BENEFITS AND COST OF DUAL-TASKING IN A VIGILANCE TASK:

A LABORATORY AND DRIVING SIMULATOR INVESTIGATION

BY

C2008

Mark Sai Leong Chan

Submitted to the graduate degree program in Cognitive Psychology
and the Faculty of the Graduate School of the University of Kansas
in partial fulfillment of the requirements for the degree of
Master's of Arts

____

Paul Atchley, Ph.D., Chair

Committee Members:

____

Ruth Ann Atchley, Ph.D.

____

Steven D. Schrock, Ph.D., P.E.

Date Defended: Dec 9th 2008

The Thesis Committee for Mark Sai Leong Chan certifies

That this is the approved version of the following thesis:

BENEFITS AND COST OF DUAL-TASKING IN A VIGILANCE TASK:

A LABORATORY AND DRIVING SIMULATOR INVESTIGATION

____

Paul Atchley, Ph.D., Chair

Committee Members:

____

Ruth Ann Atchley, Ph.D.

____

Steven D. Schrock, Ph.D., P.E.

Date Approved: ______________________

Abstract

It is believed that under certain conditions, the presence of a secondary task such as a cell phone conversation would minimize a decrease in vigilance. The current study investigates this assumption by using two different vigilance paradigms. Further investigations were done by applying the same secondary task conditions to a monotonous driving scenario in a simulator. Results from the vigilance studies showed robust effects of dual task interference, and improvement in task performance for participants engaged in dual task from beginning to end. It was noted that the benefit of an improvement in task performance did not outweigh its cost as the reported improvement only reached a level similar to that of an individual who was low in vigilance. Results from the driving simulator indicated a possible driving improvement with the presence of a secondary task during later stages of the driving task as indicated by smaller lane keeping variability. The perceived improvement was questioned as there was a significantly poorer recall memory under dual task conditions. In general, it might be suggested that a secondary task may improve task performance under vigilance conditions, but the reported benefit may not outweigh its costs.

Acknowledgements

The project would not have happened without the help and guidance of the many people I've had the opportunity to come in contact with. I would like to first thank my advisor, Dr. Paul Atchley for his guidance and support in the development of this project. Your patience and faith in me has been most crucial through the course of this project. Thank you.

Thanks and appreciation is also extended to Dr. Ruth Ann Atchley and Dr. Steven Schrock for being on my defense committee. Your input and insightful questions have been most helpful as I continue to fine tune my craft.

Thanks also go out to my lab mates, Jeffery Dressel and David Marshall for their input into the project. I would also like to thank my graduate student colleagues who have often shared kind words of encouragement throughout the entire process. You guys have made my time here so much more enriching and enjoyable. Special mention also goes out to Ting-Shiun for his kind support.

Papa, Mummy, and Jie Jie. I love you all, and I thank you all for keeping me in your thoughts and prayers. Papa and Mummy, thank you for giving me the opportunity to further my studies here in the United States. I'm not done yet, but I'm almost there!

S.D.G.

Table of Contents

List of Tables

List of Figures

1. Introduction

The study of attention dates back to the formative years of psychology. During that period, William James made the bold statement that "[e]veryone knows what attention is" (James, 1890, p.403). William James did not stop at that statement as it might have then been mistaken as a claim for a unitary, singular attentional process. Instead, he goes on further to describe various aspects of attention which at that time could not be studied for some reason or another. In a detailed review of the attentional system**,** Posner and Peterson (1990) described the attentional system as an anatomically independent system, separate from other systems such as the motor and information processing systems. They further stated that although the systems were anatomically independent, the interactions of the systems were needed for human cognition. The multifaceted attentional processes described by William James were broken into three main subsystems as suggested by Posner and Peterson. These subsystems were: orienting, alerting, and executive functioning. Orienting was defined as the ability to parse necessary information from input stimuli. Alerting was defined as the ability to maintain a state of readiness for the possible presence of incoming stimuli. Executive control was defined as the ability to resolve and select the best course of action among competing stimuli (Fan, McCandliss, Sommer, Raz & Posner, 2002; Raz & Buhle, 2006). It was noted in reviews by (Robertson, & Garavan, 2004; Raz & Buhle, 2006) that among the three attentional subsystems described by Poser and Peterson (1990), the alerting subsystem received the least amount of attention and research. This was in comparison to other aspects of attention such as the orienting and executive subsystems. Though viewed as anatomically independent, the interactions of the subsystems proposed by Poser and Peterson (1990) are central to human cognition.

1.1. History of Vigilance Research

Research into the alerting subsystem also known as sustained attention, or vigilance, began in earnest during the period of the Second World War. The advent of radar technology gave countries advanced warning of potential threats intruding into their territory and radar operators were required to maintain vigil for extended periods of time. Yet it was not known at that time what factors would affect observers in having to maintain a high level of sustained attention and how long observers would be able to maintain a high level of performance. In the mid 1940s, Norman Mackworth was commissioned by the British RAF to conduct studies into radar operator performance. The pioneering work that he had done was considered to be the first major piece of research in the area of vigilance (Stroh, 1971)**.** Mackworth's (1948) study investigated the performance of radar operators and their ability to detect infrequent changes on a clock observation task. He described vigilance as a state of readiness to detect and respond to rare and random events that occur while engaging in a specified task. Results from his study showed that participants were able to detect changes at a high level at the onset of the task. Yet, as time progressed, participants began to miss these changes, suggesting a decline in observer efficiency (Mackworth, 1948). This decline in observer efficiency over time has been ubiquitous in vigilance research and has been termed the vigilance decrement (Davies & Parasuraman, 1982).

1.2. Traditional Theories of Vigilance

Most theories of vigilance have focused primarily on the decrement of performance over time rather than overall performance. This emphasis might be attributed to the fact that given an understanding of the factors that might contribute to a decrement in task performance, contingencies can be put into place to minimize these decrements, which in turn may improve performance. Studies into the vigilance decrement have led researchers to hypothesize plausible theories that caused the decrement. These theories had their merits and shortcomings and have been able to describe certain aspects of the vigilance decrement. This section will give a broad overview of the main theories of vigilance that were proposed by researchers since Mackworth's (1948) work.

1.2.1. Inhibition & Habituation Theory

Rooted in the behaviorist-conditioning mindset of his time, Mackworth's theory of vigilance was based largely on the inhibition of responses due to the lack of reinforcement. Mackworth's inhibition theory was almost circular in nature, whereby a correct response to the target signal was reinforcing in nature. Yet at the same time, as the vigilance decrement sets in, participants failed to respond to these targets. Thus the lack of a reinforcer led to response inhibition thereby leading to poorer future responses. Mackworth further supported his theory by suggesting that knowledge of results and rest breaks allowed for the dissipation of inhibition (Cited in Davies & Parasuraman 1982). The inhibition theory did not take a firm footing in the field due to findings contrary to Mackworth's prediction as performance in tasks improved even with an increase in target signal rates (Davies, & Parasuraman, 1982).

Posner (1978) suggested an alternative inhibitory theory that implicated event rates (non target signals) as responsible for the building up of inhibition. Posner believed that visual and auditory pathways could be inhibited by repeated presentations of stimuli, which would result in a decreased detection rate leading to a vigilance decrement. Galinsky, Warm, Dember, Weiler, and Scerbo (1990) investigated Posner's hypothesis by comparing the detection rates in three different task types while manipulating the presentation rates of events. It was believed that by alternating the modality of the presentation, the effects of repeated presentations to any one pathway could be minimized. Results of Galinsky et al. (1990) found partial support for Posner's hypothesis of pathway inhibition as greater event rates had an effect on target signal detection efficiency, where the greater the event rate, the greater the decrease in target detection rates. But this was only true for participants in groups that had events presented in one modality. Participants that had signals alternating between modalities failed to show the effect of the rate of event presentation on target signal detection. In tasks where participants had to monitor both visual and auditory modalities, the effect of event rates did not differ. Therefore, Posner's hypothesis could only account for the effect of differing event rates on target detection in tasks that were presented either visually or auditorily but not on tasks that alternated modalities. Furthermore, the hypothesis was unable to give a satisfactory explanation for the vigilance decrement that was observed in all three task types. (Matthews, Davies, Westerman, & Stammers, 2000).

A well observed phenomenon that occurs in behavioral and physiological studies known as habituation was thought to be a possible explanation for the vigilance decrement. Habituation might be defined as a reduced physiological or behavioral response due to repeated stimulation (Groves, & Thompson, 1970). Parasuraman (1985) conducted a study that investigated the N1 component, which was previously found to exhibit habituation. Manipulating event rate, and the vigilance task (active or passive), he found the expected N1 amplitude decrease across all conditions. More importantly, results indicated that the rate of habituation was similar for all conditions. This finding would suggest that regardless of event rate, participants exhibited similar rates of habituation, with the only difference being the amplitude of the N1 component.

1.2.2. Expectancy Theory

The general concept behind expectancy theory suggested that observers were constantly formulating expectancies about future events based on previous experiences with the task. This formulation of expectancies would determine the rate of detection for the task, as the observer would be in a constant state of averaging and reviewing earlier presentations of the task (Davies & Parasuraman, 1982). The notion of expectancy exerts an important role in the criterion response variable in signal detection theory. For example, Deese (1955) stated that the frequency of target signals would play an important role in the formation of an observer's expectancy. Therefore if a signal appeared more often, expectancy would increase, and vice versa. This expectancy would cause an observer's criterion response to shift accordingly. Warm & Dember (1998) also suggested that when events occurred at regular predictable intervals, participants were better able to anticipate future presentations as compared to events that were presented at irregular intervals. Furthermore, Baker (1962) also suggested the provision of feedback also helped improved detection rates as feedback would allow participants to plan ahead with regard to future target signals (Matthews et al., 2000).

Another important finding regarding expectancy can be attributed to the probability of target signals being overestimated at the beginning of the task as suggested by Craig (1978). It was suggested that as time lapsed, observers began to lower their expectancy as they learned that the probability of a target signal occurring was much lower than expected. This caused an upward shift toward a more conservative response criterion causing participants to lower their responses to targets. Thus Craig argued that the observed vigilance decrement in observers was a case of probability matching, rather than an actual decrement (Matthews et al., 2000).

1.2.3. Filter Theory

The filter theory was based largely on Broadbent's (1958) model of attention. The attentional filter suggested by Broadbent was a hypothetical mechanism that functioned as sieve that selected task relevant information from the environment for processing. According to Broadbent, this filter had biases that caused it to select stimuli from novel sources. Jersion and Pickett (1964) found that in tasks where the presentation of targets was increased, target detection was decreased. Based on the filter theory, the decrease could be the result of the filter biasing itself toward novel stimuli beyond the task, thus causing the observer to miss the embedded target signals that occurred during the task. The theory also suggests that the rate of decrement in self-paced vigilance tasks would be less pronounced than those of timed vigilance task. The reason for this prediction was that in self-paced tasks, observers could take breaks if they felt that they were not paying sufficient attention to the task, thus minimizing the biasing nature of the filter.

1.2.4. Arousal Theory

There have been many different definitions of the term arousal. Some researchers viewed arousal as a spectrum with a state of comatose at one end, increasing to a state of extreme excitement (Duffy, 1962; Lindsley, 1951) while others defined arousal as a state of readiness to respond (Berlyne, 1960). This theory was one of the more troublesome theories of vigilance due to differing definitions of arousal. The general premise of this theory used by vigilance researchers was that task performance declines as time elapses due to decreased arousal. Davies and Parasuraman (1982) stated that there would be visible declines in electrocortical and autonomic activity during the performance of vigilance related tasks and this decline was a factor leading to a decrease in task performance. Parasuraman (1984) stated that though arousal might decline while engaged in a vigilance task, the decline in arousal did not necessarily translate to a vigilance decrement. For example, if an observer was hypoaroused due to sleep deprivation or alcoholic intoxication, his overall level of vigilance would be lower than normal, which would in turn lead to a lower than normal task performance. This would further implicate that arousal levels play a part in the overall level of vigilance in the observer, but not the only factor that affects vigilance. Research into the effects of alcohol induced drowsiness and vigilance performance show decreased overall levels of vigilance, but similar patterns of vigilance decrements when compared to controls (Erwin, Wiener, Linnoila, & Truscott, 1978). In general, arousal levels would decrease in most situations; for example individuals not engaged in any task would similarly exhibit decreased levels of arousal (Matthews et al., 2000).

1.2.5. Motivation Theory

This theory did not consider the psychophysical aspects of the vigilance task. Rather it focused on the intrinsic motivations of the individual observer. Smith (1966) stated that most normal and healthy individuals would have little trouble engaging themselves in a vigilance task. He believed that the willingness of the participant would play a larger role. Given a pool of observers, he believed that the observed vigilance decrement was a result of averaging the scores of willing observers who performed to their limits, and observers who felt no intrinsic motivation toward the task (Davies & Parasuraman, 1982). Some researchers (Antonelli & Karas, 1967; Warm, Epps & Ferguson, 1974) have attempted to improve extrinsic levels motivation in observers by manipulating the levels of feedback and knowledge of results. They found that by allowing observers to know the knowledge of their results, they were able to reduce, and in some instances, abolish the vigilance decrement.

Other researchers have attempted to improve extrinsic motivation with financial incentives, but have received mixed results. One group of researchers (Levine, 1966; Wiener, 1969) found that financial incentives did not improve the overall level of performance. On the other hand, other researchers (Bevan & Turner, 1965; Spiowicz, Ware & Baker, 1962) found that financial incentives had a positive effect on the overall level of performance. The results appear to be equivocal thus suggesting that motivation might account for some aspect of vigilance performance.

1.2.6. Summary of Early Theories

Most early theories of vigilance shifted from a purely behaviorist stance with the exception of Mackworth's theory of inhibition. Understandably, the reason for Mackworth's stance was due to the prevailing behaviorist mindset when he theorized a possible explanation for the decrease in vigilance. The theories that developed beyond Mackworth's theory found support using the principle of Signal Detection

Theory (Green & Swets, 1966). The general premise of the Signal Detection Theory proposed that an individual's response could be statistically derived from the signal to noise ratio, and response criterion set by the individual in any given environment. Firstly, the greater the signal to noise ratio, the lower the individual's perceptual sensitivity. For example, an individual's sensitivity is likely to be decreased, when stressed beyond normal conditions. Secondly, the response criterion set by the individual would fall within the bounds of the signal to noise ratio. This response criterion could shift to a more conservative or liberal criterion based on the task, and was believed to shift to a more conservative level as time on task increased, but might be mitigated if the probability of a signal was predictable (Davies & Parasuraman, 1982).

The psychophysical properties were broken into two orders: first order factors were described as physical parameters of the signal (Dember & Warm 1979; Warm & Jerison 1984; Davies & Parasuraman, 1982). These parameters would include sensory modality, signal intensity, signal duration, and background event rate (Parasuraman, Warm, & See, 1998). Sensory modality would be an obvious factor in any task. The vigilance decrement occurs in all modalities yet it has been noted that different modalities exhibit different levels of decrement. For example, some researchers (Sipowicz & Baker, 1961; Ware, 1961) found that the decrement function was less steep in auditory vigilance tasks as compared to analogous vigilance tasks in the tactile or visual modality. Signal intensity and signal duration can be grouped together as signal conspicuity. It was been widely accepted that the greater the signal intensity, be it in terms of volume in auditory tasks or brightness in visual tasks, the easier the detection of the signal. Loeb and Binford (1963) had participants in an auditory vigilance task detect target signals of varying intensity. They reported that signals at the highest intensity showed the greatest performance. The background event rate also known as the neutral stimuli that participants do not respond to, also affected task performance. Jerison and Pickett (1964) reported that higher event rates led to poorer task performance

Second order factors were described as characteristics of the signal that had to be inferred by the observer through experience (Davies & Parasurman, 1982). The parameters included temporal and spatial uncertainty. By varying the probability of target signals, the regularity or irregularity at which target signals were presented, and the regularity or irregularity of background non target events, it was reported that performance degradation was greatest under irregular presentations of events (Warm, Dember, Murphy, & Dittmar, 1992). Spatial uncertainty requires participants to engage in a visual search of targets that occur in varying locations. Thus, performance decrements would be greater in tasks where targets occur under unpredictability (Mouloua, & Parasuarman, 1995). The combination of these two psychophysical factors and signal detection theory helped early vigilance theorist build up a credible set of hypotheses.

1.3. Recent Changes in Classic Vigilance Theory - Resource Theory

A more recent treatment of vigilance theory incorporates the possibility of cognitive resources (Kahneman, 1973; Wickens, 1984, 2002) having a role to play in vigilance. Davies and Parasuraman (1982) noted that no singular theory would be able to account for the vigilance decrement. Therefore current vigilance theory might be viewed as a combination of earlier versions of vigilance theories, combined with current research on attention and workload.

1.3.1. Task Type

An earlier treatment of resource theory, though not explicitly stated, was the Parasuraman and Davies taxonomy of vigilance (Parasuraman & Davies, 1977; Davies & Parasuraman, 1982). One of the core aspects of the taxonomy was the distinction of the types of vigilance tasks, namely successive and simultaneous discrimination. Successive discrimination tasks present stimuli in a serial fashion. Either targets or non-target signals were presented one at a time. Observers were required to maintain a memory of the target stimuli in working memory and respond when presented with the target during the task. This required the observer to be vigilant and distinguish targets from non-targets. In simultaneous discrimination tasks there would be trials where targets would be presented simultaneously with non-targets, while on other trials, there would only be non-targets. This simultaneous presentation of targets and non-targets did not require the observer to hold a standard in working memory as comparisons were made online. This led Parasuraman and

Davies (1977) to propose that because of a greater memory load, successive discrimination tasks were more resource demanding than simultaneous discrimination tasks (Parasuraman et al., 1998). Increasing task demands would cause shifts in criterion on perceptual sensitivity which in turn would have an effect on performance.

1.3.2. Subjective workload

The need to understand the effects of workload on human performance was indicative of the usage of information processing resources in any given task (Warm, Parasuraman, & Matthews, 2008). Workload being a subjective construct was difficult to measure. One of the most consistent and valid work load questionnaire developed and used by researchers is the NASA-Task Load Index (NASA-TLX). The NASA-TLX is a multi-dimensional questionnaire that measures the physical, mental and temporal demands of the task on the observer, and how the observer perceives these aspects of the task based on scales of performance, effort, and frustration (Hart & Staveland, 1988; Warm et al., 2008). Results from the NASA-TLX often show observers rating vigilance type tasks as highly mentally demanding, even more so than other cognitive tasks such as, grammatical reasoning, simple tracking, and mental arithmetic (Warm & Dember, 1998). Thus using this subjective work load questionnaire, researchers have been able to infer that vigilance tasks being mentally taxing require attentional resources, lending support to the resource theory.

1.3.3. Neurophysiological Correlates

Neuroimaging studies (Pardo, Fox, & Raichle, 1991) have successfully shown regions of the brain that have increased blood flow and activation when engaged in vigilance type tasks. This increase in blood flow often suggests an increase in cortical activity, which implies an increased load on specific regions of the brain. Although advances have been made with regard to the localization of regions specific to sustained attention, these findings do not shed light on performance efficiency as pointed out by Parasuraman et al., (1998). The reasons for this were often related to the nature of vigilance tasks being long, tedious, and often monotonous. The high cost of the using neuroimaging equipment made it economically unfeasible for studies of long duration. Furthermore, the need to remain motionless for extended periods of time is not easily achieved (Parasuraman et al., 2008). With recent advances in cheaper non-invasive technology, researchers were been able to investigate cerebral blood flow in vigilance tasks by using transcranial Doppler sonography (TCD). What TCD lacks in spatial resolution, it makes up in temporal resolution, thus making it highly suitable for measuring real-time blood flow. In general, TCD allows researchers to record and measure the differences in blood flow velocity when a participant is engaged in a mental task. Studies using TCD in vigilance tasks (Hitchcock et al., 2003; Schnittger, Johannes, Arnavaz, & Munte, 1997) show a correlation between the decrement in blood flow velocity, and the vigilance decrement. This decrease in blood flow velocity was indicative of depletion in the availability of resources to the task, thus leading to a performance decrement.

In general, the resource theory of vigilance is driven primarily by varying task demands in a vigilance task. It has been suggested that the demand on resources should increase when a task requires the increased use of working memory or causes a shift in either response criterion or perceptual sensitivity. When a task is deemed mentally taxing, based on subjective taskload scores, it is also deemed resource intensive. Finally, when a task is able to decrease the blood flow velocity to regions subserving sustained attention, it is indicative of a reduction of resources to the area.

1.4. Preview of Traditional Vigilance Methodology

The previous section had highlighted various components of traditional vigilance methodology. As noted there are many manipulations that researchers could use to investigate vigilance in observers. The current investigation was not focused on the aforementioned manipulations. Traditional vigilance tasks follow a basic response paradigm where the observer was required to respond only to a target that was presented at random. The Multiple Vigilance Task (MVT) (Hirshkowitz, De La Cueva, & Herman, 1993) was one such test where observers were required to respond only to a target stimulus. This response paradigm would be used in comparison to an emerging view of vigilance, discussed in the next section.

1.5. Emerging Views on Vigilance

Many of these approaches to vigilance have focused on the vigilance decrement, which was characterized by a significant drop and leveling off in observer performance over time. Yet some researchers argue against the validity of the decrement outside of the laboratory (Mackie, 1987). Furthermore, as argued by Robertson and Garavan (2004), the prevailing paradigm of vigilance type tasks required observers to respond to a rare target, suggesting an exogenously cued response. Yet little has been discussed as to the inhibition of ongoing behavior, driven by an endogenously driven attentional system.

As noted in the previous section, vigilance researchers have primarily used detection rates, false alarms rates and response latencies over an extend period of time as measures of vigilance. There has been little reference to errors of omission or misses, which as noted by Ballard (1996) was due to the fact that researchers combine false alarm rates (errors of commission) and misses (errors of omission) as a unitary concept. Furthermore, as noted by other researchers (Helton, 2008; Manly et al., 1999; Manly et al., 2002) these errors, especially errors of omission often reach ceiling levels thus are often overlooked by researchers. This is especially true in vigilance type studies. By ignoring these errors, researchers might miss effects that may hint at other measures of performance degradation.

Most work in vigilance has centered on performance decrement for obvious reasons as there has been a need to minimize human error due to a failure of attention in mission critical tasks. Understandably, these studies lean toward a more exogenous view of attention where the individual’s task was to detect the target and respond accordingly. This was evidenced by psychophysical factors described by Dember and Warm (1979), suggesting an exogenously modulated attentional system. Furthermore, "much greater emphasis has been placed on the development of answers to the [vigilance decrement] than to the overall level of performance, and most theories of vigilance are devoted exclusively to an explanation of the vigilance decrement" (Davies & Parasuraman, 1982, p.9). There has been little research that investigates the function of endogenous attention in vigilance task, possibly due to the invoking of the oft maligned supervisory attentional system. Furthermore, with the focus of vigilance being put on task performance decrement over time, there has been a lack of work in the investigation of the overall levels of vigilance in individuals, and attentional lapses in individuals in everyday situations. Work by Robertson, Manly, Andrade, Baddeley, and Yiend (1997) sought to redirect vigilance research, especially with respect to the endogenous aspect of sustained attention.

1.5.1 Mindlessness & Task Unrelated Thoughts

Traditional vigilance research tasks require observers to monitor streams of events and respond only to the appropriate target event. Recent studies have begun to switch focus from the vigilance decrement, to slips of attention in everyday situations.

To investigate these lapses of attention, researchers (Robertson et al., 1997) needed to create an environment that would induce a state where slips of attention could easily occur in the observer, and shift behavior towards an automatic, rather than controlled process. The Sustained Attention Response Task (SART) was designed by Robertson et al., (1997) as an alternative measure to investigating lapses of sustained attention in individuals who had sustained traumatic brain injury (TBI).

The SART required observers to respond to single digits ('go') appearing at a constant speed on the computer monitor and make key press responses to these digits. The repeated and fairly rapid "mindless stimulus-press, stimulus-press style" (Manly et al., 1999, p. 662) behavior was believed to induce a state of automaticity, which in turn would draw away the need for controlled endogenous attention to the task. During the task, a random and infrequent, 'no-go' target would be presented. The presentation of the 'no-go' required observers to withhold their response to the target. The ability to withhold a response was believed to be modulated by the supervisory attentional system as described by Stuss, Shallice, Alexander, and Picton (1995). Results indicated that the SART was a sensitive measure and could discriminate between TBI patients and normal controls. This was evidenced by significantly increased error rates in patients, and the inability to modify their response times even after making an error (Robertson et al., 1997). This difference in performance was believed to be due to the effects of diffuse damage to regions of the brain that modulated sustained attention in TBI patients (Manly et al., 2002).

Further support for the SART being an effective test in discriminating between TBI patients and non-TBI individuals was conducted by Chan (2001). The study investigated whether age, gender and education would have an effect on SART performance. Results from his study indicated minimal effects of age, gender on SART performance, while also finding that TBI patients did poorer on the SART as compared to controls. To further establish the credibility of the SART as a test of being able to endogenously sustain attention, Manly et al., (1999) conducted a series of experiments that systematically controlled for regularity of targets. Results from their study suggested that the when targets were more regular and predictable, observers' accuracy scores were significantly higher, and reaction times to non-targets were longer. The results led Manly and colleagues to suggest that the increase of target frequency led to the diminishing of endogenous control and an increase in exogenous control in the experiment.

Recent interest in lapses of attention led researchers to entertain the possibility of investigating automaticity playing a role in attentional lapses and mind-wandering episodes when engaged in a 'mindless' task such as the SART. Smallwood et al., (2004) conducted a series of experiments that used a variety of measures that investigated task engagement and disengagement during sustained attention. For the purposes of their study, observers were grouped into high or low task unrelated thought (TUT), and high or low task related intrusions (TRI), based on a post task retrospective questionnaire. Regardless of whether observers rated themselves as having engaged in either high or low TUTs, reaction time was faster under a low probability target presentation. This result suggests that a byproduct of TUTs would be a drift of attention. This drifting occurred in all individuals regardless of their perceived degree of TUTs. This finding further supports Robertson and colleagues (Manly et al., 1999; Robertson et al., 1997) claim on the relationship between faster reaction time and attentional drifts where faster reaction times was an indicator of waning attention. Secondly, it was also reported that observers who rated themselves highly on task related intrusions (TRI) were quicker to respond to targets as compared to those who rated themselves lower on TRI. This suggests that thinking of the task may also lead to drifts in attention. Furthermore, observers who rated themselves as having more task related intrusions, also made more errors as compared to observers who rated themselves lower on the TRI scale.

Another study that used the SART to investigate lapses of attention and everyday cognitive failures was conducted by Cheyne, Carriere, and Smilek (2006). Cheyne et al. sought to test the efficacy of a self-report measure (Attention-Related Cognitive Errors – ARCES) that was developed to assess "everyday performance failures arising directly or primarily from brief failures of sustained attention" (Cheyne et al., 2006, p578). The driving factor behind the need to design the new questionnaire was due to the belief that a popular questionnaire, the Cognitive Failure's Questionnaire (CFQ – Broadbent, Cooper, FitzGerald, and Parkes, 1982) was too non-specific in nature, and assessed factors beyond attention related failures. Their self-report questionnaire was designed specifically to assess errors from attentional lapses. This was in contrast to another measure, the Mindful Attention Awareness Scale (MAAS – Brown & Ryan, 2003), which was a scale that assessed the ability to maintain sustained awareness in everyday life (Cheyne et al., 2006)**.** Initial psychometric comparisons between the MAAS and ARCES showed that both measures were negatively correlated, as predicted by the authors. This suggested that these measures would be able to assess the differences between attentional lapses and the ability to maintain sustained awareness. Based on the results of a path analysis using the SART as a measure of attentional lapses and failures of attention, it was showed that the MAAS was associated with speeded reaction times. Speeded reaction times in the SART literature suggests that as attention begins to wane, participants begin to respond automatically to the repetitive stimuli indicating that attention is no longer directed to observing the stimuli, leading to waning attention. The ARCES was associated with SART errors whereby participants were unable to inhibit their automatic responses due to a failure in sustaining their attention. This failure to sustain attention, leading to task errors was attributed to the waning of attention. It was believed that as attention waned, participants would have a more difficult time to inhibit their automatic responses when the target occurs. Therefore, results from their study further support the notion that the SART could be a measure of both attentional lapses and failures of attention.

Further support for the SART paradigm was found by Smallwood, Beach, Schooler, and Handy (2008), when they investigated episodes of mind wandering induced during the SART by using simple thought probes that queried the state of mind that the participant was in, and the event-related potentials that were elicited by responding to the non-targets that preceded a target event. It was believed that the amplitude of the P300 component reflected the amount of attentional resources directed to a task. The greater amplitude was suggestive of greater attention. It was found that the P300 amplitude was greater in trials where observers were able to successfully withhold their responses as compared to trials where an error was mode. Furthermore, the P300 amplitude was found to be greater when observers reported, via thought probes, that they were on task, as compared to instances where they were inattentive to the task. Results from Smallwood et al. (2008) support the SART by showing greater P300 amplitudes in trials where observers were successful in withholding their response to the target suggesting a greater degree of endogenous attention, as compared to weaker P300 amplitudes in SART errors. Secondly, periods of inattention or mind wandering were reflected in weaker P300 amplitudes, suggesting that the endogenous nature of mind wandering could also potentially have an effect on task performance.

1.6. Disagreement between Views

There are currently two competing views in the vigilance literature: the mindlessness theory, and the resource theory. Some of points raised by resource theorist that attempted to refute the mindlessness theory of vigilance include:
i) similar subjective workload, ii) signal regularity, iii) signal salience.

A study conducted by Grier et al. (2003) attempted to investigate potential differences in subjective workload. Observers in the study either completed a traditional vigilance study lasting 50 minutes or the exact vigilance study but using the SART response scheme. Results from the study indicated no differences in subjective workload. The supposed mindless response to stimuli was rated similarly taxing as traditional vigilance tasks indicating that vigilance tasks using a SART response scheme was similarly taxing. Supporters of mindlessness do not deny the high task demands in their studies. It was noted that simple SART like tasks where observers had to maintain an acceptable level of performance when faced with boring and ‘low-demand’ activities were viewed as demanding (Manly et al., 2003; Wilkins, Shallice, & McCarthy, 1987). Furthermore, the notion of mindlessness does not indicate a low workload. Rather, the term ‘mindlessness’ focuses on the act of responding in a mindless fashion, and having to then inhibit the automatic behavior. Thus the high workload scores might be indicative of the difficulty in inhibiting the automatic behavior.

Secondly, a study by Helton et al. (2005) investigated the effects of signal regularity on task performance. Observers were either engaged in a traditional vigilance task lasting 40 minutes or the exact task, but employed the SART response scheme. Within each task, participants either encountered regular signal events or irregular signal events. Results indicated that regular signals led to a significantly lower detection rate, which was contrary to the findings of Manly et al., (1999). Helton et al. claimed that regardless of their results, the notion that observers were monitoring the temporal presentation of signals, casts doubt on the mindlessness theory. It was believed that if a task were truly mindless, participants would not be actively monitoring the stimuli, thus signal regularity should not have an effect on task performance. Yet, Manly et al. (1999) had previously argued that increased signal regularity allowed for exogenous support to the task, rather than taxing the endogenous attentional system as brought about by the SART.

Thirdly, Helton and Warm (2008) suggested that signal salience could play a role in task performance, based on earlier work by (See, Howe, Warm, & Dember,

1995) who reported that an increase in task difficulty would lead to greater performance decrement. Observers in their study were engaged in an abbreviated 12 minute traditional vigilance task response scheme, where the salience of the signals differed between the groups. Performance results from both groups were compared with observers self-reports of TUTs and energetic arousal. Results from their study indicated that observers in the low salience condition reported less TUT, but still reported greater detection failures This finding might be contrary to the mindlessness theory as it purports that the easier the task, the greater the chance of mindless behavior, and TUTs leading to errors in detection. Yet the basis of the mindlessness theory is not based on the salience and simplicity of the signals. Rather it is the act of responding to non-targets in an automatic way that leads to mindlessness rather than the act of observing targets (Manly et al., 1999).

The disagreement between the mindlessness theorist and resource theorist might be pinned to two key factors. Firstly, resource theorists have focused primarily on performance decrement with time on task, rather than an overall decrement, and attentional lapses. Secondly, resource theorists have often employed a signal detection paradigm in investigating their results. This disagreement is understandable and as stated previously, resource theorist have traditionally been focused on performance levels, as their work has been rooted primarily in the human performance community, and the need to understand the vigilance decrement is primary. Furthermore, applying a signal detection framework allows for focused research into specific psychophysical factors that affect the vigilance performance. In contrast, mindlessness theorists have come primarily from the clinical and cognitive neuroscience fields of psychology. Their focus has been toward lapses in attention, as evidenced by the SART (Robertson et al., 1997). More recently their focus has broadened to include task unrelated thoughts and mindlessness as a result of automatic behaviors induced by the task that were no longer being modulated by an endogenous attentional system. Mindlessness theorists have used subjective scales that require introspection, much to the ire of pure experimentalist. Jack and Roepstorff (2003) argued that the unwillingness of cognitive scientists to investigate introspective studies will not aid in the validation or invalidation of these measures leading to no progress. It has been noted that mind wandering is a very common phenomenon and would account for some portion of an individual's time when performing a cognitive task (Smallwood & Schooler, 2006). It might then be suggested that these are two different processes with regard to vigilance, with resource theorists focused on the vigilance decrement, while the mindlessness theorists focused on the lapses of attention.

1.7. Vigilance and Driving

Of interest to the current line of research is the effect of having to engage in a boring monotonous drive on driver performance. Schmidt et al. (2001) observed a performance decrement in drivers who were engaged in a three hour monotonous drive. Using an auditory odd-ball reaction time task as a measure of performance, it was reported that drivers showed a marked linear increase in reaction time to the target tone as the duration of the task increased. Interestingly, participant reported subjective states of sleepiness, inattention, and monotony showed a significant quadratic trend over time, suggesting that observers might not be fully aware of their current state and might overestimate their ability even when showing signs of slowed reaction times. Papadelis et al. (2007) investigated possible neurophysiological measures that could serve as potential indicators of driver fatigue. Though not subjected to statistical analysis, visual interpretation of EEG measurements showed sudden bursts of alpha wave activity prior to a driving error, especially in observers who reported high levels of fatigue. EOG measurements also showed increased duration of eye blinks in the later stages of the driving task. The literature has been consistent with regard to changes in behavioral and neurophysiological measures when drivers were engaged in a monotonous drive, thus implying that driving in a monotonous environment could be compared to a vigilance task (Thiffault & Bergeron, 2003). Whether monotonous driving might be viewed as a mindless automatic task or a demanding cognitive task is yet to be seen.

1.8. Talking on Cell Phones While Driving

Driving has become an integral part of life in most developed nations, and with sufficient practice, most individuals do not think about their driving behaviors. As of this writing, there were an estimated, 266 million cell phones subscribers in the United States alone (CITA, 2008). A report by the Department of Transportation (Glassbrenner, 2005) noted that at any given daylight moment, there were approximately 974,000 vehicles on the road that had drivers actively using a hand-held cell phone. The report did not account for drivers using hands-free cell phones, thus masking a potentially larger number of users. Furthermore this number would have obviously increased since then as the report stated a 13% increase from an earlier survey. In one of the most influential and often-cited research to date that investigated the relationship between cell phone usage and motor vehicle accidents was a paper written by Redelmeier and Tibshirabi (1997). They reported that drivers who engaged in cell phone conversations while driving were four times more likely to get into accidents than drivers who were not. Furthermore, this statistic did not differ between the demographic of cell phone users. In an attempt to minimize the risk of vehicular accidents, some legislative bodies have banned the use of hand-held cell phones while driving. Yet, research into distracted driving suggests that regardless of the method drivers' use, hands-free or hand-held, the risk of getting into an accident was no different.

1.8.1 Effects of Cell Phone Conversations on Simulated Driving Tasks

Strayer and Johnston (2001) had participants engage in a tracking task while either conversing with a researcher via a cell phone, or passively listening to a radio broadcast. Results showed that observers in the cell phone task were more likely to miss target signals in the tracking task. Reaction times to targets were slower in the cell phone condition as compared to the passive radio broadcast condition. In a separate study, they also found greater tracking error in observers who were engaged in a speech generation task when compared to observers in a speech shadowing task. Combined results from the study indicated that performance was equally poor when observers were engaged in a conversation regardless of whether the conversation took place using a hand-free or hand-held device. Secondly, word generation tasks rather than simple speech shadowing showed greater task errors. These results further implicated the negative effect of cell phone conversations. Furthermore, research by Atchley and Dressel (2004) investigated the effects of cell phone conversations on an individual's functional field of view (FFOV). The FFOV was defined as a "restriction of a person's field of view imposed by limits in his or her visual information processing capacity" (Atchley & Dressel, 2004, p. 665). Results from their study indicated that when observers were engaged in a cell phone conversation, their already limited field of view was further minimized, as compared to situations where no conversations took place. This shrinking of the FFOV implied that distracted drivers might miss events that occurred in their peripheral field of view. It was also reported that when observers were engaged in both tasks simultaneously, suboptimal results were found for both the conversation task, and the FFOV task.

1.8.2. Effects of Cell Phone Conversations on Driver Performance in Driving Simulators

Besides simulated driving tasks, researchers have sought to further improve the ecological validity of their studies by using driving simulator studies. Rakauskas, Gugerty, and Ward (2003) had observers in their study driving in a small simulated environment that had stationary vehicles parked on the side of the road. The research design had the driver engaged in no conversation, easy, and difficult cell phone conversation, and furthermore, had to react to various hazardous events. Results from their study showed that when drivers were engaged in conversation, speed maintenance was significantly impaired, suggesting that drivers had to constantly adjust their speed when engaged in a conversation. Though there were no significant results with respect to hazardous road events, it was believed that the lack of significance could have been attributed to the lack of power, possible practice effects that were likely to occur in an enclosed course.

Kubose et al. (2006) conducted two experiments to investigate the potential difference in driving performance between speech production and speech comprehension while driving in a simulated rural highway. Results from their study indicated that concurrent speech production or speech comprehension while driving had an effect on a driver's ability to maintain constant velocity when compared to drivers who were not engaged in conversation. They also found that under dual-task conditions, drivers had a larger headway variability and headway distance from a lead car, suggesting a difficulty in the ability to make judgments on distance. Results from the literature continue to suggest that under dual-task conditions, drivers have a greater difficulty in vehicle handling as compared to drivers who were engaged only in the driving task.

1.9. Dual Tasking and Vigilance

As previously discussed, the effect of having to perform two tasks (e.g. engaging in a cell phone conversation while driving) concurrently had significant negative effects on task performance (e.g. Atchley & Dressel, 2004; Strayer & Johnston, 2001). Dual task studies in vigilance research have focused on the effects of having to perform vigilance tasks on two different modalities simultaneously. In such tasks, it was noted that having to monitor multiple sources would degrade performance efficiency in both modalities (Warm & Dember, 1998). The aim of the current research project was to investigate the effect of a non vigilance secondary task i.e. cell phone conversation, on the primary vigilance task performance. The auditory portion of the task was not deemed to be a vigilance task where observers had to respond to a target stimulus. Rather, it was deemed to be a potential energizer to the observer engaged in the visual vigilance tasks used in our studies.

1.10. Purpose

The need to understand lapses in attention in human behavior is of utmost importance. Research into vigilance has lead to development of potential methods that could minimize the monotony of the task with the intention of minimizing performance decrements. Based on the current understanding of human performance of vigilance tasks, various methods of potentially improving performance have been suggested. Minimizing background event rates would potentially improve detection performance (Galinsky et al., 1990). The provision of feedback to the observer would also improve performance (Baker, 1962). The presentation of false signals could also serve to improve task performance by piquing overall arousal temporarily (Parasuraman, 1984). These methods are likely to improve performance, based on previous research findings, but these factors are seldom under the control of the observer.

The current research focuses on what might possibly be a factor that could improve performance on a vigilance task, both in a laboratory setting, and in an applied setting by using driving simulator data. Drivers often state that besides using stimulants such as caffeine, engaging in a cell phone conversation was believed to also help keep them awake. In this study, two different vigilance tasks were used, a traditional vigilance paradigm where observers had to respond only to the target stimuli (MVT) and the SART response paradigm where observers had to withhold response to the target stimuli. This would constitute the laboratory component of the vigilance task. The results from the laboratory component would be used to support potential findings from the driving simulator data.

It was hypothesized that a vigilance decrement would occur in both, traditional and SART response paradigms, but the detrimental effects on performance would be modulated with the presence of the secondary conversation task. It was hypothesized that driving performance might be improved with the introduction of a secondary conversational task. It was believed that the presence of a conversation might break the monotony of the drive thus leading to better driving performance.

2. Rationale for Analysis

Analysis of data focused only on the differences in the first and last blocks of the experiments. The rationale for such a decision was because the study sought to investigate overall performance decrement, rather than a gradual performance decrement. In instances where there were no significant interactions between the independent variables, follow up analyses were still conducted. The rationale for such a decision was two-fold. Firstly this thesis sought to compare methodological differences under the pretext of an exploratory investigation to measure sustained attention. Hence in some cases, even with the lack of significant main effects and/or interaction effects, further analyses were still conducted, lest a Type II error was made. Secondly, the applied nature of this research would imply that all significant results might have an applied impact. Thus failure to report and account for these masked results might weaken the case of an applied study.

3. Experiment One – SART

3.1. Methods

3.1.1. Participants

One hundred and twenty students from the University of Kansas undergraduate research pool, and students from an upper level psychology class, 73 men and 47 women, averaging 19.84 years (*SD* = 2.28), participated for course credit. Participants were randomly assigned into one of three task conditions: no conversation, 25 men and 15 women; full conversation, 24 men and 16 women; late conversation, 24 men and 16 women. Almost all participants reported English as their primary/native language. A Snellen Visual Acuity Chart was used to test for visual acuity in participants. All participants had normal to corrected visual acuity of 20/32 or better.

3.1.2. Materials and apparatus

SART. A modified version of the Sustained Attention Response Task (SART) developed by Robertson et al (1997) was programmed on E-prime (Psychology Software Tools Inc., Pittsburgh, PA) and ran on a Windows 2000 PC. Stimuli were presented on a standard 15 inch CRT monitor with a refresh rate of 75Hz, and at a screen resolution of 640x480 pixels. Viewing distance was 40cm from screen to participant. No restrictions were placed on participant's range of motion so long as one hand was kept on the spacebar at all times. The SART was designed by Robertson et al. (1997) as a means of comparing differences in sustained attention between traumatic brain injury (TBI) patients and normal controls.

The SART lasted a total of 21.5minutes, with 5 periods of watch each lasting 4.3 minutes. Within each block, 225 single digits (Numbers 1 through 9) were presented in a pseudorandom order (randomization was controlled by E-Prime). Each digit was presented for 250ms, and was followed by a structural mask (Numbers 1-9 overlaid above each other). The mask duration lasted 900ms. Participants were told to respond to all digits with a keypress on the spacebar, and withhold their keypress when the number "3" (target) was presented. There was a total of 25 targets (11% of trials) and 200 non-targets (numbers 1, 2, & 4-9), in each timing block. All stimulus was presented in the center of the monitor (See Figure 1 for example of SART).To ensure that targets did not appear back to back, E-prime was programmed to randomize the presentation in a manner that would present at least 2 non-targets before presenting a target again. In total, there were a total of 1125 digits, of which 125 were targets. E-prime was programmed to collect accuracy scores (hits, misses, false alarms) and reaction time data. The digits were presented centrally on the monitor with white numerical digits on a black background. Luminance of these presentations was noted to be 100 cd/m^2. The font of choice was Courier New, and the font size of the stimuli was 24. The experimenter remained in the same room for the duration of the study to ensure that the participant was actively involved with the task as much as possible. The experimenter did not converse with the participant at any time during the entire duration of the experiment. Participants were told to turn off their cell phones, and no task duration information was made known to them except that the task would last no longer than 30 minutes, as stated in the consent form.

Due to the unique presentation procedure of the SART, definitions of hits, misses and false alarms differ from traditional vigilance task. Hits were defined as a purposeful withholding of response when the target appears. Misses were defined as a lack of response to a non-target, this was also known as an error of omission. A false alarm was defined as a keypress response to the target when the correct response would be to withhold the keypress, also known as an error of commission. Reaction time was scored and logged as a successful keypress when a stimulus was presented and the participant responds before the presentation of the next stimuli. Therefore a response during the mask would still count as a response, false alarm keypresses were also factored into the reaction time calculations. For the purposes of our study, we did not investigate hit rate. It was noted that false alarms were inversely proportional to hits, thus one would expect that results for either dependant measure would be similar.

Conversation Task. 71 words were chosen from the ANEW wordlist (Bradley, & Lang, 1999) to serve as stimuli for the conversational task. (See Appendix A for wordlist). The words were noted to be of low valance and of positive affect. The words were presented randomly via E-prime, and the wordlist was programmed not to repeat any given word until the current wordlist completed its cycle; following which another randomized wordlist would be generated. Participants were told to free associate their responses to the stimulus word. There were no responses that were deemed correct or incorrect. Words were presented one at a time mimicking a back and forth conversational style, with the participant giving a one word response to the word that was presented. If participants failed to respond due to an inability to comprehend the word or the inability to produce a word within the interstimulus-interval (ISI), the software would automatically continue with the next word. The presentation of words was controlled by E-prime with a constant ISI of four seconds. This rate of presentation was decided upon through informal pilot testing where it was noted that four seconds would be sufficient for participants to generate a response. The prerecorded words were spoken by a native-speaker of English. The computer responsible for the conversation task was housed in a separate room from the participant.

Panasonic KX-TG5422 cordless telephones were used to facilitate communication between the participant and the computer generating the prerecorded words. One cordless phone would be placed in the room with the computer running the conversation. This phone would be responsible for transmitting the word to the participant who was seated in a separate room. The auditory stimuli were always presented to the right ear of the participant by means of a hands free kit. All verbal responses from the participant were spoken into the attached microphone from the hands free kit. Participants in the no conversation task did not engage in any conversation throughout the task. Participants in the full conversation task were engaged in both the SART and conversation task for the entire duration. Participants in the late conversation task engaged in the conversational task only in the last time block, 17.2 minutes into the task.

3.1.3. Procedure

Upon completion of a consent form and demographic datasheet that queried cell phone usage behavior, participants were given an eye test to check visual acuity. Following which, participants engaged in a practice trial that lasted approximately 1 minute. During this trial a total of six targets (number “3”) were presented. At the end of the practice block, the experimenter would ask all participants to equip the hands free kit by placing the headphone over their right ear, and extending the boom microphone to their mouth. The experimenter informed all participants that they may or may not be engaged in a conversational task. Once participants understood the task instructions and were ready to begin the study, the participant would say ‘Go’ into the microphone. This would remotely activate the computer responsible for the conversation task. Upon the completion of the experiment, participants completed an electronic version of the NASA-TLX (Hart & Staveland, 1988). The entire experiment was conducted in a darkened room with no ambient light source. Luminance levels were 0 cd/m^2 as measured by a Minolta LS110 Luminance Meter.

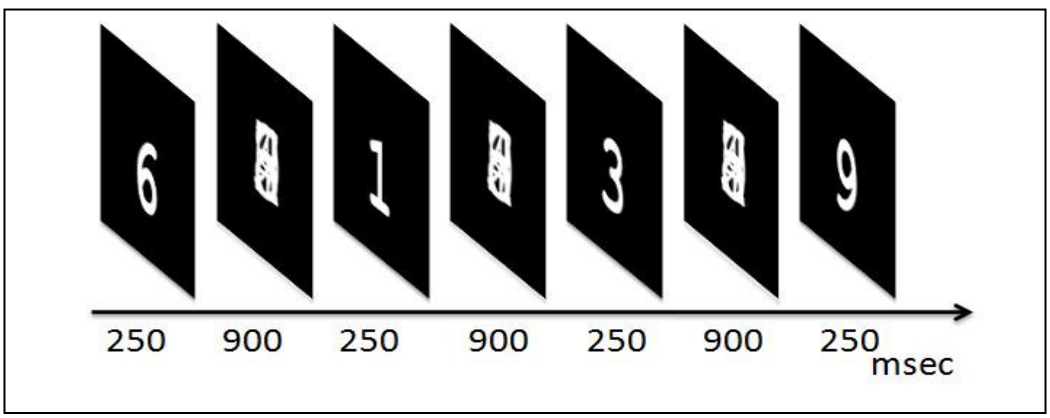

Figure 1. Example of SART progression

3.2. Results

Three separate 3(Task type)x2(Watch period) Mixed-Analysis of Variance (ANOVA) were conducted with Task type as the between subjects factor and Watch period as within subjects factor for the dependent variables.

3.2.1. Reaction Time

Reaction time scores were recorded whenever a participant made a key press to a stimulus before the presentation of subsequent stimuli. Therefore, a response that occurred during the masking period, but before the next stimuli was counted as a key press. False alarm key presses where participants failed to withhold their response to the target were not factored into the overall reaction time scores.

Results indicated no significant main effect for task type, $F(2, 117) = 1.22$, $p = 0.29$, $MSE = 9965.39$; partial-$\eta^2 = .02$ and watch period $F(1, 117) = 0.47$, $p = 0.49$, $MSE = 2809.05$; partial-$\eta^2 = .004$. It was noted that there was a significant interaction between task type and watch period, $F(2, 117) = 12.92$, $p < .001$, $MSE = 2809.05$; partial-$\eta^2 = .18$.

Table 1: Mean values (and standard deviations) of reaction times (in milliseconds) for SART

Task Type	Watch Period 1	Watch Period 2	*Average*
No Conversation	278.58 (59.82)	268.22 (74.38)	*273.40 (67.10)*
Full conversation	315.85 (93.66)	271.65 (99.03)	*293.75 (96.35)*
Late conversation	275.41 (42.67)	315.86 (93.88)	*295.64 (68.28)*
Average	*289.95 (65.38)*	*285.24 (80.09)*	

Planned comparisons with least significant difference adjustments revealed that there was a significant decrease in reaction time for the full conversation task between the first and second watch periods $F(1, 117) = 13.91$, $p < .001$. Participants in the full conversation task were getting faster in their responses over time. An opposite effect was observed for the late conversation task, where reaction times significantly increased between the first and second watch periods $F(1, 117) = 11.65$, $p < 001$, suggesting that participants were slowing down in their responses. There was no significant difference between the first and second watch period for the no conversation group.

Simple effects analysis using univariate ANOVAs on both watch periods showed significant differences within the first watch period, $F(2, 117) = 4.28$, $p < .05$, $MSE = 4724.24$; partial-$\eta^2 = .07$ and the second watch period. $F(2, 117) = 3.51$, $p < .05$, $MSE = 8050.66$; partial-$\eta^2 = .06$. Pairwise comparisons between task types within each watch period were used to further investigate the differences. In the first watch period, reaction times for the full conversation task were significantly slower than the late conversation task, $F(1, 117) = 6.93$, $p < .01$, and the no conversation task $F(1, 117) = 5.88$, $p < .05$. There was no significant difference between the no conversation task and late conversation task. There were also significant changes in reaction time differences between some task types in the second watch period. During this period, the reaction times for the late conversation task were significantly slower than the full conversation task $F(1, 117) = 5.12$, $p < .05$, and the no conversation task $F(1, 117) = 6.11$, $p < .05$. There was no significant difference between the no conversation task, and the full conversation task.

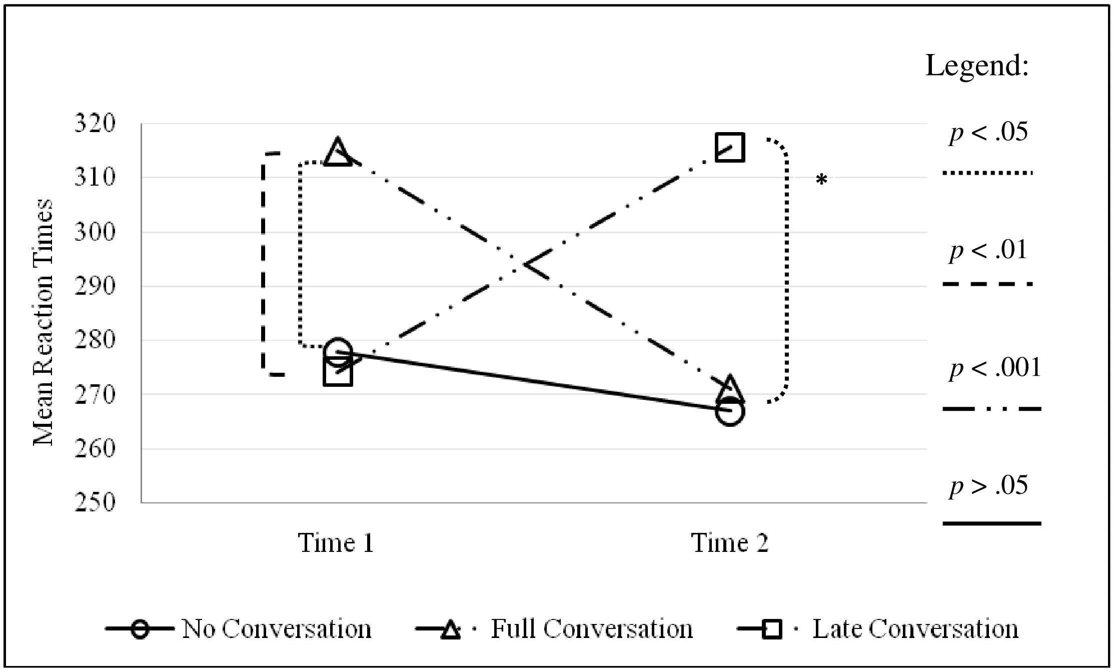

Figure 2. Mean Reaction Times for SART over Two Time Blocks

3.2.2. Misses

Misses were scored as periods of inactivity where participants failed to respond to the stimuli during the task. This was also known as an error of omission. This would include failures to respond to non-target stimuli or the masking image within the specified inter-stimulus duration.

Results indicated a significant main effect for task type, $F(2, 117) = 7.89$, $p < .001$, $MSE = 299.96$; partial-$\eta^2 = .12$, where full the conversation task ($M = 20.57$) had more misses than the no conversation task ($M = 9.7$) and the late conversation task ($M = 15.68$). There was also a significant main effect for watch period $F(1, 117) = 20.93$, $p < .001$, $MSE = 129.02$; partial-$\eta^2 = .15$. It was noted that there were more misses within the first watch period (*M* = 11.96) as compared to the second watch period (*M* = 18.67). More importantly, it was noted that there was a significant interaction between task type and watch period, $F(2, 117) = 24.76$, $p < .001$, $MSE = 129.02$; partial-$\eta^2 = .29$.

Table 2: Mean values (and standard deviations) of Misses for SART

Task Type	Watch Period 1	Watch Period 2	*Average*
No Conversation	6.25 (7.35)	13.15 (12.28)	*9.7 (9.82)*
Full conversation	23.58 (21.23)	17.55 (15.72)	*20.57 (18.48)*
Late conversation	6.05 (4.65)	25.3 (19.04)	*15.68 (11.85)*
Average	*11.96 (11.08)*	*18.67 (15.68)*	

Planned comparisons with least significant difference adjustments revealed that there were significant changes in mean number of misses for all task types over the course of the experiment. There was a significant increase in misses for the no conversation task, $F(1, 117) = 7.38$, $p < .01$, and the late conversation task $F(1, 117) = 57.44$, $p<.001$. Conversely, an opposite effect was observed for the full conversation task, where mean number of misses decreased over time $F(1, 117) = 5.63$, $p < .05$.

Simple effect analysis using univariate ANOVAs on both watch periods indicated significant differences in the first watch period $F(2, 117) = 23.08$, $p < .001$, $MSE = 175.46$; partial-$\eta^2 = .28$ and second watch period $F(2, 117) = 5.97$, $p < .01$, *MSE* = 253.52; partial-η^2 = .09. Pairwise comparisons between task types within each watch period were used to further investigate the differences. There were significantly less misses in the first watch period for the no conversation task $F(1, 117) = 34.23$, p <. 001, and late conversation task $F(1, 117) = 35.01$, $p < .001$, when compared to the full conversation task. There were no significant differences between the late conversation task and no conversation task in this period. Results for the second watch period were different from the first watch period, whereby it was observed that the no conversation task $F(1, 117) = 11.65$, $p < .001$, and the full conversation task $F(1, 117) = 35.01$, $p < .001$ made less misses than the late conversation task. There were no significant differences between the no conversation task and full conversation task in the second watch period.

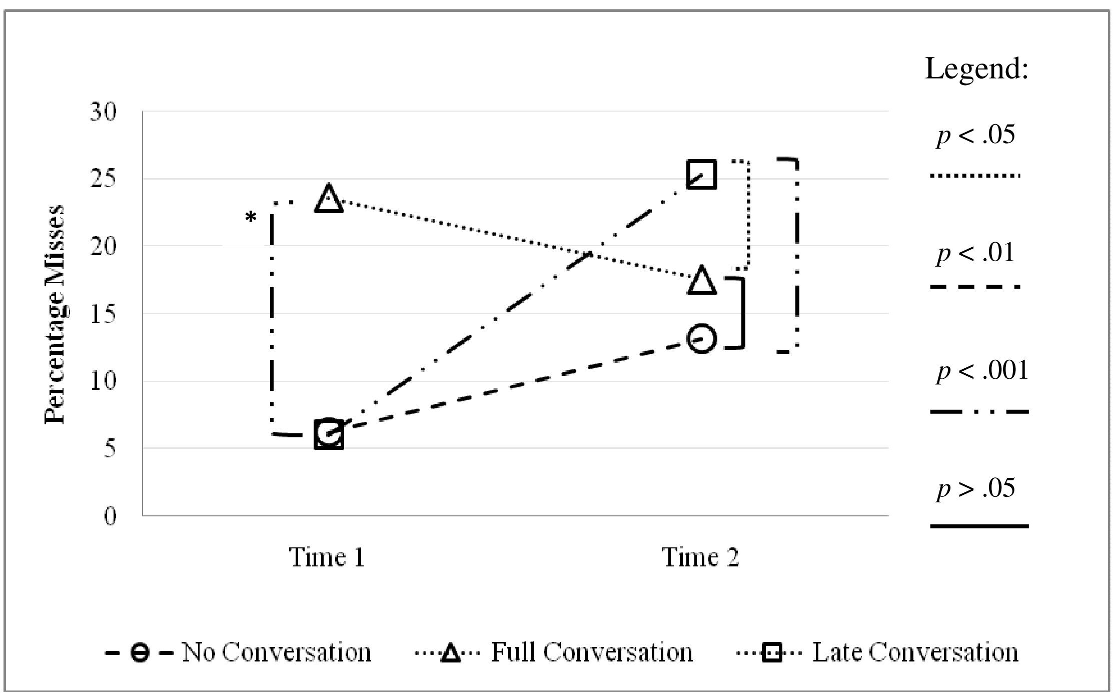

Figure 3. Percentage Miss for SART over Two Time Blocks

3.2.3. False Alarms

False alarms were scored as the failure to inhibit a response to the target stimuli. This was also known as an error of commission. Scoring for false alarms included responses made to the target stimuli during the masking period i.e. during the inter-stimulus interval.

Results indicated no significant main effect for task type, $F(2, 117) = 1.95$, $p = .15$, $MSE = 36.31$; partial-$\eta^2 = .03$. The main effect for watch period was significant $F(1, 117) = 15.42$, $p < .001$, $MSE = 10.49$; partial-$\eta^2 = .12$. There were more misses in the first watch period ($M = 17.37$) as compared to the second watch period ($M = 28.54$). There was no significant interaction between task type and watch period, $F(2, 117) = 1.8$, $p = .17$, $MSE = 10.49$; partial-$\eta^2 = .03$.

Table 3: Mean values (and standard deviations) of False Alarms for SART

Task Type	Watch Period 1	Watch Period 2	*Average*
No Conversation	16.15 (5.72)	18.45 (5.45)	*17.3 (5.59)*
Full conversation	18.13 (4.89)	20.22 (3.66)	*19.18 (4.28)*
Late conversation	17.85 (4.47)	18.38 (4.54)	*18.12 (4.51)*
Average	*17.38 (5.08)*	*19.02 (4.65)*	

Planned comparisons with least significant difference adjustments revealed that there was a significant increase in false alarms for the no conversation task $F(1, 117) = 10.09$, $p < .01$, and full conversation task $F(1, 117) = 8.41$, $p < .01$, over the duration of the study. There was no significant change in false alarms for the late conversation group.

Analysis of simple effects using univariate ANOVAs on both watch periods indicated no significant differences between groups in the first watch period $F(2, 117) = 1.79$, $p = .17$, $MSE = 25.56$; partial-$\eta^2 = .03$ and second watch period $F(2, 117) = 2.06$, $p = .131$, $MSE = 21.23$; partial-$\eta^2 = .03$. No further analyses were conducted due low probability of any potential effects.

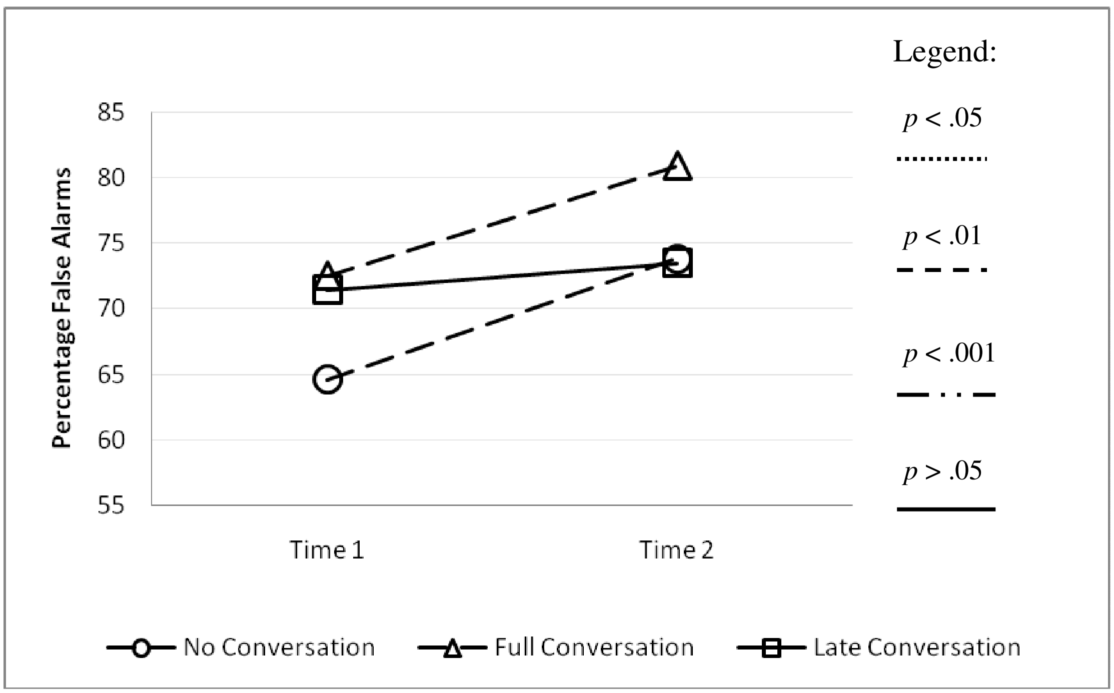

Figure 4. Percentage False Alarms for SART over Two Time Blocks

3.2.4. NASA-TLX Ratings

A One-way ANOVA was used to measure potential differences of participant's perceived task load. Results indicated that all three task types rated (*M* =

74.36, *SD* = 14.67) the SART as mentally demanding, and there were no significant differences between tasks. $F(2, 117) = 1.49$, $p = .23$.

3.3. Discussion

The SART was designed as a means of investigating failures and lapses of everyday attention (Manly, 1999; Robertson et al., 1997), and it does so by having participants engage in a seemingly mindless, monotonous and repetitive task that required little control once the participant has sufficiently learned the task and entered into a state of automatic responses (Schneider & Shiffrin, 1977). Results from the current experiment support results from previous studies that applied the SART methodology (Manly et al., 1999, 2002; Robertson et al., 1997).

3.3.1. Reaction Time

The SART predicts that as attention to the task begins to wane, participant responses to the non-target stimuli would become faster due to the setting in of automaticity. Reaction times in the first watch period for the full conversation task were higher than the no conversation task, and the late conversation task. This was likely due to the early onset of a conversation. Participants in full conversation task had to timeshare resources in an attempt to cope with both the conversation task and the SART. It has to be noted that based on SART methodology, an increase in reaction time also suggests a greater amount of attention focused on the task, and this might be the case for the full conversation task as participants might be attempting to learn how to successfully timeshare resources.

During the second watch period, an interesting reversal of reaction times occurred between the late conversation task and full conversation task. In this period, reaction times for the full conversation task were significantly lower than the previous watch period. This could be due to the participants having learned to timeshare thus leading them to behave in a manner similar to participants in the no conversation task, and as predicted by the SART would lead to faster reaction time. On the other hand, with the onset of a conversation late in the task, participants in the late conversation group were showing reaction times similar to that of participants in full conversation tasks during the first watch period. This increase in reaction time was likely due to participants having to learn how best to timeshare resources, which was similar to what was observed in participants in the full conversation task in the first watch period.

According to SART methodology faster reaction times signifies waning attention. When participants in the full conversation task had learned to timeshare resources, their ability to timeshare puts them on equal standing with participants who were not engaged in any conversation for a prolonged period of time. In other words, continuous conversation while engaged in a vigilance task does not improve performance substantially. Instead it only raises performance to a level relative to an individual who experienced a decrease in vigilance.

3.3.2. Misses

Misses in SART were rarely discussed due to the low rate of misses made by participants. In the current study misses were used as a means of investigating the ability to maintain attention to the SART. Misses were generally low in the first watch period; approximately 3% for both the no conversation and late conversation tasks respectively. A substantial difference was observed in the full conversation task where participants made significantly more misses than tasks that did not have any conversational element at the onset of the task. Participants in the full conversation task were faced with having to learn how to dual task. Thus in an attempt to timeshare resources, participants in the full conversation task made more misses to the stimuli presented. This finding was not surprising due to the effects of dual-task interference on task performance.

There were significant changes in the results of the second watch period. Firstly, the large increase in the misses for the late conversation task between the first watch period and second watch period could likely be attributed to the late onset of a conversation. Participants in this group were unlikely to have learned how to effectively timeshare resources between the SART and conversation task, resulting in the substantial increase in misses. This was similar to the behavior of participants in the full conversation task in the first watch period. Secondly, the significant decrease in misses for the full conversation task could be attributed to an improvement in the ability to timeshare resources. Yet this improvement was only comparable to participants in the no conversation task. Under the no conversation task, the significant increase in misses in the second watch period could be indicative of a decrease of attention to the task due to boredom. Thus this result may again suggest that performance might increase while performing a prolonged secondary task, the level of performance was only equivalent to a disinterested individual. This result supports the hypothesis that a secondary task may improve performance, but it also suggests that the benefits of a secondary task do not outweigh its costs.

3.3.3. False Alarms

The significant overall increase in false alarms in the no conversation task and full conversation task was indicative of waning attention. Such a result was not unexpected due to the nature of vigilance tasks. The lack of a significant increase in false alarms in the second watch period for the late conversation task could be a point of interest. Based on the hypothesis, where the presence of a secondary task might improve performance, the late conversation task group did not show a significant decrease in false alarms between the first and second watch periods. This may suggest that the presence of a late conversation might mitigate the effects of boredom. But a more parsimonious view would be the fact that participants in the late conversation group were observed to have a significant increase in misses. In a task such as the SART where the target “3” was embedded in the stream of stimuli, a significant increase in misses might also lead to a decrease or null effect on misses as participants failed to engage in the task completely.

3.3.4. *NASA-TLX Ratings*

Consistent with literature in vigilance research, participants in all task conditions rated the task in the upper scale of the NASA-TLX (Grier, et al., 2003), suggesting that the task was indeed mentally taxing.

Table 4. Summary table of SART results (effects in parenthesis)

Performance At Start			
	No Conversation	Full Conversation	Late Conversation
Reaction Time	Baseline	Slower (Dual Task)	Baseline
Misses	Baseline	More (Dual Task)	Baseline
False Alarm	Baseline	No Change (Dual Task?)	Baseline
Performance Over Time			
	No Conversation	Full Conversation	Late Conversation
Reaction Time	No Change (Floor)	Faster (Decreased Vigilance)	Slower (Dual Task)
Misses	More (Decreased Vigilance)	Less (Effective Timeshare)	More (Dual Task)
False Alarm	More (Decreased Vigilance)	More (Decreased Vigilance)	No Change (Time share?)

4. Experiment Two – MVT

4.1 Methods

4.1.1. Participants

Sixty students from the University of Kansas undergraduate research pool, and student from an upper level psychology class, 25 men and 35 women, averaging 19.32 years (*SD* = 2.06), participated for course credit. Participants were randomly assigned into one of three task conditions: no conversation, 7 men and 13 women; full conversation, 11 men and 9 women; late conversation, 7 men and 13 women. Almost all participants reported English as their primary/native language. A Snellen Visual Acuity Chart was used to test for visual acuity in participants. All participants had normal to corrected visual acuity of 20/32 or better.

4.1.2. Materials and apparatus

MVT. The MVT was designed by Hirshkowitz, De La Cueva, and Herman (1993). It was programmed on E-prime (Psychology Software Tools, Pittsburgh, PA) and ran on a Windows 2000 PC. Stimuli were presented on a standard 15 inch CRT monitor with a refresh rate of 75Hz, and at a screen resolution of 640x480 pixels. Viewing distance was 40cm from screen to participant. No restrictions were placed on participant's range of motion so long as one hand was kept on the spacebar at all times.

The MVT was used as a means of studying the effects of a vigilance type task on a clinical population that suffered from sleep disorders. The study lasted 30 minutes and was broken down into 5 periods of watch, each lasting 6 minutes. The 6 minutes blocks were subsequently broken down into 1 minute sub-blocks.). Within each 1 minute sub-block, a total of 2 targets and 6 non-targets would be presented. Within each 6 minute watch period, there was a total of 12 targets (25% of trials) and 36 non targets. In total, 60 targets and 180 non target stimuli were presented. Within each sub-block, a geometric shaped "H" (non-target) or a geometric shaped "I" (target) was presented in random order with the presentation duration of either target or non-target set at 250ms. Participants were required to make a keypress response as quickly as possible upon presentation of the target, and withhold responding to a non-target. Following which a masking grid with durations lasting between 4 to 11 seconds would appear before the presentation of the next shape. The duration of the masking grid was random to minimize possible temporal expectancy from the participants. The size of the grid and geometric shapes was set at 320x320 pixels, and was presented in the center of the monitor. (See Figure 2 for an example of the MVT). The luminance of the stimuli was noted to be 104.5 cd/m^2. The experimenter remained in the same room for the duration of the study to ensure that the participant was actively involved with the task as much as possible. The experimenter did not converse with the participant at any time during the entire duration of the experiment. Participants were told to turn off their cell phones, and no task duration information was made known to them except that the task would last no longer than 35 minutes, as stated in the consent form.

Conversation Task. The conversation task and apparatus in this experiment was similar to the first experiment, including the wordlist and timing conditions used for the conversation task. The only difference in this experiment was the onset for the presentation of words in the late conversation task. For this experiment, the late conversation began 18 minutes into the task.

4.1.3. Procedure

The procedure for the second experiment was similar to the first experiment. The only difference was that participants engaged in a practice trial that lasted 2 minutes, with 4 presentations of the target stimuli. The entire experiment was conducted in a darkened room with no ambient light source. Luminance levels were 0 cd/m^2 as measured by a Minolta LS110 Luminance Meter.

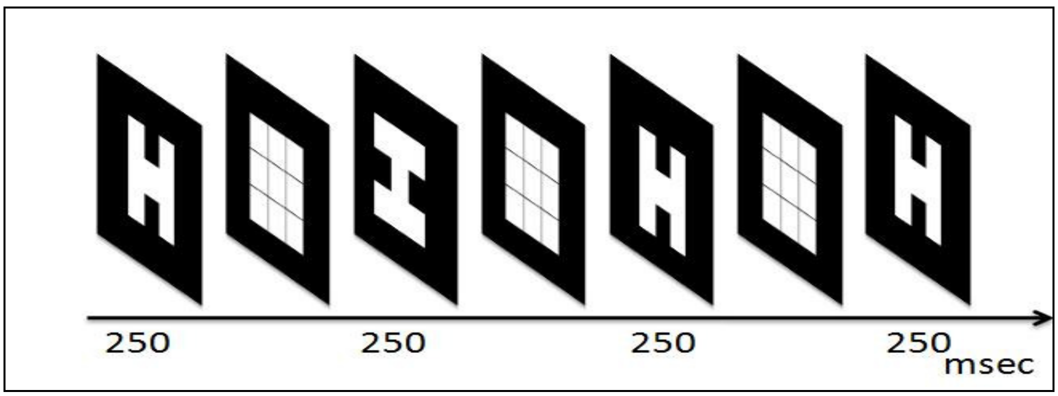

Figure 5. Example of the MVT progression

4.2. Results

Three separate 3(Task type)x2(Watch period) Mixed-ANOVA were conducted with Task type as between subjects factor and Watch period as within subjects factor for the dependent variables.

4.2.1. Reaction Time

Reaction time for the MVT was scored similarly to the SART. A key press made by the participant during the experiment was scored. Therefore, only responses to the target was recorded. Responses made during the presentation of the masking stimuli were also included.

Results indicated a significant main effect for task type, $F(2, 57) = 3.64$, p <.05, $MSE = 12388.14$; partial-$\eta^2 = .11$. Reaction times for the full conversation task was longer ($M = 628.84$) than the no conversation task ($M = 563.77$) and the late conversation task ($M = 582.02$). There was also a significant main effect for watch period $F(1, 57) = 20.33$, $p < .001$, $MSE = 4326.93$; partial-$\eta^2 = .26$. Reaction times became significantly slower between the first watch period ($M = 564.67$) and the second watch period ($M = 618.42$). There was no significant interaction between task type and watch period, $F(2, 57) = 1.49$, $p = 0.23$, $MSE = 4326.93$; partial-$\eta^2 = .05$.

Table 5: Mean values (and standard deviations) reaction times (in milliseconds) for MVT

Task Type	Watch Period 1	Watch Period 2	*Average*
No Conversation	530.86 (77.7)	596.68 (93.48)	*563.77 (85.59)*
Full conversation	616.59 (88.28)	641.09 (101.47)	*628.84 (94.88)*
Late conversation	546.56 (102.78)	617.49 (81.95)	*582.02 (92.37)*
Average	*564.67 (96.28)*	*618.42 (92.3)*	

Planned comparisons with least significant difference adjustments revealed that reaction time increased significantly for the no conversation task, $F(1, 57) = 10.01$, $p < .01$ and the late conversation task $F(1, 57) = 11.63$, $p < .001$, over the course of the study. There was no significant change in reaction time for the full conversation task.

Simple effects analysis using univariate ANOVA between tasks on watch periods indicated that there was significant difference between task types in the first watch period, $F(2, 57) = 5.13$, $p < .01$, $MSE = 8131.89$; partial-$\eta^2 = .15$, but none in the second watch period $F(2, 57) = 1.15$, $p = .32$, $MSE = 8583.85$; partial-$\eta^2 = .04$. Pairwise comparisons between task types in the first watch period were conducted to investigate the differences. Results indicated that reaction times were slower in the first watch period for the full conversation task as compared to the full conversation task, $F(2, 57) = 9.04$, $p < .01$, and the late conversation task $F(2, 57) = 6.03$, p < .05. There was no significant difference in reaction time between the late conversation ask and the no conversation task.

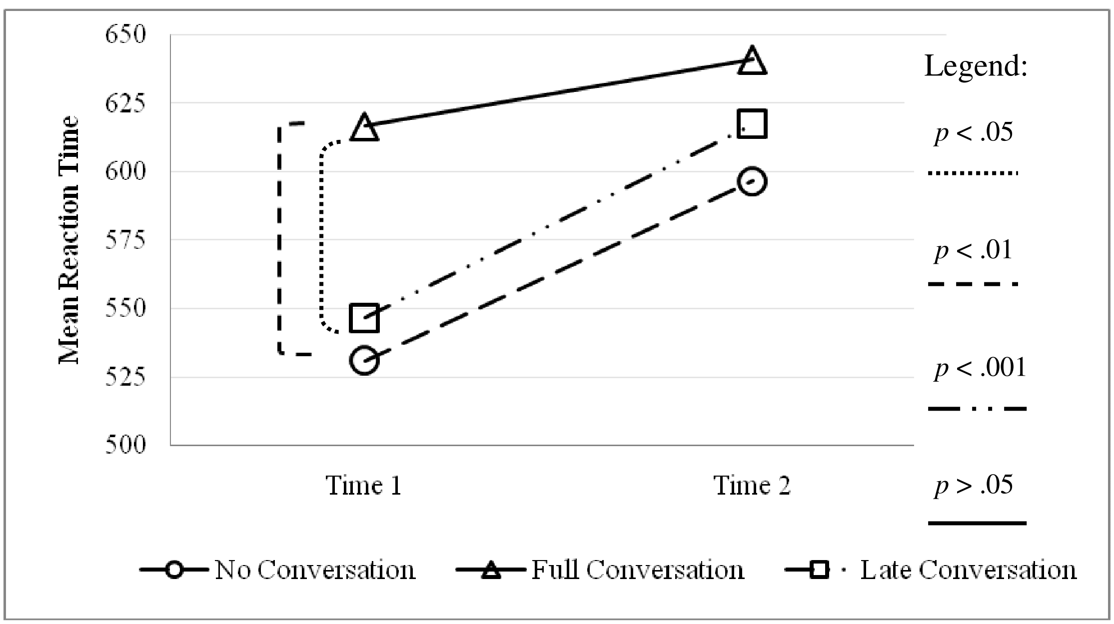

Figure 6. Mean Reaction Times for MVT over Two Time Blocks

4.2.1. Misses

Misses in the MVT were scored as instances where participants failed to make a response upon presentation of the target stimuli within the inter stimulus interval.

Table 6: Mean values (and standard deviations) misses for MVT

Task Type	Watch Period 1	Watch Period 2	*Average*
No Conversation	0.15 (0.37)	0.55 (0.83)	*0.35 (0.60)*
Full conversation	0.65 (0.88)	0.80 (1.15)	*0.73 (1.02)*
Late conversation	0.20 (0.69)	0.30 (0.73)	*0.25 (0.71)*
Average	*0.33 (0.65)*	*0.55 (0.90)*	

Results of misses in the MVT were close to ceiling. This has been observed in most vigilance tasks paradigms that used a traditional response paradigm as noted by Helton and Warm (2008). The overall miss rate was 2.75% for the first watch period, and 4.58% in the second watch period. Results of analysis indicated a significant main effect for task type $F(2, 57) = 3.32$, $p < .05$, $MSE = 0.76$; partial-$\eta^2 = .61$, where the full conversation task ($M = 0.65$) made misses than the no conversation task ($M = 0.15$) and the late conversation task ($M = 0.20$). There was no main effect of watch period nor was there an interaction between watch period and task type.

Simple effect analysis using univariate ANOVA between tasks on watch periods indicated that there was significant difference between task types in the first watch period, $F(2, 57) = 3.28$, $p < .05$, $MSE = 0.46$; partial-$\eta^2 = .60$. There was no significant difference between task types in the second watch period. $F(2, 57) = 1.47$, $p = .24$, $MSE = 0.42$; partial-$\eta^2 = .05$. Follow up pairwise comparisons were conducted on first watch period to investigate differences between the tasks. Results indicated that participants in the full conversation task made significantly more misses than participants in the no conversation task $F(2, 57) = 5.41$, $p < .05$, and the late conversation task $F(2, 57) = 4.39$, $p < .05$. There was no significant difference between the no conversation task and late conversation task in this watch period.

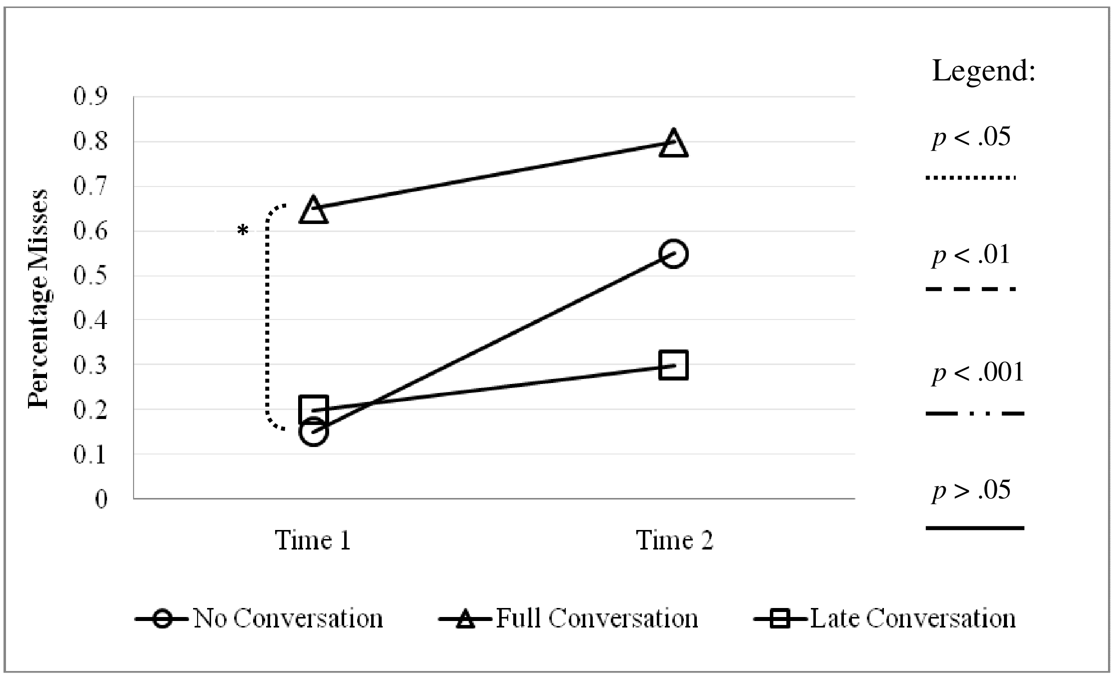

Figure 7. Percentage False Alarms for SART over Two Time Blocks

4.2.3. False Alarms

Results indicated a significant main effect for task type, $F(2, 57) = 7.09$, $p <.05$, $MSE = 1.61$; partial-$\eta^2 = .19$, where the full conversation task ($M = 1.46$) led to more false alarms than the no conversation task ($M = 0.45$) and the late conversation task (*M* = 0.7). There was also a significant main effect for watch period, $F(1, 57)$, = 10.39, $p < .01$, $MSE = 1.62$; partial-$\eta^2 = .15$, where there was an overall decrease in misses from the first watch period (*M* = 1.25) and the second watch period (*M* = 0.5). There was also a significant task type by time interaction, $F(2, 57) = 4.63$, $p < .05$, $MSE = 1.62$; partial-$\eta^2 = .14$.

Table 7: Mean values (and standard deviations) of False Alarms for MVT

Task Type	Watch Period 1	Watch Period 2	*Average*
No Conversation	0.6 (0.82)	0.3 (0.47)	*0.45 (0.65)*
Full conversation	2.35 (2.62)	0.6 (0.75)	*1.46 (1.69)*
Late conversation	0.8 (0.96)	0.6 (0.68)	*0.7 (0.82)*
Average	*1.25 (1.83)*	*0.5 (0.65)*	

Planned comparisons with least significant difference adjustments revealed that misses decreased significantly over the duration of the experiment for the full conversation task, $F(1, 57) = 10.01$, $p < .01$. There were no significant changes in false alarm rates for the no conversation task and late conversation task.

Simple effect analysis using univariate ANOVA between tasks on watch periods indicated that there was significant difference between task types in the first watch period, $F(2, 57) = 6.52$, $p < .01$, $MSE = 2.82$; partial-$\eta^2 = .19$. There was no significant difference between task types in the second watch period. $F(2, 57) = 1.44$, $p = .25$, $MSE = 0.42$; partial-$\eta^2 = .05$. Follow up pairwise comparisons were conducted on first watch period to investigate differences between the tasks. Results indicated that participants in the full conversation task made significantly more false alarms than participants in the no conversation task $F(2, 57) = 10.87$, $p < .01$, and the late conversation task $F(2, 57) = 8.50$, $p < .01$. There was no significant difference between the no conversation task and late conversation task in this watch period.

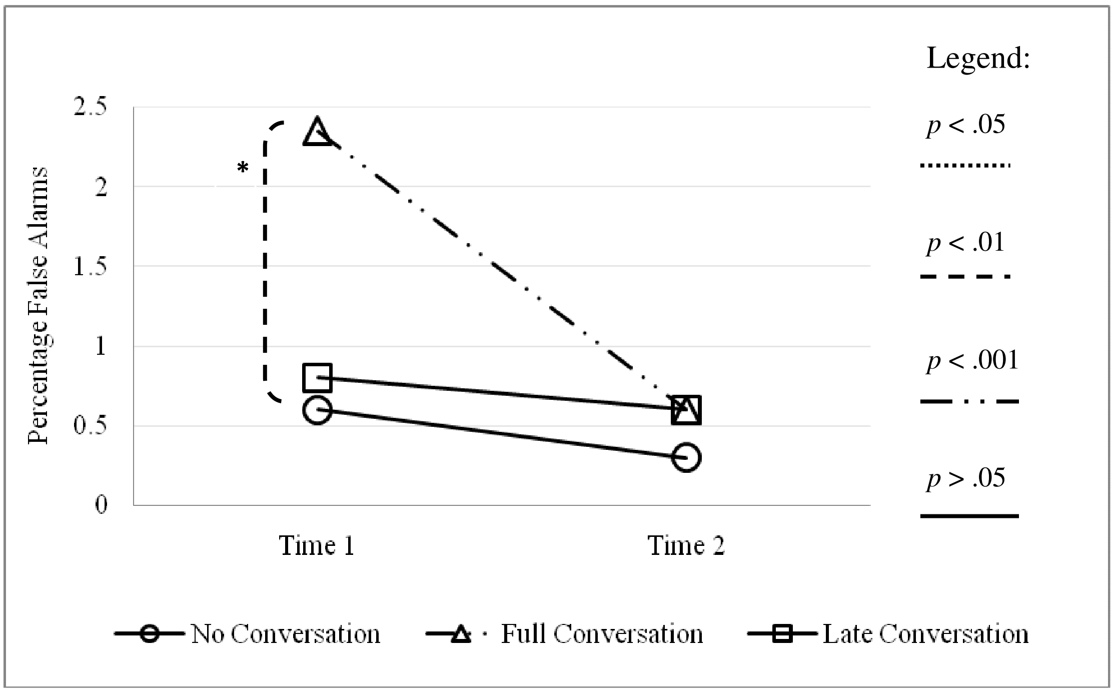

Figure 8. Percentage False Alarms for MVT over two time blocks

4.2.4. NASA-TLX Ratings

A one-way ANOVA was used to measure potential differences of participant's perceived task load. Results indicated that all three task types rated (M = 65.15, SD = 12.41) the MVT as mentally demanding, further analysis indicated that there was a significant between task difference in TLX Scores. $F(2, 57) = 3.64$, p <

.05 Follow up pairwise comparisons indicated that the participants in the full conversation task (M = 70.15, SD = 12.29), rated the MVT as more mentally demanding than participants in the late conversation task (M = 60.08, SD = 10.84), $F(1, 57) = 7.26$, $p < .01$. MVT ratings did no differ significantly in between the other tasks.

4.3 Discussion

As previously discussed, the MVT was a vigilance task used to investigate vigilance in patients with sleep disorders (Hirshkowitz et al., 1993). Its testing methodology was similar to traditional vigilance tasks where participants were only required to respond to the target stimuli as quickly as possible and ignore the non-target stimuli.

4.3.1. Reaction Time

A faster reaction time in the MVT was indicative of the participant's level of attention. Quicker reaction times would indicate higher levels of attention to the task. Results suggest that reaction times increased substantially over the duration of the task, and support the notion of a vigilance decrement as time on task did indeed have an effect on reaction time.

The substantially longer reaction times in the first watch period for the full conversation task as compared to the no conversation and late conversation tasks can be attributed to the effect of the conversation. As previously discussed, this increase in reaction time was likely due to dual task interference, and participants under the full conversation task condition had to time share their resources.

Reaction times in the second watch period were significantly longer as compared to the first watch period for the no conversation and late conversation task. The longer reaction times for the late conversation could be attributed to the presence of a conversation, where participants were unable to successfully timeshare their resources. Statistically, all tasks were equal in the reaction time performance, thus further supporting the notion that the presence of a secondary task during a vigilance task does not lead to better performance.

4.3.2. Misses

Misses will not be discussed due to extremely low miss rates in all groups. Though it could be suggested that hits be analyzed instead due to having larger values, the low miss rates would imply high hit rates as these measures were inversely proportional. Thus the differences between groups even if significant would be an artifact of statistical analysis rather than an actual effect of manipulation of the independent variables.

4.3.3. False Alarms

Low levels of false alarms in the MVT would indicate a higher level of attention. Results suggest that this was the case in the first watch period for the no conversation and late conversation tasks, as compared to the full conversation task. The significantly higher number of misses in the full conversation task could be attributed to the presence of the conversation task during the onset of the task.

During the second watch period, misses were all statistically equal to each other for all task groups. Participants in the full conversation task appeared to have successfully learned to timeshare resources. The non significant change in false alarms for the late conversation seems to suggest that there was no dual task cost as might be expected. A possible reason might be attributed to sufficient practice. Prior to the onset of the conversation, participants would have undergone the task for a period of time. Thus with the onset of the conversation they were able to successfully redeploy attentional resources to the secondary task at no cost to the primary vigilance task.

4.3.4. NASA-TLX Ratings

Results from the NASA-TLX ratings were consistent with the other vigilance tasks, as all task types rated the MVT in the upper scale. The difference between full conversation and late conversation tasks could be due to participants rating the tasks as a whole. This would imply that the onset of a conversation late in the task could possibly be less taxing than having to sustain a prolonged conversation.

Table 8. Summary Table of MVT results (effects in parenthesis)

Performance At Start			
	No Conversation	Full Conversation	Late Conversation
Reaction Time	Baseline	Slower (Dual Task)	Baseline
Misses	Baseline	More (Dual Task)	Baseline
False Alarm	Baseline	More (Dual Task)	Baseline
Performance Over Time			
	No Conversation	Full Conversation	Late Conversation
Reaction Time	Slower (Decreased Vigilance)	No Change (Floor)	Slower (Dual Task)
Misses	No Change (Ceiling)	No Change (Ceiling)	No Change (Ceiling)
False Alarm	No Change (Ceiling)	Less (Effective Timeshare?)	No Change (Timesharing)

5. Experiment 3 – Vigilance Drive

5.1. Methods

5.1.1. Participants

Thirty students from the University of Kansas undergraduate research pool, and student from an upper level psychology class, 10 men and 20 women, averaging 21.06 years (*SD* = 1.41), participated for course credit. They had an average of 5.3 years of driving experience (*SD* = 1.58). Participants were randomly assigned into one of three task conditions: no conversation, 4 men and 6 women; full conversation, 3 men and 7 women; late conversation, 3 men and 7 women. All participants reported English as their primary/native language. A Snellen Visual Acuity Chart was used to test for visual acuity in participants. All participants had normal to corrected visual acuity of 20/32 or better. One participant was disqualified and replaced due to not having a driver's license.

5.1.2. Materials and apparatus

Driving simulator. The driving scenario was simulated using STISIM Drive (Systems Technology Inc. Hawthorne, CA) simulator software (Version 2.08.02). The simulator was a fixed-base unit, where participants were seated in Playseats Classic gaming seat. Vehicular control was done via a Thrustmaster™ force feedback steering and pedal set. The vehicle was set on automatic transmission. Participants viewed the simulated road on a single 17 inch LCD display. A Fresnel lens was placed between the LCD display and the participant to create an artificial 3D space to improve the realism of the scenario. Viewing distance varied according to how participants adjusted the seat.

Driving scenario. Participants from all three task types drove in the same simulated environment. The roadway was a four-lane rural highway. Each lane was 12 ft (3.66m) in width. A 12 ft (3.66 m) median separated traffic from the opposite direction. There were occasional curves and hills on the roadway to ensure a degree of realism. There were periods of intermittent traffic from the opposite direction and in the participant's lane. Traffic in the participant's lane travelled at speeds lower than the participant, thus allowing for passing to take place. The total distance of the drive was 140000 ft (42.67 km)**.** No data was collected in the first 5000 ft (1.52 km) of the drive to allow participants to reach the specified speed of 65 mph (104.6 kph). To mimic highway speeds by using cruise control, participants were told to fully depress the accelerator to achieve the specified speed. Data collection began from the 5000 ft (1.52 km) mark till the 137000 ft (41.76 km) mark. Data collection was broken into 5 blocks, each block lasting 26400 ft (8.05 km). Participants took an average of 25 minutes to complete the scenario.

5.1.3. Events and tasks in scenario

Wind gusts. To increase the difficulty of the task, a continuous wind gust lasting 30 seconds was introduced within each block except the penultimate block.

Participants were thus required to make effortful control of the car to ensure safe driving.

Critical event. To further investigate whether participants were paying attention to the drive, critical events occurred in the first and last block of the drive. In both blocks, a car parked on the road shoulder would suddenly pull out onto the road when the participant's vehicle was one second away from it (the timing was based on the participant's current speed).

Memory task. Within each driving block, a billboard image of a popular local fast food chain would be presented on the right shoulder. This served two purposes, firstly, it added to the realism of the scenario as billboard advertisements appear on most highways. Secondly, this served as a memory task for participants at the end of the drive.

Conversation Task. The conversation task and apparatus in this experiment was similar to the first experiment, including the wordlist and timing conditions used for the conversation task. The only difference in this experiment was the onset for the presentation of words in the LC condition. For this experiment, the late conversation began 17.5 minutes into the drive.

5.1.4. Procedure

Upon completion of the consent form and demographic datasheet that queried cell phone usage behavior, and driving experience, participants were given an eye test to check visual acuity. Following which, participants engaged in a practice drive that lasted approximately 2 minutes. This practice drive allowed participants to familiarize themselves with the simulator and the handling of the steering. Participants were instructed to fully depress the accelerator pedal to achieve the specified speed of 65mph (104.6 kph), and were told to try their best to maintain that speed. At the end of the practice block, the experimenter would ask all participants to equip the hands free kit by placing the headphone over their right ear, and extending the boom microphone to their mouth. The experimenter informed all participants that they might or might not be engaged in a conversation task. Once participants understood the conversation task instructions and were ready to begin the task, the participant would say ‘Go’ into the microphone. This would remotely activate the computer responsible for the conversation task. Upon the completion of the experiment, participants completed an electronic version of the NASA-TLX (Hart and Staveland, 1988), and were two minutes to recall as many billboards as they could remember while driving. The entire experiment was conducted in a darkened room with no ambient light source, except light emanating from the scenario. Luminance levels were 0 cd/m^2 as measured by a Minolta LS110 Luminance Meter.

5.2. Results

5.2.1. Driving Performance.

Participants were required to control their vehicle laterally by trying to stay in the middle of the right lane, relative to the roadway dividing line. Larger positive values indicate deviation away from the roadway dividing line, toward the right shoulder and smaller positive values indicate deviation toward the roadway dividing line. Standard deviation data from lane position data was used to indicate variability within the lane. Greater numbers indicate greater movement within the lane. Lane position was measured in feet. Data was collected approximately 10 times per second, or 10 Hz. The initial design was to analyze driver performance based on a critical event (described later in this section), but results indicated that not all drivers were affected by the critical event, i.e. hit the car that suddenly moved into the roadway, as expected by the experimenter. Remaining in line with the effects of dual tasking and driving, this analysis focused on the driving ability of participants when presented with the critical wind gust event. Thus, rather than investigating results between time blocks, analysis focused on the driver's ability to maintain control of the vehicle 30 seconds prior, and 30 seconds after a the wind gust event. This would allow for a clearer investigation on the effects of a secondary conversational task on driving performance when presented with event that required attention to ensure safe driving behavior.

Separate 3(Task type)x2(30seconds prior, 30 seconds post) Mixed-ANOVAs were conducted with task type as between subjects factor and 30 seconds prior and 30 seconds post wing gust as within subjects factor. Two Mixed-ANOVAs analyzed lane keeping performance for watch period 1 and 2 respectively, and two other Mixed-ANOVAs analyzed lane variability for watch periods 1 and 2.

Lane keeping performance for watch period 1. Results did not indicate any significant or marginally significant main effects or interaction effects for the lane position of the vehicle relative to the roadway dividing line.

Lane keeping performance for watch period 2. Results indicated a significant main effect for task type $F(2, 27) = 6.15$, $p < .01$, $MSE = 27.33$; partial-$\eta^2 = .31$. Participants in the late conversation task ($M = 20.07$) were driving closer to the roadway diving line as compared to participants in the full conversation task ($M = 24.54$), and the no conversation task ($M = 25.49$). There was no significant main effect for time, and there were no significant interactions between factors.

Table 9: Mean lateral lane position (and standard deviations) 30 seconds before and after a wind gust event at both watch periods (All values in feet)

	Watch Period 1			Watch Period 2		
	30s Prior	30s After		30s Prior	30s After	
Task Type			*Average*			*Average*
No Conversation	25.58(1.74)	24.52(3.59)	*25.05(2.67)*	25.58(1.74)	25.41(5.14)	*25.49(3.44)*
Full conversation	24.39(3.11)	24.19(4.6)	*24.29(3.86)*	24.39(3.11)	24.68(5.93)	*24.54(4.52)*
Late conversation	24.72(4.42)	24.27(4.5)	*24.49(4.46)*	18.27(5.32)	21.86(5.83)	*20.07(5.58)*
Average	*24.89(3.2)*	*24.33(4.11)*		*22.44(4.83)*	*23.98(5.67)*	

Simple effect analysis using univariate ANOVA indicated that there were significant differences between task types 30 seconds prior to encountering the wind gust event. $F(1, 27) = 11.29$, $p <.001$, $MSE = 13.65$; partial-$\eta^2 = .46$ but no differences between task types 30 seconds after the wind gust event $F(1, 27) = 1.1$, *p*

= .35, *MSE* = 31.87; partial-η^2 = .08 (See Figure 8). Follow up pairwise comparisons were conducted on the first 30 seconds to investigate differences between the tasks. Results indicated that in the initial 30 seconds, drivers in the no conversation task were driving significantly further from the roadway dividing line compared to the drivers in late conversation task $F(1, 27) = 19.6$, $p < .001$. It was also observed that drivers in the full conversation task were also driving further to the right of the roadway dividing line as compared to drivers in the late conversation task $F(1, 27) = 13.37$, $p < .001$. There was no significant difference between drivers in the no conversation task and drivers in the full conversation task.

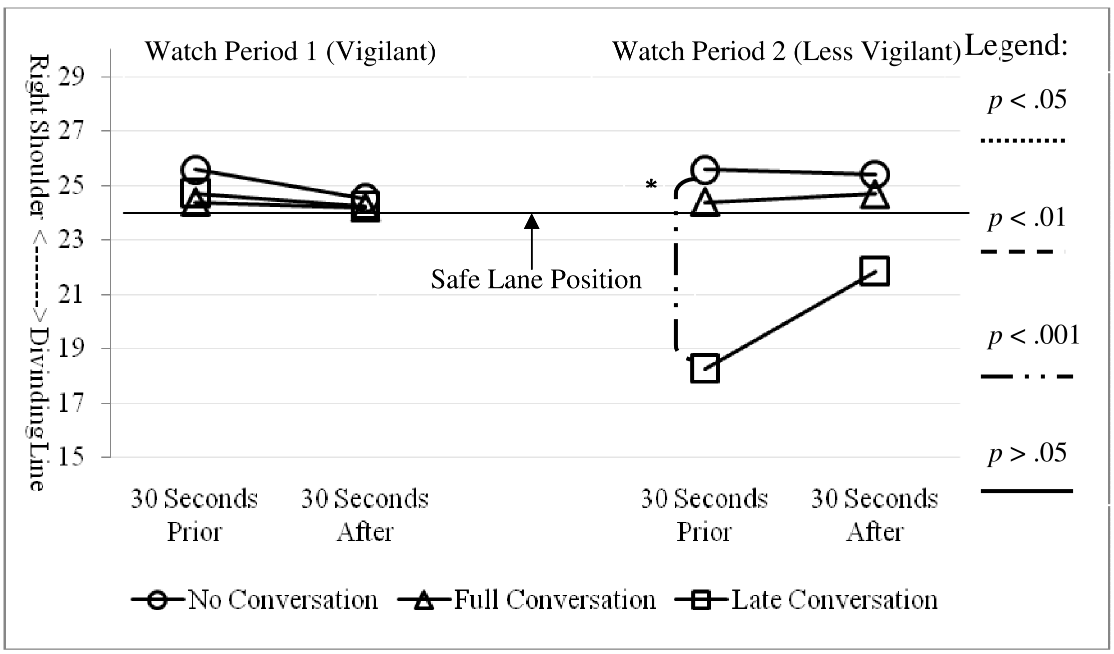

Figure 9. Lane position before and after wind gust for two time blocks

Mean standard deviation lane position for watch period 1. Results indicated a marginally significant main effect for time, $F(1, 27) = 3.93$, $p = .06$, *MSE* = 1.64;

partial-η^2 = .13. There was more variability in the 30 seconds prior (M = 3.22) to the wind gust event as compared to the 30 seconds post (M = 2.56) wind gust event. There was no main effect for task type, nor was there an interaction effect.

Planned comparisons with least significant difference adjustments revealed that lane variability decreased significantly 30 seconds prior to the wind gust as compared to 30 seconds post wind gust for the full conversation task, $F(1, 27) = 6.42$, $p < .05$. There were no significant changes in lane variability in the no conversation task and late conversation task.

Mean standard deviation lane position for watch period 2. Results indicated a significant main effect for time $F(1, 27) = 10.89$, $p < .01$, $MSE = 2.54$; partial-η^2 = .29. Lane variability was greater 30 seconds prior to the wind gust (M = 2.87) as compared to 30 seconds post wind gust (M = 1.51) There was also a marginally significant interaction effect, $F(1, 27) = 3.12$, $p = 0.06$, $MSE = 2.54$; partial-η^2 = .19. There was no main effect for task type.

Planned comparisons with least significant difference adjustments revealed that there was a significant decrease in lane variability 30 seconds after the wind gust event as compared to 30 seconds before the wind gust event for the no conversation group $F(1, 27) = 6.29$, $p < .05$. This decrease in lane variability was also observed in the full conversation group $F(1, 27) = 10.84$, $p < .01$. The late conversation group did not show a significant change in driving variability.

Table 10: Mean lane deviation (and standard deviations) 30 seconds before and after a wind gust event at both watch periods (All values in feet)

	Watch Period 1			Watch Period 2		
	30s Prior	30s After		30s Prior	30s After	
Task Type			*Average*			*Average*
No Conversation	3.07 (1.22)	2.81 (1.61)	*2.94 (1.42)*	3.07 (1.22)	1.28 (0.74)	*2.18 (0.98)*
Full conversation	3.68 (1.35)	2.23 (1.42)	*2.96 (1.39)*	3.68 (1.35)	1.33 (0.89)	*2.51 (1.14)*
Late conversation	2.89 (1.41)	2.64 (1.11)	*2.77 (1.26)*	1.86 (1.49)	1.92 (2.46)	*1.89 (1.98)*
Average	*3.22 (1.33)*	*2.56 (1.37)*		*2.87 (1.52)*	*1.51 (1.54)*	

Simple effects analysis using univariate ANOVA indicated that there were significant differences between task types 30 seconds prior to the wind gust event. $F(1, 27) = 4.66$, $p < .05$, $MSE = 1.84$; partial-$\eta^2 = .26$ but no differences between task types 30 seconds after the wind gust event $F(1, 27) = .51$, $p = .6$, $MSE = 2.46$; partial-$\eta^2 = .03$ (See Figure 9). Follow up pairwise comparisons were run on the initial 30 seconds prior to the wind gust. Results indicated that in the initial 30 seconds, drivers in the full conversation task were driving more erratically compared to the drivers in late conversation task $F(1, 27) = 8.99$, $p < .01$. There was also a marginal difference between the no conversation task and late conversation task $F(1, 27) = 3.95$, $p = .06$. There was no significant difference between the no conversation task and full conversation task during this period.

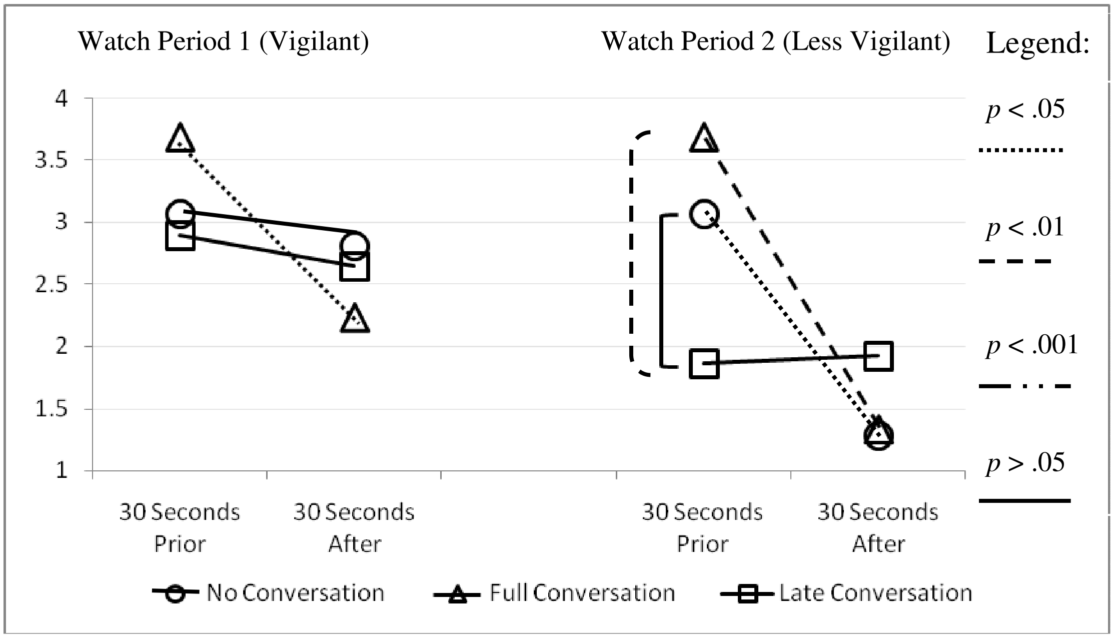

Figure 10. Lane variability before and after wind gust for two time blocks

5.2.2. Memory Task

A one-way ANOVA was used to measure potential differences in driver's ability to recall billboards that appeared on the road shoulder during the drive. Results indicated a difference between task types, $F(2, 27) = 4.71$, $p < .05$ Follow up pairwise comparisons indicated that the participants in the no conversation task ($M = 3$, $SD = 0.82$), were able to recall more billboard images than participants in the full conversation task ($M = 1.7$, $SD = 1.16$), $F(1, 27) = 6.52$, $p < .05$. Results also indicated that participants in the late conversation task ($M = 3.1$, $SD = 1.37$), were able to recall more billboard images than participants in the full conversation task ($M = 1.7$, $SD = 1.16$), $F(1, 27) = 7.56$, $p < .05$.

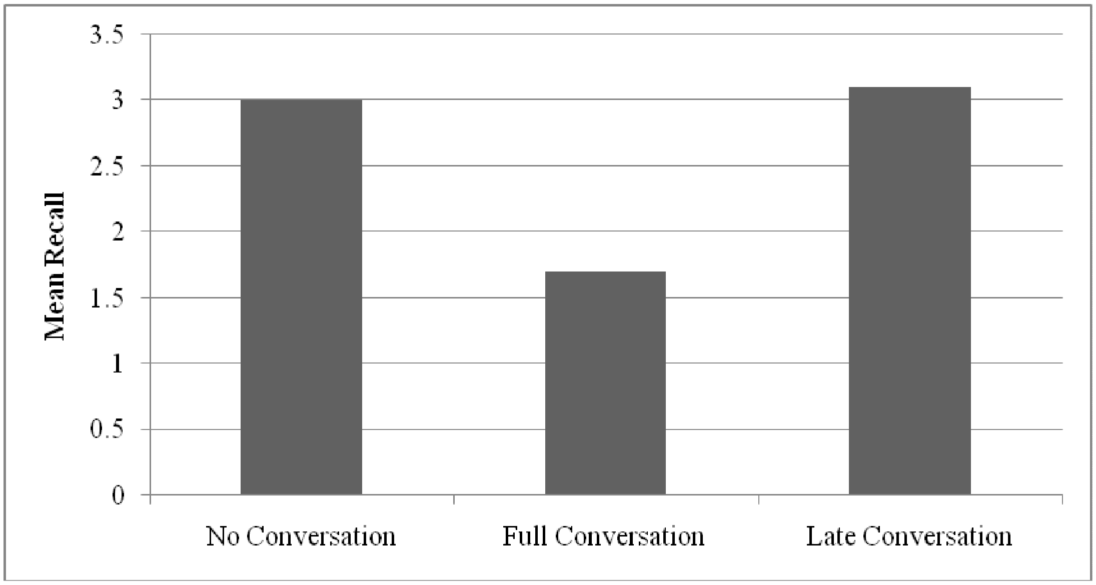

Figure 11. Mean recall ability between task types

5.2.3. Critical Events

Separate one-way ANOVAs was used to measure potential differences as to whether drivers were able to successfully avoid hitting an errant car that jumped onto the road way. There were no significant differences in crash avoidance between groups in the first watch period , (M = .4, SD = .49), $F(2, 27) = 0.39$, $p = .68$, and the second watch period (M = .3, SD = .46), $F(2, 27) = 0.44$, $p = .65$.

Results from both critical events were both non significant thus as noted earlier, the investigation did not use this event to measure driver behavior. The rationale for such as decision was due to the fact that since this measure was to be a within subjects measure, all drivers were expected to collide with the vehicle, but that was not the case, some drivers successfully avoided collision with the vehicle, thus the decision not to use this event.

5.2.4. Subjective Task-load scores - NASA-TLX

A one-way ANOVA was used to measure potential differences of driver's perceived task load. Results indicated that all three task types rated (M = 58.03, SD = 13.18) the drive as fairly demanding, and there were no significant differences between tasks. $F(2, 27) = 3$, p =.07.

5.3 Discussion

The driving task was designed to be boring in an attempt to simulate monotonous highway driving. Besides having to drive the vehicle in a safe and consistent manner, participants, though not explicitly informed, were also required to be aware of the driving environment. It was believed that events such as wind gusts or a critical braking event would be able to bring about changes in driving behavior.

5.3.1. Driving Performance

Mean lane position. There were no significant changes in driving behavior during the first watch period, before or after the wind gust. This was surprising as no dual task interference effects were observed in the full conversation task during this period. During the second watch period, participants in the late conversation task appeared to be driving closer to the roadway dividing line as compared to the no conversation task and full conversation task prior to the wind gust event. This poorer lane position could be attributed to the additive effect of a decrease in vigilance and the presence of the conversation task. This would suggest that conversation did not improve drivers' ability to maintain a safe lane position (See Table 11).

Table 11. Summary table of lane position results before and after wind gusts (effects in parenthesis)

Before Wind gust			
	No Conversation	Full Conversation	Late Conversation
Lane Position (Time 1)	Stable	Stable (No Dual Task?)	Stable
Lane Position (Time 2)	Stable	Stable	Further Left (Dual Task)
After Wind gust			
Lane Position (Time 1)	No significant change	No significant change	No significant change
Lane Position (Time 2)	No significant change	No significant change	No significant change

Mean lane deviations. Drivers appeared to be driving erratically before the wind gust event relative to after the wind gust in first and second watch period suggesting a possible drift of attention. However, 30 seconds after the wind gust, drivers appear to be less erratic in their driving. The overall difference would suggest a possible energizing effect from the wind gust thus leading drivers to minimize lane variability. It was noted that participants in the full conversation task showed a significant decrease in lane variability in the first watch period. This result was interesting as drivers in the full conversation task were dual tasking, yet were able to achieve a tighter driving pattern after the wind gust event indicating that they were aware of the potential danger of the wind gust. Furthermore, it could also be suggested that the wind gust was able to redirect the attention of drivers back to the driving task as indicated by a decrease in lane variability (See Table 12).

Interestingly, during the second watch period, drivers in the late conversation task did not show any significant changes in lane variability, as compared to drivers in the other tasks. Firstly, driver variability 30 seconds prior to the wind gust was lesser in the late conversation task as compared to the other tasks. This was interesting in the light of being engaged in a conversation, as it was expected that drivers would be more erratic in their driving as observed in participants in the first watch period of the full conversation task. It could be possible that engaging in a conversation later in the drive energized participants as it drew them out from a state of boredom. Secondly, this consistent lane keeping behavior was also noted 30 seconds after the wind gust event, suggesting that the wind gust had no effect on participants in the late conversation task. The energizing effect of the wind gust was also clearly visible in the no conversation and full conversation task as their lane keeping variability decreased significantly 30 seconds after the wind gust.

Table 12. Summary table of lane variability results before and after wind gusts (effects in parenthesis)

Before Wind gust			
	No Conversation	Full Conversation	Late Conversation
Lane Variability (Time 1)	Stable	Stable(No Dual Task?)	Stable
Lane Variability (Time 2)	More Variable (Vigil Decrement)	More Variable (Vigil Decrement)	Less Variable (Dual task)
After Wind gust			
Lane Variability (Time 1)	No significant change	Decreased Variability (but Dual Task)	No significant change
Lane Variability (Time 2)	Decreased Variability	Decreased Variability (but Dual Task)	No significant change

5.3.2. Memory Task

The current study also lends further support to data found by Strayer, Drews and Johnston (2003). Participants in the dual task conditions were unable to recognize as many signs when compared to a single task condition. Results from the current study indicate the same results. This finding was further extended whereby participants in the late conversation condition had seen all the signs prior to the conversation task, indicating that prior to the conversation, participants were able to encode the signs seen during the drive in the same fashion as drivers not engaged in conversation.

5.3.3. Critical Events

As discussed earlier, there was little difference between groups with regard to the critical events for either watch period. Thus this measure was not used

5.3.4. Subjective Task-load scores - NASA-TLX

Participants rated the drive as fairly demanding, suggesting that the drive might not be as demanding as a pure laboratory test of vigilance.

6. Conclusions & Recommendations

Two vigilance tasks with different methodologies were used to investigate the effects of a secondary conversational task on the primary vigilance task. Broadly speaking, results from both studies were fairly consistent in two areas. Firstly, in both experiments, the effects of time on task were clearly seen whereby task performance decreased as time on task progressed, suggesting a decrease in vigilance. Secondly, the detrimental effects of a secondary conversational task were also clearly observed in instances such as the full conversation task in the first watch period of both the SART and MVT. These results appear to suggest that both tasks were indeed testing vigilance levels of participants. Furthermore, these results were buffered by the uniquely high scores on the NASA-TLX that could only be attributed to demanding tasks such as a vigilance task. Yet, there were admittedly some differences between the SART and MVT too.

One of the secondary goals of this study was to investigate potential differences between the two methodologies used to study vigilance. Results from the current set of studies suggest that task manipulations appear to be more amenable toward the SART rather than the MVT. By using the SART, the study was able to show that the presence of a prolonged secondary task did, in fact, improve performance, but only to a level comparable to that of an individual who had a low level of vigilance. On the other hand, the late conversation failed to improve performance, though it could be argued that participants were trying to pay more attention to the primary task as attested by their slower reaction times. Investigation into the quality of conversation could show whether or not participants were attending to the conversation more than they were attending to the primary task. It might then be suggested that the SART could be used to investigate lapses of attention and the MVT more amenable to the classic vigilance decrement.

The current sets of results do not invalidate the MVT or any traditional forms of vigilance studies. Instead, it might indicate that tasks such as the MVT require longer durations, as originally designed, before clearer differences are to be detected using the current study's manipulations. Furthermore, both tasks were investigating vigilance albeit taking different points of view. The designers of the SART envisioned vigilance as the ability to control the endogenous aspect of attention, which would lead the observer to inhibit an on-going behavior when needed. (Robertson, & Garavan, 2004). While the designers of traditional vigilance studies envisioned vigilance as the ability to successfully detect and respond to changes in the environment, contingent on an exogenously driven state of constant watch keeping. Whether or not vigilance tasks were found to be mindless or resource demanding was not fully answered in the current set of studies. Future studies would be needed to clarify this.

The driving simulator aspect of the study was an attempt to further expand the used to the plausibility of a possible benefit of talking on a cell phone while driving on a monotonous route. The current results were inconclusive when taken in the light of a monotonous drive. Results suggest that drivers who talked on the cell phone were more likely to miss stimuli that occur in the periphery, as noted by a poorer recollection of the billboards on the drive. This result also supports the findings by Atchley and Dressel (2004), that conversing on a cell phone shrinks the useful field of view. This finding would indicate poorer visual attention, but does not address driving ability when the driver was bored. On the other hand, data suggests that drivers who engaged in a conversation only in the second watch period were better able to control their car, before and after the wind gust event as compared to drivers in the other conditions. This result could possibly imply a benefit of short intermittent conversations while engaged in a monotonous drive. More work will have to be done to further investigate this claim. It was also observed that upon encountering the wind gust, drivers in all task types were able to significantly decrease their lane variability. This suggests that even drivers who were engaged in conversation were able to return their attention to the driving task when a potentially dangerous event occurred. Further studies should investigate whether or not this return to attention could be a matter of 'too little, too late'. Furthermore, to be able to truly claim effects of a cell conversation on a driving task, future studies may require participants to drive for more than 30 minutes.

Another aspect of task performance that was not explicitly investigated in this project was that of fatigue. It is believed that fatigue could have a detrimental effect on task performance, but the duration of the current set of studies were not long enough to induce fatigue in participants. As with future studies in the simulator, future studies in vigilance will investigate the effects of fatigue on vigilance task performance by increasing the duration of the tasks.

In conclusion, results from these three studies have covered a broad scope. Firstly, the SART has been shown to be more robust than the MVT in showing the benefits of a secondary task on a vigilance task. But the benefits do not appear to outweigh the cost as performance only reaches a level of an individual who was low in vigilance. Furthermore this benefit only appears after an extend period of conversation. Secondly, regardless of task, performance decreases over time. Finally, results from the driving simulator have shown some degree of benefit of talking on a cell phone while driving, though it would be only prudent to further investigate this reported benefit, as results from an applied study could have profound impact if misreported or misrepresented. It would be wise to err on the side of caution and not attempt to dual task in any circumstance, including times when one feels bored and possibly fatigued.

7. References

Antonelli, D. C., & Karas, G.G. (1967). Performance on a vigilance task under conditions of true and false knowledge of results. *Perceptual and Motor Skills, 25*, 129-138.

Atchley, P., & Dressel, J. (2004). Conversation limits the functional field of view. *Human Factors, 46*(4), 664-673.

Baker, C. H. (1962). On temporal extrapolation. *Canadian Journal of Psychology, 16*, 37-41.

Ballard, J. C. (1996). Computerized assessment of sustained attention: a review of factors affecting vigilance performance. *Journal of Clinical and Experimental Neuropsychology, 18*(6), 843-863.

Berlyne, D. E. (1960). *Conflict, arousal, and curiosity*. New York: McGraw-Hill.

Bevan, W., & Turner, E.O. (1965). Vigilance performance with a qualitative shift in verbal reinforcers. *Journal of Experimental Psychology, 70*, 83-86.

Bradley, M.M., & Lang, P.J. (1999). Affective norms for English words (ANEW). Gainesville, FL. The NIMH Center for the Study of Emotion and Attention,University of Florida.

Broadbent, D.E. (1958). *Perception and Communication*. London: Pergamon Press.

Broadbent, D. E., Cooper, P.F., Fitzgerald, P., & Parkes, K.R. (1982). The cognitive failures questionnaire (CFQ) and its correlates. *British Journal of Clinical Psychology, 21*, 1-16.

Brown, K. W., & Ryan, R.M. (2003). The benefits of being present: mindfulness and its role in psychological well-being. *Journal of Personality and Social Psychology, 84*(4), 822-848.

Chan, R. C. K. (2001). A further study on the sustained attention response to task (SART): the effect of age, gender, and education. *Brain Injury, 15*(9), 819-829.

Cheyne, J. A., Carriere, J.S.A., & Smilek, D. (2006). Absent-mindedness: Lapses of conscious awareness and everyday cognitive failures. *Consciousness and Cognition, 15*, 578-592.

CITA. (2008). *Estimated Current US Wireless Subscribers*. Retrieved October 19, 2008, from http://www.ctia.org/

Craig, A. (1978). Is the vigilance decrement simply a response adjustment toward probability matching. *Human Factors, 20*, 441-446.

Davies, D. R., & Parasuraman, R. (1982). *The Psychology of Vigilance*. London: Academic Press Inc

Deese, J. (1955). Some problems in the theory of vigilance. *Psychological Review, 62*(5), 359-368.

Duffy, E. (1962). *Activation and behavior*. New York: Wiley.

Erwin, C. W., Wiener, E.L., Linnoila, M.I., & Truscott, T.R. (1978). Alcohol-induced drowsiness and vigilance performance. *Journal of Studies on Alcohol, 39*(3), 505-516.

Fan, J., McCandliss, B.D., Sommer, T., Raz, A., & Posner, M.I. (2002). Testing the efficiency and independence of attentional networks. *Journal of Cognitive Neuroscience, 14*(3), 340-347.

Galinsky, T. L., Warm, J.S., Dember, W.N., Weiler, E.M., & Scerbo, M.W. (1990). Sensory alternation and vigilance performance: the role of pathway inhibition. *Human Factors, 32*(6), 717-72

Glassbrenner, D. (2005). *Traffic safety facts: research note. Driver cell phone use in 2005 - overall results*. Research Note No. DOT HS 809 967. Washington, D.C: U.S. Department of Transportation.

Green, D. M., & Swets, J.A. (1966). *Signal detection and psychophysics*. New York: Wiley.

Grier, R. A., Warm, J.S., Dember, W.N., Matthews, G., Galinsky, T.L., Szalma, J.L., et al. (2003). The vigilance decrement reflects limitations in effortful attention, not mindlessness. *Human Factors, 45*, 349-359.

Groves, P. M., & Thompson, R.F. (1970). Habituation: a dual-process theory. *Psychological Review, 77*(5), 419-450.

Hart, S. G., & Staveland, L.E. (1988). Development of NASA-TLX (task load index): Results of empirical and theoretical research. In P. A. Hancock, & Meshkati, N. (Ed.), *Human mental workload* (pp. 239-250). Amsterdam: North Holland Press.

Helton, W. S. (2008). Implusive responding and sustained attention and response task. *Journal of Clinical and Experimental Neuropsychology, iFirst*, 1-9.

Helton, W. S., Hollander, T.D., Warm, J.S., Matthews, G., Dember, W.N., Wallaart, M., et al. (2005). Signal regularity and the mindlessness model of vigilance. *British Journal of Psychology, 96*, 249-261.

Helton, W. S., & Warm, J.S. (2008). Signal salience and the mindlessness theory of vigilance. *Acta Psychologia, 129*, 18-25.

Hirshkowitz, M., De La Cueva, L., & Herman, J.H. (1993). The multiple vigilance test. *Behavior Research Methods, Instruments, and Computers, 25*(2), 272-275.

Hitchcock, E. M., Warm, J.S., Matthews, G., Dember, W.N., Shear, P.K., Tripp, L.D., et al. (2003). Automation cueing modulates cerebral blood flow and vigilance in a simulated air traffic control task. *Theoretical Issues in Ergonomics Science, 4*(1-2), 89-112.

Jack, A., & Roepstorff, A. (2003). Why trust the subject? *Journal of Consciousness Studies, 10, (9-10)*, v-xx (1).

James, W. (1890). *The Principles of Psychology. Vol. 1*. New York, NY: Henry Holt and Company.

Jerison, H. J., & Pickett, R.M. (1964). Vigilance: The importance of the elicited observing rate. *Science, 143*, 970-971.

Kahneman, D. (1973). *Attention and Effort*. Englewood Cliffs, NJ: Prentice Hall.

Kubose, T. T., Bock, K., Dell, G.S., Garnsey, S.M., Kramer, A.F., & Mayhugh, J. (2006). The effects of speech production and speech comprehension on simulated driving performance. *Applied Cognitive Psychology, 20*, 43-63.

Levine, J. M. (1966). The effects of values and costs on the detection and identification of signals in auditory vigilance. *Human Factors, 8*, 525-537.

Lindsley, D. B. (1951). Emotion. In S. S. Stevens (Ed.), *Handbook of experimental psychology* (pp. 473-516). New York: Wiley.

Loeb, M., & Binford, J.R. (1971). Modality, difficulty and coupling in vigilance behavior. *American Journal of Psychology, 84*, 529-541.

Mackie, R. R. (1987). Vigilance research - are we ready for countermeasures? *Human Factors, 29*(6), 707-723.

Mackworth, J. F. (1968). Vigilance, arousal, and habituation. *Psychological Review, 75*(4), 308-322.

Mackworth, N. H. (1948). The breakdown of vigilance during prolonged visual search *The Quarterly Journal of Experimental Psychology, 1*(1), 6-21.

Mackworth, N. H. (1957). Some factors affecting vigilance. *Advancement of Science, 53*, 389-393.

Manly, T., Lewis, G.H., Robertson, I.H., Watson, P.C., & Datta, A.K. (2002). Coffee in the cornflakes: time-of-day as a modulator of executive response control. *Neuropsychologia, 40*, 1-6.

Manly, T., Robertson, I.H., Galloway, M., & Hawkins, K. (1999). The absent mind: further investigations of sustained attention to response. *Neuropsychologia, 37*, 661-670.

Manly, T., Owen, A.M., McAvinue, L., Datta, A., Lewis, G.H., Scott, S.K., et al. (2003). Enhancing the sensitivity of a sustained attention task to frontal damage: convergent clinical and functional imaging evidence. *Neurocase, 9*(4), 340-349.

Matthews, G., Davies, D.R., Westerman, S.J., & Stammers, R.B. . (2000). *Human Performance: Cognition, stress and individual differences*. Philadelphia, PA: Psychology Press.

Mouloua, M., & Parasuraman, R. (1995). Aging and cognitive vigilance: effects of spatial uncertainty. *Experimental Aging Research, 21*, 17-32.

Parasuraman, R. (1984). Sustained attention in detection and discrimination. In R. Parasuraman, & Davies, D.R. (Ed.), *Varieties of attention* (pp. 243-291). Ontario, FL: Academic Press.

Parasuraman, R. (1985). Sustained attention: a multifactorial approach. In M. I. Posner, & Marin O.S.M. (Ed.), *Attention and Performance XI*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

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, NY: Plenum Press.

Pardo, J. V., Fox, P.T., & Raichle M.E. (1991). Localization of a human system for sustained attention by positron emission tomography. *Nature, 349*, 61-64.

Papadelis, C., Chen, Z., Kourtidou-Papadelis, C., Bamidis, P.D., Chouvarda, I., Bekiaris, E., et al. (2007). Monitoring sleepiness with on-board electrophysiological recordings for preventing sleep-deprived traffic accidents. *Clinical Neurophysiology, 118*, 1906-1922.

Posner, M. I. (1978). *Chronometric explorations of mind.* Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Posner, M. I., & Petersen, S.E. (1990). The attention system of the human brain. *Annual Reviews in Neuroscience, 13*, 25-42.

Rakauskas, M. E., Gugerty, L.J., & Ward, N.J. (2004). Effects of naturalistic cell phone conversations on driving performance. *Journal of Safety Research, 35*, 453-464.

Raz, A., & Buhle, J. (2006). Typologies of attentional networks. *Nature Reviews Neuroscience, 7*, 367-379.

Redelmeier, D. A., & Tibshirani, R.J. (1997). Association between cellular-telephone calls and motor vehicle collisions. *The New England Journal of Medicine, 336*(7), 453-458.

Robertson, I. H., & Garavan, H. (2004). Vigilant attention. In M. S. Gazzaniga (Ed.), *The Cognitive Neurosciences* (3rd ed., pp. 563-578): MIT Press.

Robertson, I. H., Manly, T., Andrade, J., Baddeley, B.T., & Yiend, J. (1997). 'Oops!': Performance correlates of everyday attentional failures in traumatic brain injured and normal subjects. *Neuropsychologia, 35*(6), 747-758.

Schmidt, E. A., Kincses, W.E., Schrauf, M., Haufe, S., Schubert, R., & Curio, G. (2007). *Assessing drivers' vigilance state during monotonous driving.* Paper presented at the Driving Assessment, Stevenson, Washington, USA.

Schnittger, C., Johannes, S., Arnavaz, A., & Munte, T.F. (1997). Relation of cerebral blood flow velocity and level of vigilance in humans. *NeuroReport, 8*, 1637-1639.

See, J.E., Howe, S.R., Warm, J.S., & Dember, W.N. (1995). Meta-analysis of the sensitivity decrement in vigilance. *Psychological Bulletin, 117(2), 230-249.*

Shiffrin, R.M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review, 84(2),* 127-190.

Sipowicz, R. R., & Baker, R.A. (1961). Effects of intelligence on vigilance: a replication. *Perceptual and Motor Skills, 13*, 398.

Sipowicz, R. R., Ware, J., & Baker, R.A. (1962). The effects of reward and knowledge of results on the performance of a simple vigilance task. *Journal of Experimental Psychology, 64*, 58-61.

Smallwood, J., Beach, E., Schooler, J.W., & Handy, T.C. (2008). Going AWOL in the brain: mind wandering reduces cortical analysis of external events. *Journal of Cognitive Neuroscience, 20*(3), 458-469.

Smallwood, J., Davies, J.B., Heim, D., Finnigan, F., Sudberry, M., O'Connor, R., et al. (2004). Subjective experience and the attentional lapse: Task engagement and disengagement during sustained attention. *Consciousness and Cognition, 13*, 657-690.

Smallwood, J., & Schooler, J.W. (2006). The restless mind. *Psychological Bulletin, 132*(6), 946-958.

Smith, R. L. (1966). *Monotony and motivation: A theory of vigilance*. Los Angeles, CA: Dunlop and Associates.

Strayer, D. L., Drews, F.A., & Johnston, W.A. (2003). Cell phone-induced failures of visual attention during simulated driving. *Journal of Experimental Psychology: Applied, 9*(1), 23-32.

Strayer, D. L., & Johnston,W.A. (2001). Driver to distraction: Dual-task studies of simulated driving and conversing on a cellular telephone. *Psychological Science, 12*(6), 462-466.

Stroh, C. M. (1971). *Vigilance: the problem of sustained attention*: Pergamon Press.

Stuss, D. T., Shallice, T., Alexander, M.P., & Picton, T.W. (1995). A multidisciplinary approach to anterior attentional functions. *Annals of the New York Academy of Sciences, 769*, 191-211.

Thiffault, P., & Bergeron, J. (2003). Monotony of road environment and driver fatigue: a simulator study. *Accident Analysis and Prevention, 35*, 381-391.

Ware, J. R. (1961). Effects of intelligence on signal detection in visual and auditory monitoring. *Perceptual and Motor Skills, 13*, 99-102.

Warm, J. S., & Dember, W.N. (1998). Tests of vigilance taxonomy. In R. Hoffman, Sherrick, M.F., & Warm, J.S. (Ed.), *Viewing psychology as a whole: The* *integrative science of William N. Dember*. Washington, DC, USA: American Psychological Association.

Warm, J. S., Dember, W.N., Murphy, A.Z., & Dittmar, M.L. (1992). Sensing and decision-making components of the signal-regularity effect in vigilance performance. *Bulletin of the Psychonomic Society, 30*(4), 297-300.

Warm, J. S., Epps, B.D., & Ferguson, R.P. (1974). Effects of knowledge of results and signal regularity on vigilance performance. *Bulletin of the Psychonomic Society, 4*, 272-274.

Warm, J. S., & Jerison, H.J. (1984). The psychophysics of vigilance. In J. S. Warm (Ed.), *Sustained attention in human performance* (pp. 15-51). New York: John Wiley & Sons.

Warm, J. S., Parasuraman, R., & Matthews, G. (2008). Vigilance requires hard mental work and is stressful. *Human Factors, 50*(3), 433-441.

Wickens, C. D. (1984). Processing Resources in Attention. In R. Parasuraman, & Davies, D.R. (Ed.), *Varieties of Attention*. Orlando, FL, USA: Academic Press Inc.

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science, 3*(2), 159-177.

Wiener, E. L. (1969). Money and the monitor. *Perceptual and Motor Skills, 29*, 627-634.

Wilkins, A. J., Shallice, T., & McCarthy, R. (1987). Frontal lesions and sustained attention. *Neuropsychologia, 25*(2), 359-365

Appendix A: Word List for Conversation Task.

1. Adorable
2. Agreement
3. Angel
4. Bath
5. Beauty
6. Bed
7. Bird
8. Bless
9. Blossom
10. Brother
11. Bunny
12. Butterfly
13. Cake
14. Capable
15. Carefree
16. Caress
17. Color
18. Comfort
19. Cozy
20. Cuddle
21. Devoted
22. Dignified
23. Earth
24. Easy
25. Easygoing
26. Elegant
27. Enjoyment
28. Fantasy
29. Friendly
30. Gentle
31. Grateful
32. Grin
33. Heal
34. Honest
35. Kindness
36. Loyal
37. Luxury
38. Masterful
39. Melody
40. Nature
41. Ocean
42. Palace
43. Paradise
44. Pillow
45. Politeness
46. Protected
47. Rainbow
48. Refreshment
49. Respectful
50. Reward
51. Safe
52. Sailboat
53. Satisfy
54. Scholar
55. Secure
56. Sky
57. Sleep
58. Snuggle
59. Soft
60. Soothe
61. Soothing
62. Spouse
63. Sun
64. Sunrise
65. Sunset
66. Twilight
67. Untroubled
68. Useful
69. Warmth
70. Wise
71. Wish

The page left intentionally blank