indicate sampling error or response confusion (Stanislaw et al., 1999). B" is also calculated using percentage hits and percentage FA's. B" is a nonparametric measure of response bias. Values of B" ranges from -1 (liberal criterion, i.e., more likely to have FA's) to 1 (conservative criterion, i.e. more likely to miss the target). B" value of 0 indicates no response bias (Stanislaw et al., 1999). The median RT and RT^{SD} of correct trials, or hits, were calculated (in seconds) for every period of watch across all participants. RT^{SD} reflects the standard deviation of RT, which provides stronger effect size than the mean RT and is sensitive to group differences (Klein, Wendling, Huettner, Ruder, and Peper, 2006).

Workload measures were obtained for each vigilance task immediately following its completion using the NASA-TLX. Participants rated the degree to which they experience in Mental Demand, Temporal Demand, Physical Demand, Performance, Effort and Frustration using the weighted assessment of the subscales. Responses were recorded on a 100-point scale in increments of 5. Of the six subscales, Physical Demand received the lowest rating, 15.69. Thus, it was removed from future analysis in order to meet the independence assumption of the analysis (Temple, Warm, Dember, Jones, LaGrange & Matthews, 2000). Descriptive statistics for all experimental conditions for both the vigilance and workload data are presented in Tables 1 and 2. Although I collected data regarding music genre from each participants the diversity between and within playlists precluded any formal analysis.

Vigilance Data

The data for each dependent measure was submitted to a 2 (Order) x 2 (Music) x 3 (Periods of Watch) split plot ANOVA with repeated measures on both Music and Periods of Watch.

Percentage Hits. The main effects of Order, Music and Periods Watch did not reach significance, Wilk's lambda = .94, F (1, 43) = 3.79, $p > .05$, F (1, 43) = 2.71, $p > .05$, F (2, 42) = 1.18, $p > .05$, respectively. However, we observed a significant Order x Music interaction (F (1, 43) = 15.79, $p < .001$, $\eta_p^2 = .27$). Figure 2 displays the changes in percentage hits in both of the quiet and the music conditions. As can be seen in Figure 2, participants correctly detected more targets in the condition that was administered second. However, it is apparent that when the quiet condition is presented first, participants correctly detected fewer targets relative to all other conditions.

Percent False Alarms. We observed a significant Order x Music x Periods of Watch interaction (F (2, 42) = 12.25, p < .001, η_p^2 = .37). The three-way interaction is plotted in Figure 3. As seen in Figure 3, participants committed fewer false alarms in the second vigil regardless of the condition presented. However, when the music condition was presented first, participants experienced a greater decline in false alarms between Periods 1 and 2 and percentage of false alarms eventually decreased to the same level of the quiet vigil performed second. Whereas when the quiet vigil is administered first, percentage false alarms decreased less dramatically and never attained the same level of performance as the second vigil.

Median RT for Hits. The main effects for Order, Music and Periods of Watch did not reach significance, F (1, 43) = 2.476, p > .05, F (1, 43) = .880, p > .05, F (2, 42) = 2.153, p > .05), respectively. However, we observed a significant Order x Music x Periods of Watch interaction (F (2, 42) = 5.465, p = 0.008, η_p^2 = .206). The three way interaction is plotted in Figure 4. As observed, median RT stayed stable for both music vigils across all periods whether it was administered first or second. However, the median RT's of the quiet vigils differ greatly depending on the order in which they were administered. When the quiet vigil is presented first, participants exhibit the highest median RT (worst performance) across all conditions. However, when the quiet vigil is presented second, median RT is the lowest (best performance).

Variability of Reaction Time for Hits (RT^{SD}). We observed a moderate effect of Order (F (1, 43) = 4.37, p = .04, η_p^2 = .09). The mean RT^{SD} was significantly higher when the quiet is administered before the music condition. However, the main effect of Music and Periods of

Watch did not reach significance, F (1, 43) = .02, p > .05, F (2, 42) = .89, p > .05 respectively. There are also no significant interactions.

Perceptual Sensitivity (A'). The main effects for Order, Music and Periods of Watch did not reach significance, F (1, 43) = 2.70, p > .05, F (1, 43) = 3.36, p > .05, F (2, 42) = 2.44, p > .05), respectively. However, we observed a significant Order x Music interaction (F (1, 43) = 26.19, p < 0.001, η_p^2 = .38), which is plotted in Figure 5. As can be seen in Figure 5, participants exhibited higher perceptual sensitivity in the vigil that was performed second. However, it is apparent that when the quiet vigil is presented first, participants exhibited lower sensitivity relative to all other conditions.

Response Bias (B"). We observed a significant main effect of Period of Watch, F (2, 42) = 13.67, p = .010, η_p^2 = .39. However, the main effects of Order and Music did not reach significance, F (1, 43) = 1.26, p > 0.05, F (1, 43) = .25, p>.05, respectively. We also observed a significant Order x Music interaction (F (1, 43) = 7.17, p < .001, η_p^2= .14) as well as a significant Order x Music x Periods of Watch interaction (F (2, 42) = 4.81, p = .013, η_p^2 = .19). The three-way interaction is plotted in Figure 6. We observed higher B" values in the second vigil regardless of the condition presented, which indicates that participants were more conservative in their responses. However, when the music vigil is presented first participants experienced a greater rise in B", or becoming more conservative, between periods 1 and 2. Whereas when the quiet vigil is administered first, B" increased less dramatically.

Workload

The raw scores for the five subscales scores were submitted to a 2 (Order) x 2 (Music) x 5 (Workload Subscales) split plot ANOVA with repeated measures on Music and Workload

Subscales. We observed a significant main effect of workload, F (4, 40) = 8.06, $p < .001$, η_p^2 = .45. However, the main effect of Order and Music did not reach significance, F (1, 43) = .74, $p > .05$, F (1, 43) = .39, $p > .05$, respectively. We also observed a significant Order x Music interaction (F (1, 43) = 5.94, $p = 0.019$, $\eta_p^2 = .12$) as well as a significant Order x Music x Periods of Watch interaction (F (4, 40) = 7.34, $p < 0.001$, $\eta_p^2 = .42$).

The three-way interaction is plotted in Figure 7. The left panel displays average workload reported in each of the 5 measures when the music condition follows the quiet condition. We observed that participants reported better performance and lower workload in mental demand temporal demand, effort and frustration when the music condition follows the quiet condition. However, we observed less contrast between the two conditions when the quiet condition follows the music condition (Figure 7 – right panel). When the quiet condition is presented first it has the highest average workload (69.17), whereas when the quite condition is presented second it has the lowest (62.86). Overall, the average reported workload was similar across all conditions.

Discussion

The purpose of the present study was to examine the influence of music on vigilance task performance and perceived mental workload. This project extended the extant literature regarding the effects of music on cognitive performance by using music familiar to the participants and by adding a multidimensional assessment of perceived workload in an effort to obtain a fuller understanding of performance changes that occurred. Data only partially supported the presence of a vigilance decrement. The presence of music exerted influence through numerous interactions with the variables of Order and Periods of Watch. These interactive effects are quite robust and reveal some meaningful effects.

To determine if the task employed yielded the traditional vigilance outcomes, I assessed the vigilance decrement, or changes in performance over time, by examining the variable of Periods of Watch. In vigilance tasks, a vigilance decrement is observed when there is a decline in percentage hits, a rise in median RT, an increase in response bias or conservatism, and a decline in false alarms. Statistical analyses reveal robust effects of Periods of Watch on both the decline of percentage FA and the increase in conservativism (B") with dramatic changes occurring after Period 1. Specifically, participants committed fewer FA's and showed a rise in conservatism as the session progressed.

Contrary to predictions regarding the vigilance decrement, participants exhibited an increase in percentage hits and a decline in median RT as the session progressed. That is, participant's performance efficiency increased overtime. This phenomenon can be described as a vigilance increment (Dember, Warm, Bowers & Lanzetta, 1984; Lysaght, Dember, Warm & Loeb, 1984; Warm & Jerison, 1984). A vigilance increment is defined as an increase in detection probability over the course of the vigil (See, Howe, Warm & Dember, 1995). This phenomenon often occurs in cognitive simultaneous tasks with low event rate (See, et al., 1995). The current task includes a cognitive component (adding the two digits) and shares characteristics of a simultaneous discrimination (each event contains adequate information to determine whether or not it is a target). Although the task reflects a relatively fast event rate (45 events per minute), the overall characteristic of the task bears resemblance to those that report a vigilance increment.

These opposing findings (vigilance decrement for FA and B'', and vigilance increment for hits and response latency) suggests that the decrease in false alarms may reflect that the observers needed more time in the practice trials to become proficient. The percentage of FAs significantly decreased after Period 1, which could indicate that the observers were still learning the task during the first period. Median RT of hits also decreased after Period 2. Improved performance is observed in all depended variables in the second vigil presented as compared to the first, regardless of order. Although fewer false alarms along with an increase in B" often indicate conservativism, it could also be a sign of mastery of task when combined with an increase of hits and decrease of RT. Thus one possibility for performance improvement may have been that the Period 1 served as practice. The practice they received prior to the actual task was relatively short and without feedback. The practice consisted of nine events, of which two were targets. Pilot data suggested that participants understood the instructions with ease and did not require a lot of practice. However, they were often surprised by the speed and the locations of the stimuli. The practice trials were designed to show all nine possible locations and familiarize the participants with the fast-paced presentation of the stimuli. The number of practice trials was relatively less as compared to recent vigilance studies. Perusal of the literature reveals that many studies required over 100 trials of training before participants were allowed to proceed onto the experimental trials (Caggiano & Parasuraman, 2004; Laurie-Rose et al., 2002; Matthews, Warm, Reinerman-Jones, Langheim, Washburn, & Tripp, 2010; Shaw, Matthews, Warm, Finomore, Silverman, & Costa, 2010), whereas others required less than 20 trials (Helton & Russell, 2011, Laurie-Rose, et al., in press). The variation in practice length may be due in part to the complexity of the task. There is no industry standard for the length of a vigilance practice. Perhaps a performance criterion should be more widely adopted in vigilance studies. For example, participants must meet a 90% performance criterion in the practice trials before proceeding to the experimental trials.

Task duration may also provide a partial explanation for the opposing findings in vigilance outcomes. As Laurie-Rose, Frey, Sibata and Zamary (in press) noted, the modal length for a vigilance task was 60 minutes, with an average length of 39 minutes. The duration of the task in the current study was 24 minutes, which is considered a relatively short vigil. Vigilance tasks with rapid event rates and successive discrimination tend to have more noticeable declines in performance (Parasuraman & Davies, 1977). Short vigils especially rely on these characteristics for a decrement. The task employed in the current experiment had an event rate of 45 events per minute, whereas the event rates of tasks used in shortened vigils research frequently approach 60 events per minute (Neuchterlein, Parasuraman, & Jiang, 1983, Temple, Warm, Dember, Jones, LaGrange & Matthews, 2000). Also, as mentioned before, the current task has more characteristics of a simultaneous discrimination, not successive. Although recent studies using shortened vigils reported a significant vigilance decrement (Shaw et al., 2010) with a simultaneous task, the majority of the literature still favors the use of successive tasks for a more pronounced vigilance decrement. The current task does not fit the profile of a task to elicit a decline in performance, and this at the very least, may explain the lack of a vigilance decrement.

Similar to the vigilance outcome, the workload profile of the current experiment shares similarities and differences from that of typical vigilance experiments. I observed that the workload profile of the vigilance task in the current experiment differs from the typical workload signature of a vigilance task, of which Mental Demand and Frustration bear the highest ratings (Warm et al., 1996). Instead, participants reported highest ratings for Temporal Demand followed by Mental Demand and Effort. High rating of Mental Demand, Temporal Demand, Effort and Frustration indicate more workload. Performance ratings were low, which is expected of high workload. Surprisingly, frustration ratings were also quite low. The workload profiles for the two orders also differed slightly. Frustration levels decreased significant when the music vigil followed the quiet vigil. However, there was virtually no change in Frustration ratings when the music vigil preceded the quiet vigil.

No main effects for the presence or absence of music were observed, however, music interacted with both order and periods of watch. Significant Music x Order interactions for correct detections and perceptual sensitivity revealed stable performance in the music vigils across orders. I observed great discrepancies, however, between the quiet conditions for both of these dependent measures. Performance is poorest when the quiet condition precedes the music condition, whereas performance is best when the quiet condition follows music. In addition to these first-order interactions, I also observed robust interactions between Order, Music and Periods of Watch for median RT of hits, response criterion, and percentage false alarm. In line with the first-order interactions, the three-way interaction for median RT reveals that performance in the quiet condition is not only dependent on the order it was administered, but also the amount of time that had elapsed in the vigil. When the quiet vigil precedes music, median RT remains stable in the first two Periods of Watch and drops after Period 2. However, when the quiet vigil follows music, median RT decreased after Period 1 and increased after Period 2. That is, participants reacted most quickly in Period 2 in the quiet vigil that followed music. Meanwhile, performance in the music conditions remained relatively stable regardless of whether observers performed it first or second in the session. However, the drastic difference between the quiet conditions can be only explained by the presence of music. That is, by listening to music first, participants responded more accurately, more quickly and with better perceptual sensitivity in the subsequent quiet vigil. Workload data corroborated these performance findings. When the quiet vigil was presented first workload is higher in every subscale than when it is presented second. This may suggest that music lowers one's workload, and therefore, participants perform best in the quiet condition after listening to music of their own choice.

I also observed a Music x Order x Periods of Watch interaction for percentage FA and response criterion B" that reveals a similar pattern as the 3-way interaction of median RT. Significantly larger percentage FA and more liberal criteria are observed in Period 1 as compared to the rest of the Periods of Watch. Immediately after Period 1, percentage FA decline considerably and response criteria became drastically more conservative and both variables remain stable for the remainder of the task. This phenomenon was observed in both orders. Regardless of which condition was administered first, Period 1 always displayed the highest percentage FA. However, when the music vigil was presented first, percentage FA and response bias reached the same level as the quiet vigil that followed by Period 3. When the music vigil was presented second, percentage false alarms and response bias remained different between the two vigils. That is, the scores of the quiet vigil never reached the performance level of the music vigil. If participants' performances were gradually increasing, music seems to speed up the process of improvement. That is, when the music vigil is presented first, participants reach performance level of the second vigil (better performance) during the first vigil, whereas the first-administered quiet vigil never reached the same level as the second vigil.

Workload seems to play an important role in both the influence of music and the vigilance outcome of the current experiment and sheds light on these complex performance interactions. The condition presented first always elicited higher workload ratings as compared to the condition presented second (exception for Performance ratings, which were always lower).

However, there exist greater differences in ratings between the vigils when the quiet vigil was presented first. When the music vigil precedes the quiet vigil, there were virtually no changes in the Mental Demand and Frustration ratings. Also, there are fewer differences between the music conditions as compared to the quiet conditions. All of the workload findings correspond with the results of the main effects and interactions mentioned above. The changes in workload are reflected in the participants' performance—the lowest workload ratings (and high Performance rating) and highest hits/A' occurs when quiet follows music, whereas the highest workload occur along with lowest hits/A' when quiet precedes music. Percentage FA was higher and response bias was more liberal in the first vigil presented regardless of the condition. Workload data also displays higher workload in the first vigil presented regardless of order. Overall, lower workload tends to yield more proficient performance.

There exists a long tradition within the human factors literature of examining the effects of extraneous noise on performance. In such studies, mixed results of increment, decrement and no changes in vigilance efficiency were found (Warm, 1984). To structure the findings, noise effect on vigilance performance was divided into factors of noise level (high or low), noise quality (white or varied) and processing demand (high or low) (Lysaght, 1982). In general, improved performance efficiency was found in tasks with low processing demand, low noise level (under 90 dB) and varied noise quality. A decrement is found in tasks with high processing demand, high noise level and while noise. Tasks with white noise and low processing demand showed no change in performance regardless of noise level. The vigilance task used in current investigation would be considered to have high processing demand and varied "noise" quality. However, the noise level is difficult to determine because participants were free to adjust the volume.

An alternate line of research concerns the facilitative effect of music on learning and performance. Previous studies experimented with listening to music prior to completing a task, and the findings are equivocal. Rauscher, Shaw and Ky (1993) revealed that subjects scored significantly higher on a spatial IQ reasoning test after listening to Mozart's *Sonata for Two Pianos* as compared to relaxation tape and silence. The study was largely discredited; however, later meta-analyses (Chabris, 1999; Pietsching, 2010) revealed a small, positive, significant effect of listening to Mozart prior to completion of the spatial task. Rauscher and his colleagues attributed this finding to music's ability to enhance spatial abilities. The present findings, however, provide an alternative explanation—music reduced workload and thus performance in the task following music was enhanced.

The effect of music could vary because of its genre or certain aspects of the music itself (i.e., tempo, volume, etc.). However, in the current investigation we only focused on the preference of music. It is important for the participants to listen to their own music because I wanted to mimic a realistic study or work environment in which participants control what they hear. Participants were instructed to bring in music they would choose to listen to while studying. However, not all participants regularly listen to music while studying. Previous research on reading comprehension, which allowed self-provided music, found that participants scored significantly worse on a reading comprehension test when they listen to non-preferred music (Johansson, et al., 2012). They suspect that non-preferred (and probably unfamiliar) music was distracting and participants' unawareness of the distraction led to poorer performance. Perham and Withey (2012) used similar methods in studying spatial rotation performance and allowed participants to bring in their own "liked" music to maximize the chance that "liked" music was truly preferred. They reported increased performance in the "liked music" condition as compared to listening to music one dislikes. By listening to music they know and prefer from their own device, I anticipated that such conditions would provide data more pertinent to the original question of the current investigation.

A previous study on music preference provided preferred music based on genre preference to exert some control over the task and to ensure a uniformed experience for all participants (Daoussis & McKelvie, 1986). However, the results reported differed from studies using self-provided music. Daoussis and McKelvie (1986) reported negative effect of preferred music on reading comprehension. The results showed significantly better performance in the no music condition as compared to preferred music condition in introverts on a reading comprehension task. The studies mentioned above (Johansson, et al., 2012; Perham et al., 2012), which allowed participants to bring music from their own collection, reported positive effect of preferred music and negative effect of non-preferred music. Even though I did not observe main effect music in the current study, there are also several significant interactions of music with Order and Periods of Watch that showed potential facilitative effect of preferred music. Clearly, there is a discrepancy in findings between self-provided and experimenter-provided preferred music. This opposing findings suggests that when preferred music is not chosen from participants' own collection, it functions similarly to non-preferred music. Therefore, to ensure that preferred music is truly preferred, it is important for the participants to provide the music for the experiment.

Limitations

The current study contributes to the understanding of the influence of music on vigilance, presenting compelling data on how music and order interact to affect performance and workload.

Nonetheless, some limitations remain. First, the study used silence as the control condition. As discussed previously, the literature differs over whether silence is an acceptable control. Some researchers suggested that silence reduces stimulation and arousal (Thompson, Schellenberg, & Husain, 2001) and white noises or recorded sounds in natural environment would be better alternatives. Given the goal of the experiment was to mimic a natural study environment, we elected to utilize silence. Silence is the most probable alternative to listening to music while studying under natural circumstances. Further, Johansson et al (2011) found no significant differences among preferred music, non-preferred music and café noise on participants' reading comprehension abilities. However, participants scored significantly higher in the silent condition. This finding suggests that noise has similar effect as music, whereas silence is significantly different. Therefore, silence was elected as the better choice of control.

Participants providing their own music in the current study proved both a strength and a limitation. Self-provided music guarantees familiarity and preference, however, the lack of systematic control is problematic. Certainly in this sample, the wide diversity of genre both within and between playlists raises concerns as to whether these often vastly different playlists exerted similar influence across all participants. Further effort is necessary to develop a method that allows systematic control of music while the music remains familiar and preferred. One possible method could be providing a large pool of songs. Participant would form a playlist from the existing song within the pool. However, a substantial amount of songs must be included to cover all genre of music and therefore, the content of the playlist would be just as diverse and self-provided playlists. Also, it may not achieve the same level of preference and familiarity as self-provided playlists. Another method is to issue surveys of detailed musical preference prior to the experiment and prepare playlists based on the results of the surveys. However, it is challenging to create a survey that accurately reflects one's musical interests.

Conclusions

Even though there were significant findings, the role of music regarding vigilance and workload is still uncertain. Previous literature showed that different types of music (preferred, fast tempo, classical, etc.) can improve performance in areas beyond vigilance tasks. The findings regarding order effect in the current study were surprising, but enlightening. Recent study revealed significant effect of the position of the vigilance task within a larger battery (Laurie-Rose, et al., in press), but order effect has not been explored in conjuncture with the presence or lack of music. Further research is indicated to clarify that the changes were solely due to the presence of music, and not simply an effect of learning. A second experiment is needed to test the facilitative effect of music by removing the music vigils. One group of participants will listen to preferred and familiar music for a short period of time and proceed to complete a vigilance task, while the other group will sit in silence before complete the vigilance task. Workload rating will be collected at the end of the tasks.

Reference (I am making it)

Bradt, J. and Dileo, C. (2009) Music for stress and anxiety reduction in coronary heart disease patients. *Cochrane Database Syst. Rev*. 2, CD006577

Caggiano, D. M., & Parasuraman, R. (2004). The role of memory representation in the vigilance decrement. *Psychonomic Bulletin & Review*, 11, 932-937.

Chabris, C. F., Steele, K. M., & Dalla Bella, S. (1999). Prelude or requiem for the 'Mozart effect'?. *Nature*, 400(6747), 826-828.

Chanda, M. L., & Levitin, D. J. (2013). The neurochemistry of music. *Trends In Cognitive Sciences*, (4), 179.

Crust, L. (2004). Carry-over effects of music in an isometric muscular endurance task. *Perceptual And Motor Skills*, (3), 985.

Daoussis, L., & McKelvie, S. J. (1986). Musical preferences and effects of music on a reading comprehension test for extraverts and introverts. *Perceptual & Motor Skills*, 62283-289.

Dember, W. N., Warm, J.S., Bowers, J.C., & Lanzetta, T. (1984). Intrinsic motivation and the vigilance decrement. In A. Mital (Ed.), *Trends in Ergonomics/Human Factors I* (pp. 21-26). North Holland: Elsevier.

Dileo, C. and Bradt, J. (2007) Music therapy: applications to stress management. In Principles and Practice of Stress Management (Lehrer, P.M. et al., eds), pp. 519–544, Guilford Press.

Hart, S., & Staveland, L. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock and N. Meshkati (Eds.), *Human mental workload* (pp. 139-184). Amsterdam: Elsevier.

Helmreich, R. (2000). On error management: Lessons from aviation. *British Medical Journal*, 320, 781–785.

Helton, W. S., & Russell, P. N. (2011). Working memory load and the vigilance decrement. *Experiemental Brain Research, 212,* 429-437.

Johansson, R., Holmqvist, K., Mossberg, F., & Lindgren, M. (2012). Eye movements and reading comprehension while listening to preferred and non-preferred study music. *Psychology of Music*, 40(3), 339-356.

Klein, C., Wendling, K., Huettner, P., Ruder, H., & Peper, M. (2006). Intra-subject variability in attention-deficit hyperactivity disorder. *Biological Psychiatry*, 60, 1088 –1097.

Koelsch, S. and Stegemann, T. (2012) The brain and positive biological effects in healthy and clinical populations. *In Music, Health, and Wellbeing* (MacDonald, R.A.R. et al., eds), pp. 436–456, Oxford University Press

Krumhansl, C. L. (1997). An exploratory study of musical emotions and psychophysiology. *Canadian Journal of Experimental Psychology*, (4), 336.

Laurie-Rose, C., Bennett-Murphy, L., Byard, L., & Nikzad, K. (2002). The role of the big-five personality factors in vigilance performance and workload. *European Journal of Personality*, *16*, 185-200.

Laurie-Rose, C., Frey, M., Ennis, A., & Zamary, A. (2014). Measuring Perceived Mental Workload in Children. *The American Journal of Psychology*, (1). 107.

Laurie-Rose, C., Frey, M., Sibata, E., & Zamary, A (in press). The effects of memory load and test position on short duration sustained attention tasks. *The American Journal of Psychology.*

Levin, S., France, D. J., Hemphill, R., Jones, I., Chen, K. Y., Richard, D., . . . Aronsky, D. (2006). Tracking workload in the emergency department. *Human Factors*, 48, 526–539.

Lysaght, R. J., Warm, J. S., Dember, W. N., & Loeb, M. (1984). Sex differences on a cognitive vigilance task within the context of noise exposure and relative quiet. In A. Mital (Ed.), *Proceedings of the First Mid-Central Conference on Human Factors*. North Holland, Amsterdam.

Matthews G, Warm, J. S., Reinerman-Jones, L. E., Langheim, L. K., Washburn, D. A., & Tripp, L. (2010). Task engagement, cerebral blood flow velocity, and diagnostic monitoring for sustained attention. *Journal of Experimental Psychology: Applied, 16,* 187-203.

Moroney, W. F., Biers, D. W., & Eggemeier, F. T. (1995). Some measurement and methodological considerations in the application of subjective workload measurement techniques. *International Journal of Aviation Psychology*, 5, 87–106.

Neuchterlein, K. H., Parasuraman, R., & Jiang, Q. (1983). Visual sustained attention: Image degradation produces rapid sensitivity decrement over time. *Science, 220,* 327-329.

Panksepp, J. (1995). The Emotional Sources of 'Chills' Induced by Music. *Music Perception: An Interdisciplinary Journal*, (2). 171.

Parasuraman, R., & Davies, D. R. (1977). A taxonomic analysis of vigilance performance. In R. R. Mackie (Ed.), *Vigilance: Theory, operational performance, and physiological correlates* (pp. 559-574). New York: Plenum.

Perham, N., & Withey, T. (2012). Liked music increases spatial rotation performance regardless of tempo. *Current Psychology*, (2), 168.

Pietschnig, J., Voracek, M., & Formann, A. K. (2010). Mozart effect–Shmozart effect: A meta-analysis. *Intelligence*, 38314-323.

Pittman, S. and Kridli, S. (2011) Music intervention and preoperational anxiety: an integrative review. *Int. Nurs. Rev*. 58, 157–163

Rickard, N. (. (2004). Intense emotional responses to music: A test of the physiological arousal hypothesis. *Psychology of Music*, 32(4), 371-388.

See, J. E., Howe, S. R., Warm, J. S., & Dember, W. N. (1995). Meta-analysis of the sensitivity decrement in vigilance. *Psychological Bulletin, 117*, 230-249.

Shaw, T. H., Matthews, G., Warm, J. S., Finomore, V. S., Silverman, L., & Costa, P. T. (2010). Individual differences in vigilance: Personality, ability and states of stress. *Journal of Research in Personality, 44,* 297-308.

Sloboda, J. (1991). Music structure and emotional response: Some empirical findings. *Psychology of Music*, 19(2), 110-120.

Sloboda, J., Ivaldi, A., & O'Neill, S. (2001). Functions of music in everyday life: An exploratory study using the experience sampling method. *Musicæ Scientiæ: The Journal of The European Society For The Cognitive Sciences of Music*, *5*(1), 9-32.

Soto, D., Funes, M. J., Guzmán-García, A., Warbrick, T., Rotshtein, P., Humphreys, G. W., & Thompson, B. (2009). Pleasant Music Overcomes the Loss of Awareness in Patients with Visual Neglect. *Proceedings of the National Academy of Sciences of the United States of America,* (14). 6011.

Stanislaw, H., & Todorov, N. (1999). Calculation of Signal Detection Theory Measures. *Behavior Research Methods, Instruments, & Computers,* 31(1), 137-149.

Stone, N. J. (2008). Human Factors and Education: Evolution and Contributions. *Human Factors*, 50(3), 534-539.

Temple, J. G., Warm, J. S., Dember, W. N., Jones, K. S., LaGrange, C. M., & Matthews, G. (2000). The effects of signal salience and caffeine on performance, workload and stress in an abbreviated vigilance task. *Human Factors*, 42, 183–194.

Tenenbaum, G., Lidor, R., Lavyan, N., Morrow, K., Tonnel, S., Gershgoren, A., & ... Johnson, M. (2004). The effect of music type on running perseverance and coping with effort sensations. *Psychology of Sport & Exercise*, 589-109.

Thompson, W. F., Schellenberg, E. G., & Husain, G. (2001). Arousal, Mood, and the Mozart Effect. *Psychological Science*, (3). 248.

Warm, J. S. (1984). *Sustained attention in human performance*. Chichester [West Sussex] ; New York : Wiley, c1984.

Warm, J. S., & Jerison, H. J. (1984). The psychophysics of vigilance. In J. S. Warm (Ed.), *Sustained attention in human performance* (pp.15-20). Chichester: John Wiley & Sons.

Warm, J. S., & Jerison, H. J. (1984). The psychophysics of vigilance. In J. S. Warm (Ed.), *Sustained attention in human performance* (pp.15-20). Chichester: John Wiley & Sons.

Warm, J.S., Dember, W. N., & Hancock, P. A. (1996). Vigilance and workload in automated systems. In M. Mouloua & R. Parasuraman (Eds.), *Automation and human performance: Theory and application* (pp. 183-200). Mahwah, NJ: Erlbaum.

Waterhouse, J., Hudson, P., & Edwards, B. (2010). Effects of music tempo upon submaximal cycling performance. *Scandinavian Journal of Medicine & Science In Sports*, 20(4), 662-669.

Wiebe, E. N., Roberts, E., & Behrend, T. S. (2010). An examination of two mental workload measurement approaches to understanding multimedia learning. *Computers in Human Behavior*, 26, 474–481

Wierwille, W. W., & Eggemeier, F. T. (1993). Recommendations for mental workload measurement in a test and evaluation environment. *Human Factors*, 35, 263–281.

Wininger, S. R., & Pargman, D. (2003). Assessment of Factors Associated with Exercise Enjoyment. *Journal of Music Therapy*, 40(1), 57-73.

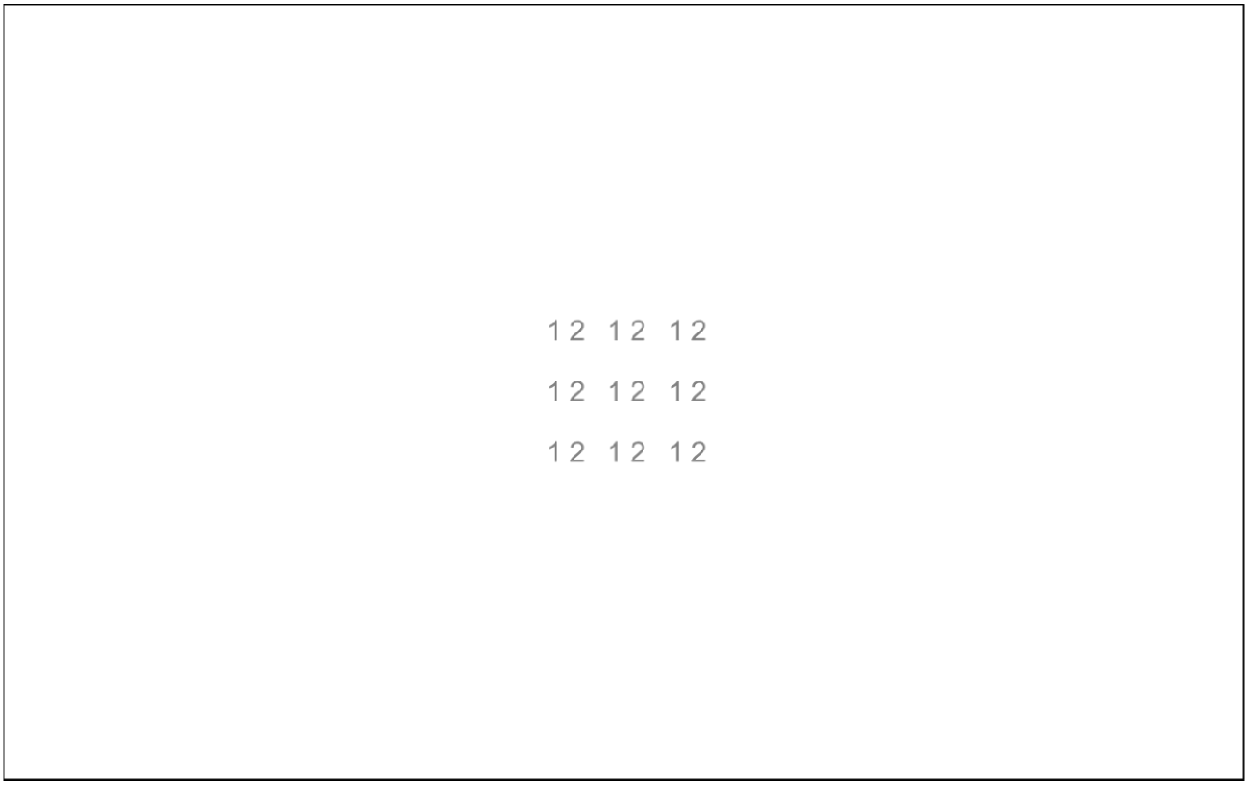

Figure 1. Stimuli Locations.

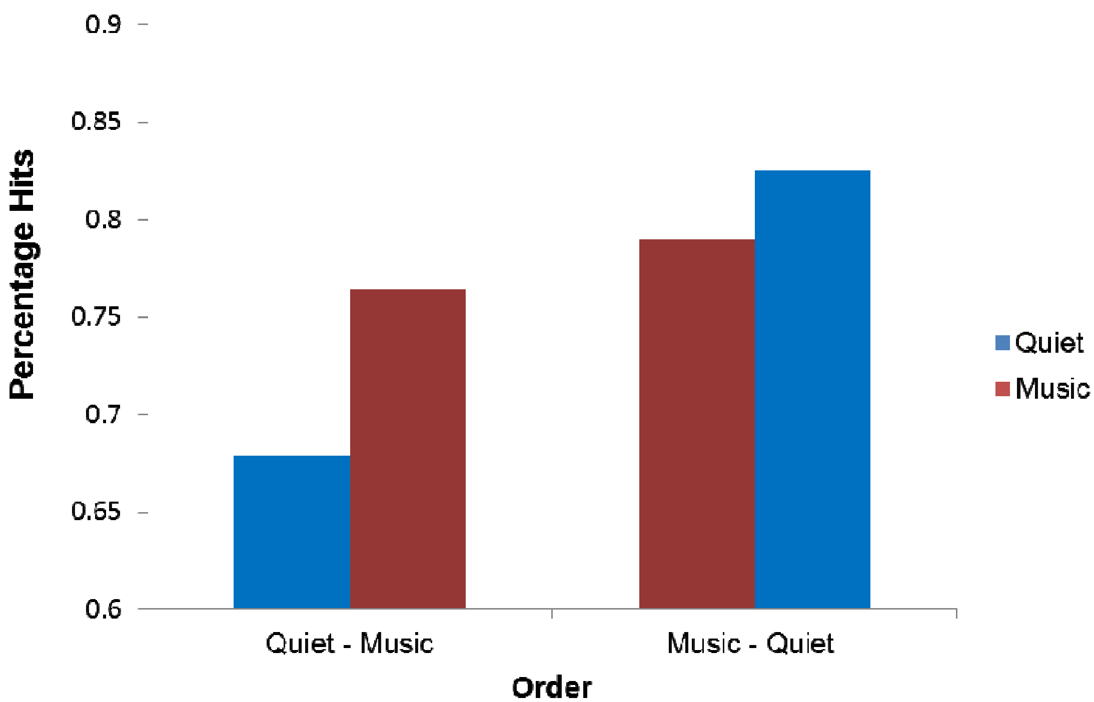

Figure 2. Summary plots of percentage hits for all conditions in both orders.

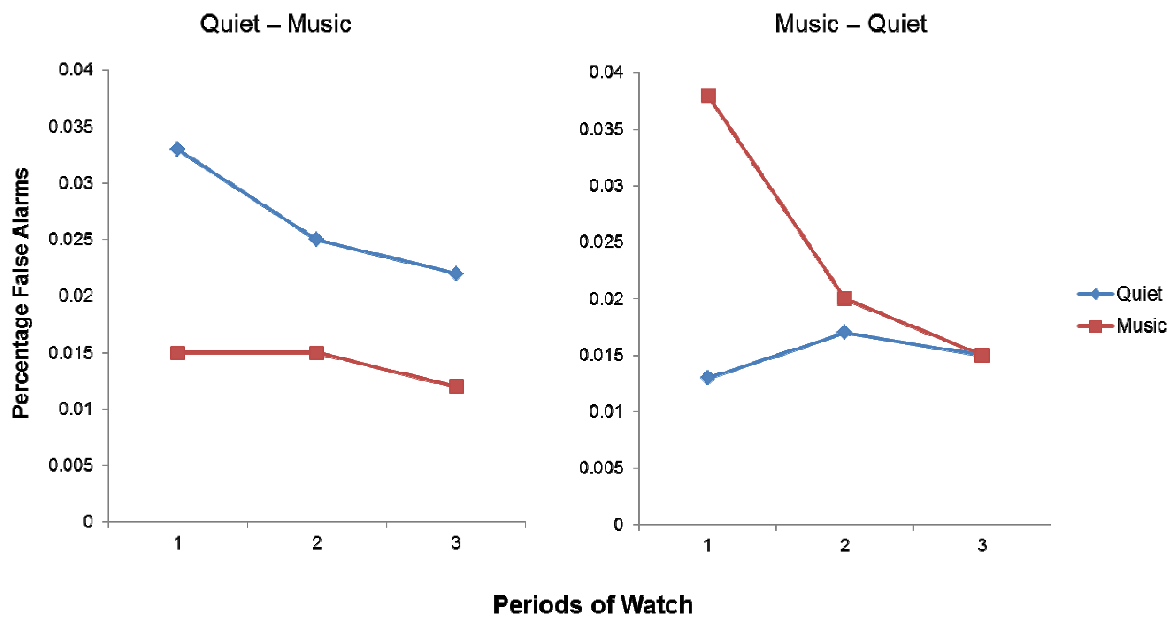

Figure 3. Summary plots of percentage false alarms for all periods of watch across both music conditions in both orders.

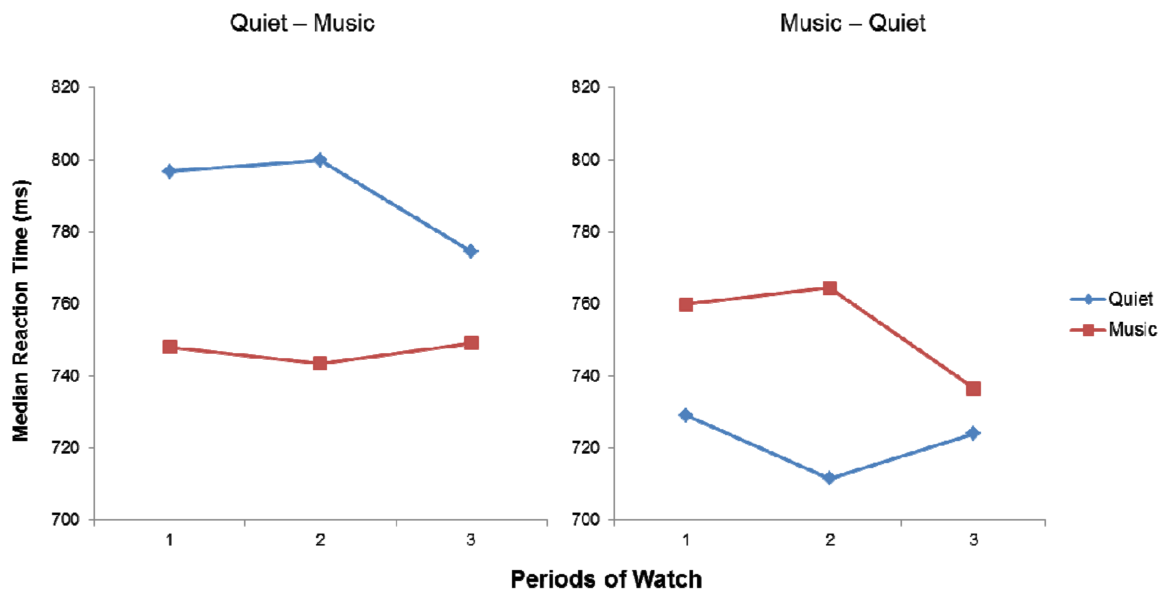

Figure 4. Summary of median reaction time for all periods of watch across both conditions in both orders.

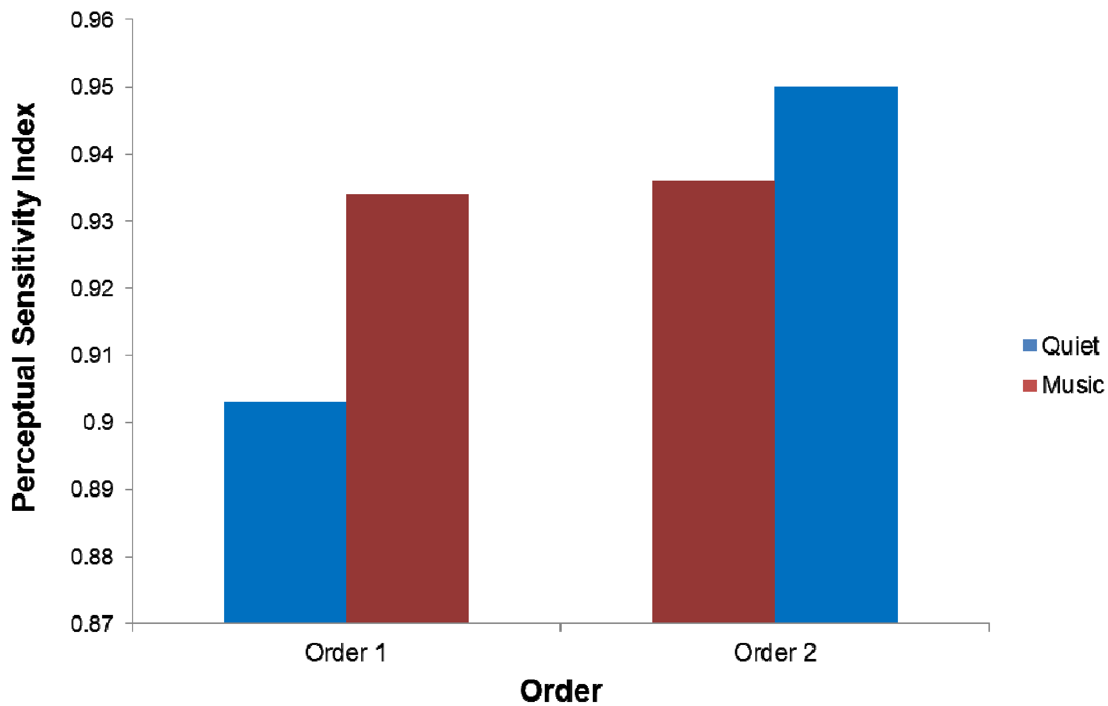

Figure 5. Summary of Perceptual Sensitivity for all music conditions across both orders.

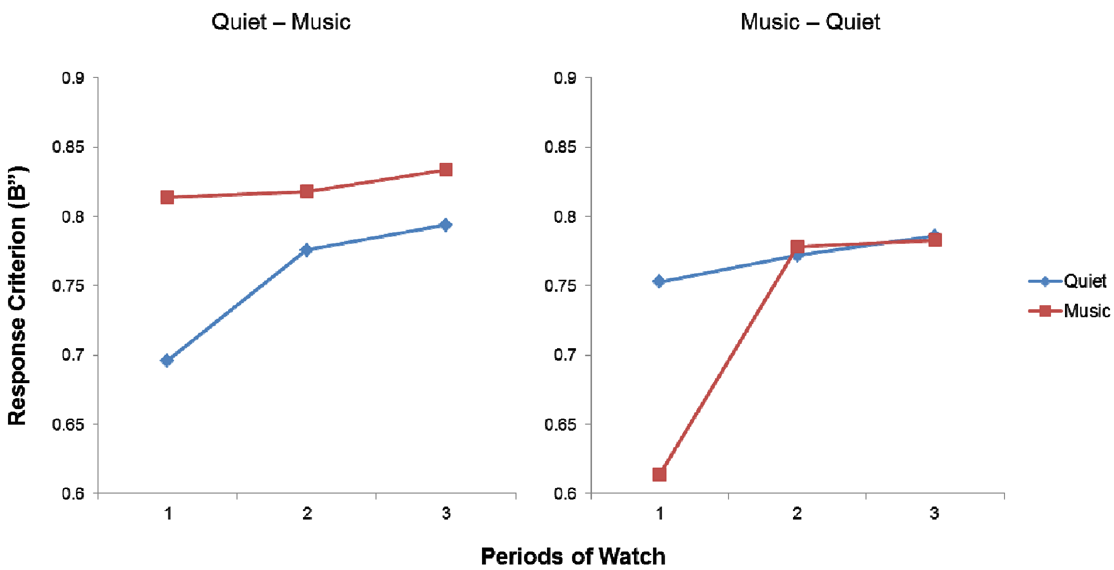

Figure 6. Summary of Response Criterion for all periods of watch across all music conditions in both orders.

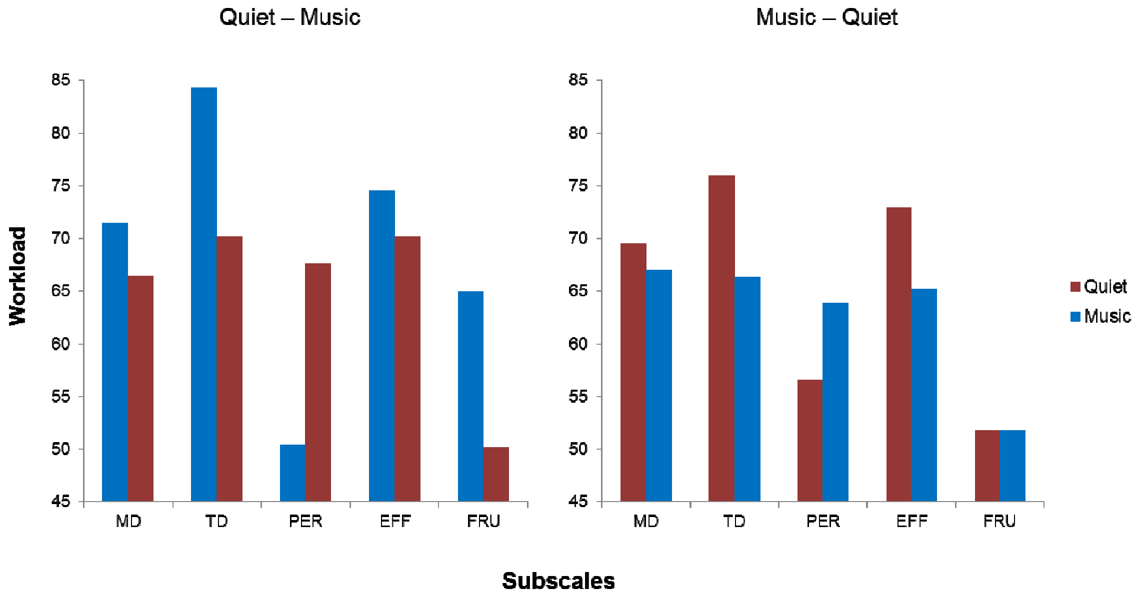

Figure 7. Summary of Workload for all music conditions across both orders.