12-2016

The effects of time of day and circadian rhythm on performance during variable levels of cognitive workload

Kathryn A. Feltman
University of North Dakota

THE EFFECTS OF TIME OF DAY AND CIRCADIAN RHYTHM ON PERFORMANCE DURING VARIABLE LEVELS OF COGNITIVE WORKLOAD

by

Kathryn Ann Feltman
Bachelor of Science, University of North Dakota, 2010
Master of Arts, University of North Dakota, 2014

A Dissertation

Submitted to the Graduate Faculty

of the

University of North Dakota

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Grand Forks, North Dakota

December
2016

This dissertation, submitted by Kathryn Feltman in partial fulfillment of the requirements for the Degree of Doctor of Philosophy from the University of North Dakota, has been read by the Faculty Advisory Committee under who the work has been done and is hereby approved.

Richard F. Ferraro, PhD

Thomas Petros, PhD

Dmitri Poltavski, PhD

Adam Derenne, PhD

Warren Jensen, MD

This dissertation is being submitted by the appointed advisory committee as having met all of the requirements of the School of Graduate Studies at the University of North Dakota and is hereby approved.

Grant McGimpsey
Dean of the School of Graduate Studies

Date

PERMISSION

Title	The Effects of Time of Day and Circadian Rhythm on Performance during Variable Levels of Cognitive Workload
Department	Psychology
Degree	Doctor of Philosophy

In presenting this dissertation in partial fulfillment of the requirements for a graduate degree from the University of North Dakota, I agree that the library of this University shall make it freely available for inspection. I further agree that permission for extensive copying for scholarly purposes may be granted by the professor who supervised my dissertation work or, in his absence, by the Chairperson of the department or the dean of the School of Graduate Studies. It is understood that any copying or publication or other use of this dissertation or part thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of North Dakota in any scholarly use which may be made of any material in my dissertation.

Kathryn Feltman

November 8, 2016

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Table Page

ACKNOWLEDGEMENTS

I wish to express my appreciation to the members of my advisory committee for the guidance, patience, and tolerance of me during my time in the doctoral program at the University of North Dakota. I also wish to express my gratitude to my father for having patience with me throughout the entire process of graduate school. I could not have done it without his never ending support.

ABSTRACT

The present study examined the effects of time of day of testing on a simulated aviation task. The tasks required the participants to engage in multitasking while electroencephalogram (EEG) data was collected to objectively measure participants' workload. Task demands were altered throughout the testing period to expose participants to both high and low workload conditions. Additionally, individual differences in circadian rhythm were explored by assessing participants' circadian typology. No significant differences in performance were found resulting from time of day differences. However, performance and EEG differences were found based on phase of testing and workload manipulations. Subjective workload measures were influenced by time of day, with a moderating effect of circadian typology. Implications are discussed.

CHAPTER 1

INTRODUCTION

Many human biological processes, ranging from gene expression to behavior, follow a natural rhythm, with fluctuations occurring throughout a 24-hour period that influence alertness and cognitive performance (Duguay & Cermakian, 2009; Jasper et al., 2010). In addition to experiencing regular fluctuations in circadian rhythm, many individuals will also demonstrate variations in time of day preference, where their task performance and alertness will peak at certain times during the day and decrease at others (Schmidt et al., 2007; Taillard, Philip, & Bioulac, 1998). These variations in cognitive performance and alertness can impact how well an individual is able to accurately complete tasks, such as those requiring attention for sustained periods and multitasking.

Scheduling needs within various occupational fields often require employees to work shifts around a 24-hour period, and as a result many are required to work during times incompatible with their own circadian typology, such as a ‘morning person’ working night shifts. Pilots, both civilian and military, are often scheduled during nighttime hours or early morning hours, and this regularly occurs with little opportunity to rest in between flights. When one is required to make changes in sleep cycles and is given little time to recooperate, circadian desynchronization can result.This desynchronization may cause fatigue and errors inperformance as the individual needs to maintain wakefulness despite a feeling of sleepiness (Winget, DeRoshia, Markley, & Holley, 1984).

In addition to the experience of an increase in errors from incompatible scheduled times, pilots are exposed to variable workload levels throughout a flight, which also creates difficulties for performance (Caldwell, 2004; Wilson, Caldwell, & Russell, 2007). Workload levels will vary throughout a flight with pilots typically experiencing a high workload level during the take-off and landing phases, and a low workload level during the cruise phase (Di Nocera, Camilli, & Terenzi, 2007). Periods of high workload are typically known for being most problematic for pilot performance, since extended periods of high workload diminish cognitive resources and increase the likelihood of errors (Warm, Matthews, & Parasuraman, 2008). However, pilots who are fatigued have also been shown to demonstrate poor performance and increased errors during the cruise phase, when workload is lower (Cabon, Coblentz, Mollard, & Fouillot, 1993). Poor performance during the cruise phase is seen most frequently in long-haul flights, when pilots are fatigued from circadian rhythm disruptions as a consequence of the flight schedule (Caldwell, 2004). Such decrements in performance resulting from fatigue increases the likelihood of accidents, with circadian rhythm disruptions having been identified as a causal factor in multiple, recent aviation accidents, indicating that this remains an issue for the aviation community despite changes to work-rest regulations (National Transportation Safety Board, 2014a, 2014b).

Scheduling of pilots in aviation remains problematic, with many pilots scheduled to work hours incompatible with their preferred times of the day. Furthermore, the Federal Aviation Administration has recently reexamined flight regulations for pilots; however, the changes brought forth continue to focus on work-hour limits rather than on sleep and circadian factors (Caldwell, 2012). Sleep and individual circadian factors should be taken into consideration when adjusting crew scheduling, as these factors are commonly the main cause of pilot fatigue

(Caldwell, 2001, 2004). Many of the currently available scheduling tools, such as the Fatigue Avoidance Scheduling Tool (FAST), which is commonly used in military and civilian aviation scheduling, do not account for individual differences, such as circadian typology. Additionally, tools such as FAST have been based on mathematical fatigue models, which are based on performance changes on reaction time tasks, such as the psychomotor vigilance task, and cognitive tests, such as arithmetic (Hursh et al., 2004). Not accounting for individual differences and consideration of the dynamic and complex tasks performed by aviators when in fatigued states have resulted in these tools not fully resolving the issues that remain in regards to pilot fatigue.

Pilot Fatigue

Both civilian and military pilots are prone to scheduling that can result in increased fatigue, with technological advances making early morning, late night, and overnight flights safe and commonplace. Two recent surveys of airline pilots have found that the experience of fatigue may be more commonplace than previously thought, with many short- and medium-haul pilots reporting high levels of fatigue, whereas this was previously considered to be more of an issue in long-haul pilots (Reis, Mestre, & Canhão, 2013; Roach, Sargent, Darwent, & Dawson, 2012). The increased experience of fatigue has been associated with earlier departure times that curtail the pilot’s sleep. Similarly, military aviation operations often occur during early morning hours. Operations during these timeframes are typically associated with a higher incidence of fatigue for pilots and aircrew members, since waking during extreme early morning hours, or remaining awake into the morning hours, does not usually coincide with one’s natural circadian rhythm (Rabinowitz, Breitbach, & Warner, 2009). In addition to working hours not coinciding with natural circadian rhythm, pilots' schedules can change frequently, without allowing sufficient time for the pilots' circadian rhythm to resynchronize. When pilots are not given enough time to resynchronize the likelihood of higher error rates from fatigue caused by temporary disruption of sleep cycles can increase (Caldwell, 2001, 2004).

Furthermore, many individuals who regularly experience circadian desynchronization and fatigue will typically underreport the actual extent of fatigue experienced and not recognize the increased inclination for errors (Van Dongen, Maislin, Mullington, & Dinges, 2003). Unawareness of one's current fatigued state can be problematic in terms of determining whether or not one is suitable to fly. For example, the disruption in sleep cycles and frequent changing of schedules can result in the occurrence of microsleeps, which is when an individual falls asleep for a very brief period, oftentimes unaware of having fallen asleep (Wright & McGown, 2001). Microsleep occurrences have been attributed to crewmembers not having sufficient time between flights to adjust to disruptions in circadian timing and obtain adequate sleep, and will often go unnoticed by the pilot (Wright & McGown, 2004). Microsleeps going unnoticed can be problematic if a problem arises during flight with the pilot remaining unaware of it and unable to respond properly. Aviation mistakes due to fatigue-related problems, such as microsleeps, can be costly. The cost of a major civilian accident can often exceed $500 million, as well as present the potential for the loss of lives (Caldwell, 2004). There are several documented flight accidents where crew fatigue, resulting from long duty hours and disruption to circadian rhythms, have been implicated as a causal factor (Caldwell, 2004; National Transportation Safety Board, 2014a, 2014b).

In order to address the issue of the effects of fatigue on pilot performance, scheduling tools such as the Fatigue Avoidance Scheduling Tool (FAST; Hursh, Balkin, Miller, & Eddy, 2004) have been developed based on fatigue models that have been validated using simple vigilance (e.g., psychomotor vigilance task) and cognitive taks (e.g., arithmetic). Additionally, field-deployable versions of the psychomotor vigilance task have been created to provide a quick assessment of performance and fatigue levels prior to allowing an individual to fly (Lamond, Dawson, & Roach, 2005). However, simple vigilance and cognitive tasks may not provide an accurate assessment of actual performance in regards to aviation tasks. The workload pilots experience will typically take the form of various visual and auditory stimuli which they must monitor and respond to, as well as monitor for environmental cues outside and within the aircraft that may affect the progress of the flight (Lee & Liu, 2003). Maintaining these various tasks can create fluctuations in the workload experienced by the pilot, and these fluctuations may affect performance in differing ways. For example, high levels of workload place demands on the pilot's cognitive capabilities or resources, which may result in performance errors as pilots' cognitive resources attempt to keep up with the demands (Wilson, 2002). Periods of low workload can create problems for pilots as well, since the amount of cognitive resources during a low workload period may be reduced (Stanton, Young, & McCaulder, 1997). After lowering the level of cognitive resources, pilots may experience difficulties when workload demands unexpectedly increase during flight (Morris & Leung, 2006). Furthermore, scheduling tools such as FAST, while accounting for variations in circadian rhythm, do not account for individual differences in response to such variations.

Circadian Rhythm

Variations in performance on cognitive tasks based on time of day differences have been recognized since Ebbinghaus (1885, 1964) first noted that individuals learn nonsense syllables better in the morning than in the evening. The variability in performance on such cognitive tasks has been attributed to changes in the sleep-wake cycle that occurs throughout the day. The sleep-wake cycle is controlled by two systems, the circadian timing process and homeostatic process, working either in synchrony or in opposition of one another throughout the 24-hour cycle to promote wakefulness and to increase sleepiness (Schmidt et al., 2007).

The circadian timing process is connected to the 24-hour day cycle and is influenced by the light-dark cycle each day. This process tends to coincide with the light-dark cycle to allow an individual to engage in activities during the light periods of the day. The circadian rhythm then results in specific sleeping and waking times, bodily temperature fluctuations, and differing levels of cognitive functioning to occur in synchrony with the light-dark cycle (Rogers, Dorrian, & Dinges, 2003). The circadian timing process works through the suprachiasmatic nuclei (SCN) of the anterior hypothalamus, which is influenced by the light-dark cycle environmental cues and generates the circadian rhythms that occur throughout the day (Rogers, Dorrian, & Dinges, 2003).

The homeostatic process corresponds with the amount of time spent awake, where longer periods of wakefulness create a greater pressure for sleep. Therefore, when individuals experience sleep deprivation, the homeostatic need for sleep increases, which is further associated with a decrease in alertness and an increase in fatigue levels (Maire, Reichert, & Schmidt, 2013).The homeostatic process is considered a sleep-promoting process, with sleep pressure continuously accumulating with time spent awake. These increases in sleep pressure during periods when an individual would normally be asleep, is often associated with changes in alertness and poor performance (Rogers, Dorrian, & Dinges, 2003).

The extent to which one is fatigued or alert at any given time is determined by the interaction of both the circadian timing process and the homeostatic process (Schmidt et al., 2007). These two processes work together by essentially counterbalancing one another. As the homeostatic process increases throughout the day, and a person experiences more sleep pressure, the circadian timing process will assist in keeping the individual awake through the daylight hours. The two processes will work together to promote sleep if the amount of time spent awake increases past normal sleeping hours, such as when an individual is required to stay awake beyond normal sleep hours. Although all individuals are affected by the circadian timing and homeostatic processing to maintain wakefulness and increase sleepiness, there are also large variations and individual differences in time-of-day fluctuations and sleep-wake preferences (Schmidt et al., 2007).

Time of Day

Individuals who demonstrate time of day preferences have been categorized as morningness (M-types) or eveningness (E-types) types (Natale & Cicogna, 2002; Taillard, Philip, & Bioulac, 1999), and are most frequently assessed using the Morningness-Eveningness Questionnaire (MEQ; Horne & Östberg, 1976). Time of day preferences have been known to affect alertness patterns, with M-type tending to have peak alertness in the morning and early day hours, and E-type in the late afternoon and late day hours (Schmidt et al., 2007; Taillard, Philip, & Bioulac, 1998). Differences in circadian typology also reflect differences in ability to cope with sleep deprivation effects, as well as with the ability to maintain wakefulness during normal sleep hours. Those who have a tendency toward the eveningness typology have typically been shown to adapt to shift work better than morningness individuals, particularly when scheduled the later shifts (Buschkens, Graham, & Cottrell, 2010; Griefahn, 2002).

Furthermore, individuals who demonstrate to be E-type have also been shown to report poorer sleep quality, as reported on the Pittsburgh Sleep Quality Index, than individuals reporting as M-type, which may translate into poorer performance on various tasks requiring attention and memory (Buysse, Reynolds, Monk, Berman, & Kupfer,1989; Roeser, Meule, Schwerdtle, Kübler, & Schlarb, 2012; Wittman, Dinich, Merrow, & Roenneberg, 2006). Given that working variable schedules is associated with circadian desynchronization, individuals who tend toward E-type or M-type may be more vulnerable to poorer sleep quality when working such schedules if these schedules do not coincide with performed time of day. The poor sleep quality may in turn lead to poor attention and memory performance.

The preference toward the M- or E-type is known to affect performance on cognitive tasks, by either improving or diminishing performance, depending on the time of day the testing occurs (West, Murphy, Armilio, Craik, & Stuss, 2002). Individual who are classified as an M- or E-type have been found to be more prone to making errors when completing tasks during times that are incompatible with their specific circadian typology (Schmidt et al., 2007). For example, one study found a time of day effect on the ability to recall passages. In their study of the influence of time of day on immediate recall of short passages, Petros, Beckwith, and Anderson

(1990), found that both M- types and E-types recalled the most when tested during times compatible with their typology (i.e., E-type tested in the afternoon), as well as recalled less when tested during incompatible times (i.e., M-type tested in the afternoon). These results support the influence that circadian typology has on performance abilities at differing points in the day.

While clear differences between the two extremes have been found, the majority of the population does not fall into the extremes, but instead fall somewhere in the middle. However, it has been demonstrated that when the MEQ scores are examined as raw scores (versus placing individuals into a category), it is possible to determine toward which end of the continuum (morningness or eveningness) an individual is classified. The MEQ scores also appear to vary with alertness levels (Natale & Cicogna, 2002), which could have implications for the optimal time to schedule pilots. This information, taken with the variability in cognitive performance seen in individuals throughout the morningness-eveningness continuum, can be used to consider individual differences in regards to scheduling to pilots work-rest cycles, particularly when considering tolerance to fatigue (Caldwell, 2012). While circadian typology is known to affect differences in fatigue tolerance and cognitive performance, its effect on differing levels of workload are less well-known.

Workload

Workload, within the context of aviation, is most frequently defined as “the combination of task demands, or load factors, and the operator’s response” (Mouloua, Gilson, & Hancock., 2003, p. 162). Determining the effects of workload on pilot behavior is often difficult due to the differences in operating during a high or low workload, and the switching between periods of high and low workload throughout flights. For example, pilots often face variable periods of workload throughout the duration of a flight, with departure and landing typically resulting in the highest workload and minimal workload experienced during the cruise phase of the flight (Di Nocera, Camilli, & Terenzi, 2007). Periods of extended high workload are often associated with performance decrements resulting from stress experienced by the pilot, but periods of low workload can be associated with boredom and decrease a pilot's performance, thereby increasing the potential for errors to occur (Miller & Parasuraman, 2007; Mouloua et al., 2003).

Additionally, pilot workload has been affected by changes in cockpit design. The implementation of the 'glass cockpit', which displays the instrument panel on one computer screen, has allowed for a decluttered instrument panel, but has also affected workload experienced by pilots. The use of the 'glass cockpit' now requires pilots to navigate the screen in order to locate pertinent information and has resulted in the increasing of workload experienced (Salas, Jentsch, & Maurino, 2010). Glass cockpits assist in reducing the amount of information to be observed at one time, but increase workload by requiring the pilot to navigate through the display to find necessary information. Technological advances such as the incorporation of glass cockpits continue to require additional research to examine and understand how these changes affect pilot performance.

The variability in workload experienced during flights, in addition to takeoff/landing and cruise phases, also results from the use of automated systems to control the aircraft. The use of automated systems affects workload by shifting the tasks controlled by the pilot, by the pilot being required to monitor the overall activity versus being in manual control of the activity (Parasuraman, 2000). Thus, the pilot will experience a new type of workload, which often places a greater demand on the pilot's information processing capacity (Warm, Dember, & Hancock, 1996; Warm, Parasuraman, & Matthews, 2008). Additionally, pilots using automation to assist in controlling the flight will occasionally experience automation complacency. Automation complacency occurs when the pilot relies too heavily on the automation to alert errors, and this overreliance can result in pilots ignoring other information that may be useful in indicating errors or automation failures (Molloy & Parasuraman, 1996).

Monitoring the overall activity, rather than being in manual control, also requires the pilot to maintain vigilance throughout the duration of the flight. This maintenance of vigilance can have a negative effect on performance, by vigilance decrements taking place (Warm et al., 1996). The maintenance of vigilance for an extended period of time is associated with a high workload, and many will experience vigilance decrements, where performance decreases over time (Gunn et al., 2005; Johnson & Proctor, 2004). It has also been reported that if the event rates are low and infrequent during a vigilant period, lower levels of workload are experienced, but performance can decline as well (Warm et al., 1996; Wiggins, 2011).

Physiological Monitoring

Electroencephalogram (EEG) measures have frequently been used to measure cognitive states in individuals. EEG measures the electrical activity of nerve cells of the brain through electrodes placed on the scalp (Zillmer, Spears, & Culbertson, 2008). The EEG will record the frequency of signal strength of neural activity, which ranges from 1 to 100 Hz, and these are separated into specific waveform patterns. Waveforms, or bands, falling within 35 Hz and above are considered gamma waves, and are most often associated with peak performance and hyper-arousal. Waves falling between 18 and 35 Hz are considered high beta waves and are associated with narrow focus, over-arousal, and anxiety. Mid-beta waves fall between 15 and 18 Hz and are associated with being active, alert, excited, or focused. Low beta waves fall between 12 and 18 Hz and are associated with a relaxed state. Alpha waves range from 8 to 12 Hz and are predominantly seen as background activity in wakeful individuals and associated with quiet, passive, resting states. Theta waves will range from 4 to 7 Hz and are mostly frequently seen in drowsiness and deeply relaxed states. Delta waves range from 0.5 to 4 Hz and are seen during sleep. EEG measures are often used to objectively determine alertness in individuals, as a means of supporting subjective measures of alertness.

Many individuals who experience chronic sleep deprivation become accustomed to the fatigued state, and as a result underreport the actual level of fatigue experienced (Balkin et al., 2008; Dinges, 2004). One study reported subjects who experienced chronic sleep deprivation, lasting a period of 14 days, demonstrated poorer performance on a psychomotor vigilance task as days of sleep deprivation accumulated, but reported low levels of fatigue during this time (Van Dongen et al., 2003). This finding suggests many individuals remain unaware of actual fatigue states and are unaware of its influence on performance. Since many individuals will unknowingly underreport fatigue levels, EEG measures have often been used as a means to objectively measure fatigue. EEG measures of fatigue have found increases in delta and theta bands are commonly seen as an individual becomes fatigued, and these increases are mostly seen in frontal and central brain areas (De Gennaro et al., 2007; Makeig & Jung, 1995). Additionally, an increase in alpha bands in frontal and parietal areas with a simultaneous decrease in beta bands is characteristic of a transition from an awake and alert state to a fatigued and drowsy state (Lal & Craig, 2000, 2002).

In addition to EEG measures, body temperature and heartrate have also been shown to vary with circadian rhythm and alertness. Body temperature fluctuations throughout the day occur in conjunction with circadian rhythm changes, Lower body temperatures are also associated with higher levels of sleepiness (Rogers et al., 2003). Body temperature has also been shown to have peak differences in M-type versus E-type individuals, supporting the differences in time of day preferences shown by M- and E-types (Bennett et al., 2008). Additionally, heart rate varies throughout the day, in accordance with the circadian rhythm (Huikuri et al., 1990; Massin et al., 2000), as well as blood pressure (Coca, 1994).

Objective measures of workload and fatigue provide valuable information to support or supplement subjective measures. Previous research has found that EEG data can be used to identify changes in an individual's cognitive state and are associated with task events (Berka et al., 2007). In order to assess an individual's cognitive state, Stikic and colleagues (2011) have developed EEG algorithms which are individualized from a participant's baseline data on three tasks, and are able to categorize second-by-second performance by giving a probability of engagement, workload, distraction, and sleep onset. By using the information provided from baseline tasks, the algorithm is able to give individualized estimates of probability for each cognitive state when a participant is performing a task. These algorithms have previously been demonstrated to detect cognitive state in conjunction with performance changes a simulator driving study (Marcotte, Meye, Hendrix, & Johnson, 2013) and in a real flight (Klyde et al., 2013).

While many individuals' performance will vary with changes in alertness influenced by circadian rhythm throughout the day and based on personal time-of-day preferences, the experience of high and low periods of cognitive workload can also impact performance. The experience of variable workload levels is known to create difficulties for the maintenance of flight performance (Di Nocera, Camilli, & Terenzi, 2007). The successful operation of an aircraft requires the pilot to manage varying workload levels, while simultaneously coping with the potential experience of fatigue.

Two prominent theories have been developed to address how individuals are able to maintain performance during cognitive workload tasks, as well as explanations for performance deteriorating during such tasks. The multiple resource theory of attention (Wickens, 2002, 2008) considers performance in terms of a fixed set of cognitive resources that are either shared during completion of a task or the task requires the use of differing resources. The malleable attentional resources theory (Young & Stanton, 2001) considers performance in terms of changing resources that adjust to the presented task demands, by increasing or decreasing in accordance with the task demands.

Multiple Resource Theory of Attention

The multiple resource theory of attention (Wickens, 2002) has frequently been used to describe the difficulties individuals experience when completing high workload tasks. Pilots often engage in multiple activities at one time when flying an airplane. When engaging in these multiple activities, varying demands will require them to utilize various cognitive resources, as well as share multiple cognitive resources to successfully complete the tasks at hand (Wickens, 1980, 2002). The sharing of multiple resources has been attributed as a causal factor in decreased performance during high workload tasks. The multiple resource theory of attention consists of four dichotomies of information processing in which a person may engage in while involved in an attention-demanding task. These dichotomies include ‘stages of processing dimension’, ‘codes of processing’, ‘modalities dimension’, and ‘visual channels’. The multiple resource theory of attention postulates that an individual’s performance will vary depending on the extent to which resources are being shared amongst these dichotomies (Wickens, 2008).

According to Wickens (2008), the stages of processing dimension states that perceptual and cognitive tasks will use different resources than the resources that are used for the selection and execution of action. This dimension has been supported with research demonstrating different brain regions responsible for perceptual and cognitive activity than for motor activity. The codes of processing dimension specifies spatial activity will use different resources than verbal activity (Baber, 1991). The modalities dimension implicates that auditory perception utilizes different resources than does visual perception. The modalities dimension has been supported with research demonstrating improved performance when task monitoring is split between auditory and visual stimuli, rather than all visual or all auditory stimuli (Wickens, Sandry, & Vidulich, 1983). The visual channels dimension distinguishes between focal and ambient vision, as the two use differing resources. Focal vision is used in object recognition, such as reading; whereas ambient vision is used in perception of orientation and movement. The visual channels dimension has been supported with these two types of vision utilizing different brain pathways (Previc, 1998).

Performance will typically remain intact as long as the presented task or tasks utilize differing resources, instead of sharing resources. However, fatigue has been implicated in depleting resource availability, and this depletion of resources has been identified as a factor behind performance decrements often seen in fatigued states (Warburton, 1986). Therefore, a fatigued individual who is already sharing resources to complete a task, may also experience depletion of resources, which could result in further performance decrements. Additionally, fatigue may play a role in an individual’s ability to allocate resources properly to maintain performance (Matthews & Desmond, 2002). Consequently, an individual in a fatigued state may experience difficulties in allocating resources to maintain performance on a difficult task.

While the multiple resource theory of attention provides a thorough explanation of the performance decrements that occur during high workload vigilance situations, it does not account for the performance decrement that often occurs in low workload situations. According to the multiple resource theory of attention, one would not expect to see performance decrements during low workload situations. Instead, it would be expected that performance would be improved as the individual would have additional resources available to use for task completion. However, performance decrements are often noted during periods of low workload (Warm et al., 1996; Wiggins, 2011). This suggests additional factors are causing the performance decrement, rather than just having ‘enough’ resources available to maintain performance (Young et al., 2015).

Malleable Attentional Resource Theory

The malleable attentional resources theory (MART; Young & Stanton, 2001) addresses the shortcomings of the multiple resource theory in regard to low workload situations. The MART, similar to the multiple resource theory, posits attention depends upon the availability of various cognitive resources. However, MART differs from the multiple resource theory by asserting that the available resources are malleable, instead of fixed, and will adjust depending on the presented task demands.

According to MART, attentional resources will shrink in order to accommodate a reduction in task demands (Young & Stanton, 2001). This theory states that as the task demands lessen, the resources used to complete the task will also temporarily lower, as not all resources are necessary to maintain performance. Performance errors will then occur when additional tasks are added or if current task demands increase, such as when turbulence is experienced during the cruise phase of a flight. Following the reduction in resources, the pilot may no longer have a sufficient pool of resources to adjust to the elevated tasks demands, and performance will remain degraded until enough resources are recruited to address the task demands (Young & Stanton, 2001). Fatigue may further increase the likelihood of the errors, with the fatigued individual having difficulties in adjusting resources needed to complete the task, particularly during periods of low workload (Matthews & Desmond, 2002).

Present Study

The present study examined whether time of day had a significant effect on performance during a simulated aviation task that required participants to engage in multitasking by responding to and monitoring four simultaneously occurring tasks. The current study was designed to address the following hypotheses:

Hypothesis One: Based on the multiple resource theory of attention, it was expected participants would perform worse during the high workload conditions as compared to the low workload conditions, particularly as time on task increased and resources became depleted (Wickens, 2002, 2008). However, based on the malleable attentional resources theory, it was expected that resources would shrink during the periods of low workload and performance would decrease. This decrease would initially continue into the high workload period, until participants recruited additional resources to meet the task demands (Young & Stanton, 2001). It was further expected that as time on task increased, worse performance would be seen during the latter half of the testing period than in the beginning, as time on task would exhaust participants' resources. It was also expected that participants' EEG classifications would vary in accordance with workload conditions and performance, such that participants would demonstrate a higher probability of workload classification and high engagement during the high workload conditions, based on EEG cognitive state classification algorithms (Stikic et al., 2011).

Hypothesis Two: It was expected participants who reported higher levels of sleepiness and poorer sleep quality would demonstrate worse performance throughout the testing period compared to those who reported lower levels of sleepiness and better sleep quality (De Gennaro, Ferrara, Curcio, & Bertini, 2001; Durmer & Dinges, 2005; Schmidt et al., 2007). It was further expected that EEG classifications would vary in accordance with sleepiness and sleep quality levels, such that those with higher daytime sleepiness and poorer sleep quality would have elevated classification probabilities for low engagement and distraction.

Hypothesis Three: Circadian typology would moderate time of day differences in performance. Individuals who tested during times conducive with their typology would perform better than those who tested during their non-preferred time, and this is was expected to occur across both workload conditions (West, Murphy, Armilio, Craik, & Stuss, 2002). It was expected that body temperature, heart rate, and blood pressure would vary between morning testing to late afternoon testing, in accordance with circadian rhythm (Bennett et al., 2008; Coca, 1994; Huikuri et al., 1990).

CHAPTER 2

METHODS

Research Design

The study used a 2 (Time of Testing: morning vs. late afternoon) X 2 (Workload: high vs. low) X 3 (Period: beginning, middle, end) mixed factorial design, with time of testing as a between-subjects factor, and workload and phase as within-subjects factors. This design is similar to previous research that has examined effects of cognitive workload and circadian rhythm (Stark, Scerbo, & Mikulka, 2000; Wilson et al., 2007). The current study also examined circadian typology, reported daytime sleepiness, and reported sleep quality as factors that may moderate the impact of circadian changes in performance. EEG data was collected to objectively measure changes in workload.

Participants were randomly assigned to either a morning testing time (0800 or 0900) or an afternoon testing time (1500 or 1600). This was done to decrease the likelihood of participants only signing up to participate during times that coincide with their circadian typology, and to increase the likelihood of an even number of individuals testing during preferred time-of-day and non-preferred time-of-day. Previous research has demonstrated performance differences in individuals tested during these morning and afternoon timeframes (Roeser et al., 2012).

Participants

Participants consisted of undergraduate students at the University of North Dakota and were of mostly Caucasian ethnicity. A total of 60 initially participated in the study; however, due to dropping out or incomplete data, only 50 participants were used in data analyses. Twenty-four of the participants were male, and 26 female, with an average age of 20 years (min = 18 years; max = 42 years). Twenty-three of the participants completed the morning testing session and 27 completed the afternoon testing session. Participants all reported as non-tobacco users, and rated their overall health as average to slightly above average. Participants were not allowed to participate if they had a psychiatric diagnosis that compromises attention or suffered a traumatic brain injury. None of the participants self-reported disqualifying diagnoses.

Participants with flight experience were recruited to participate; however, equal numbers of flight experience and non-experience were not obtained. Fourteen participants reported having had flight experience, with an average of 102.75 flight hours (*SD* = 69.51; *min* = 0.25; *max* = 206). The majority of flight experience was in fixed-wing aircraft. Since the study was unable to obtain equal numbers of experienced and inexperienced pilots in the testing groups, experience was not examined as a factor.

Participants were recruited through fliers, emails, recruitment briefs, and an online research participation system through the psychology department. Psychology students who participated were compensated with extra course credit, and students with flight experience were compensated $10 per hour of participation.

Materials

Multi-Attribute Task Battery-II

The Multi-Attribute Task Battery II (MATB-II; Santiago-Espada, Myer, Latorella, & Comstock, 2011; Santiago-Espada, 2014) is an updated version of the Multi-Attribute Task Battery (MATB) developed by Comstock and Arnegard (1992). The MATB was designed to study operator performance and workload using simultaneously presented tasks that are generalizations of piloting tasks. Several studies have demonstrated that the MATB is a valid method for assessing aviator performance (Caldwell & Ramspot, 1998; Wilson, Caldwell, & Russell, 2007). The MATB-II has been shown to be a reliable tool for examining the effects of workload on cognitive resources (Parasuraman, Bahri, & Molloy, 1992). The task itself was designed only for research purposes and is not used a training tool in aviation; therefore none of the pilot participants should be familiar with the task. The MATB-II (see appendix A for a picture of the task) consists of four tasks that require constant monitoring and occasional actions to be performed by the operator. These tasks are systems monitoring, resource management, tracking, and communications task (described below).

Systems monitoring. The system monitoring task is divided into two subtasks, which consist of warning lights and scales. During a testing session there are two warning lights, one which the participant is to keep green for the duration of the run, and the other which is to remain the background color but will turn to red throughout a run. The participant is required to maintain the green light as green by pressing the F5 key whenever it turns to the background color. Also, the participant must monitor the second light by ensuring that it remains the background color by pressing the F6 key anytime that it turns to red.

The second portion of the systems monitoring task, which is monitoring scales, requires the participant to monitor four scales that move in an up and down fashion. Each scale has a "light" on it which the participant must monitor to ensure that these stay within the middle of the scale. When a light on a scale deviates from the middle towards the upper or lower end of the scale, the participant must correct it by pressing the function key that correspond to that scale using the keyboard (F1, F2, F3, F4). The participant's time to correct the problems that arise for each of the subtasks for the systems monitoring task are recorded within the program.

The systems monitoring task records the number of missed responses for both the light and scale, the reaction time when a correct response to a light or scale change is made, and the number of times a participant responds to a light or scale when no response is needed. Scores for each of these three areas will be averaged across each of the 10 minute periods of high and low workload to make comparisons across the 60 minutes of test session.

Resource management. The resource management task is a generalized fuel management system. There are six different fuel tanks that are labeled A-F. There are also eight pumps that feed into the various fuel tanks, and these are labeled 1-8. Tanks A-D also have their remaining fuel levels next to them, which are affected by fuel consumption and the actions the participant performs on the connected pumps. The fuel levels are updated every 2 seconds. For this task, the participant is required to maintain the fuel levels in tanks A and B within +/- 500 units of 2,500 units each. The goal is to maintain as close to 2,500 units as possible, but +/- 500 units of this is an acceptable range. The box that contains the fuel amount will turn red if the amount of fuel is above or below the acceptable range. To adjust fuel levels in tanks A and B, the participant needs to press the pump number on the keyboard to turn the pump ON; pressing the key again turns on the corresponding pump in order to transfer fuel to or from the tank. The pump will turn green when it is on, remain the background color when it is off, and turn to red when it is a “failed state” and is nonfunctional. Pump flow rates are also indicated on the screen so that the participant may determine which pump to activate in order to reach the acceptable range on the tank. The resource management task records fuel unit levels for each Tank A and Tank B every 30 seconds during testing. The amount that each tank is above or below the desired 2,500 units is recorded as well.

Tracking. The tracking task requires the participant to use a joystick to keep a target within the center of a box. The tracking task switches between manual mode or automatic mode. The tracking task states in the bottom right-hand corner which mode it is, by stating either “MANUAL” or “AUTO ON”. While in manual mode, the participant is required to manually use the joystick to keep the target within the center of the box. While on automatic mode the target will remain within the box, however, “automation failures” will occur in which the target will go outside the box and the tracking task will switch into manual mode, for which the participant will need to correct it by using the joystick to manually move the target back into place. The MATB-II will collect data by calculating the root mean square deviation of the target center point from the center point in pixel units at a 15 second interval.

The tracking task gives the root mean square deviation from the center point (RMSD-C) in pixel units for every 30 seconds while the tracking task is in manual mode. This indicates how far or close the participant maintained tracking on the center point. The RMSD-C will be averaged across the 10 minute periods for each of the high and low workload conditions, and compared across time.

Communications. The communication task requires that participants listen for messages to change the frequency of the radio. The messages take the form of aircraft call sign (repeated twice) “tune your (radio name) to (frequency name)”. The possible values for the radio name are COM1, COM2, NAV1, and NAV2. However, not all of the messages stated over the radio require a response from the participant. The participant only responses to message for his or her aircraft, which is “NASA504”. When the participant hears a message intended for his or her aircraft, the participant needs to change the radio to the stated frequency by clicking the circle next to the radio name. However, if a message comes across for a different aircraft, the participant should not adjust his or her radio.

The MATB-II will be used to manipulate levels of workload and to measure performance on the management of the workload. Previous researchers have demonstrated that manipulating the number of times the MATB-II tasks require the participant to respond influences the perceived amount of workload experienced by the participant (Stark, Scerbo, Freeman, & Mikulka, 2000). The MATB-II tasks will be manipulated to require a higher number of responses from the participants during the periods of high workload, and fewer responses will be required of the participants during periods of low workload.

The communications task gives output for whether the participant selected the correct radio frequency and whether the participant does not respond to the call to change the radio frequency. Reaction times are also given when the participant correctly responds to the call for a change in radio frequency. The output also indicates when a participant makes an incorrect response by selecting a radio frequency that was not called out to be changed.

Workload Rating Scale

Subjective workload will be assessed using the Workload Rating Scale (WRS; appendix B) which is presented within the MATB-II program and is based on the NASA Task Load Index (NASA-TLX, Hart & Staveland, 1988). The WRS uses a sliding scale to rate workload from 0 to 100 on six different subscales. The subscales are as follows: mental demand, physical demand, temporal demand, performance, effort, and frustration.

EEG

EEG data was collected using the B-Alert X-24 wireless wet electrode system with 20 channels corresponding to scalp locations according to the International 10-20 system (frontal channels: Fp1, Fp2, F7, F3, Fz, F4, F8; central channels: C3, Cz, C4, T3, T4; parietal and occipital channels: P3, POz, P4, T5, T6, O1, O2). This EEG system provides cognitive state classification algorithms (engagement, distraction, and workload) to be used for data analyses. These algorithms have been previously validated (Berka et al., 2007; Johnson et al., 2011) and allow for individualization and generalization of the classification data. While two workload classifications are provided by the system, only the data from the classification based on the forward digit span task was used in analyses, as this model has been found to fit approximately 85% of the population.

The workload classifications are derived using a linear discriminant function analyses (DFA) with two classes, high and low workload. EEG data from channels C3C4, CzPO, F3Cz, FzC3, and FzPOz are used to calculate the classification (Berka et al., 2007). The workload classification provides an indication of working memory load and processing, and provides a value ranging from zero to one, with values closer to one indicative of a higher probability of the participant experiencing a greater workload. The engagement and distraction classifications are derived from the same differential channels as the workload classifications, and also provide a numeric value ranging from zero to one, with values closer to one indicating a higher probability the participant is experiencing the given cognitive state. The engagement classification is associated with active attention and vigilance constructs, whereas the distraction classification is associated with the inability to maintain passive attention. Additionally, the system provides a cognitive state classification, where given values are associated with the participants current cognitive state classification (0.3 = distraction, 0.6 = low engagement, 0.9 = high engagement).

The classifications are individualized (except for the workload classifications, which are based on a generalized model) by the subject completing three baseline tasks prior to data collection. The data collected from the baseline tasks are used to individualize the algorithms by adjusting the centroids to provide the engagement and distraction probabilities (Marcotte et al., 2013). The three tasks are used to create a participant's baseline data, which are the three-choice vigilance task, eyes open task, and eyes closed task. The three-choice vigilance task requires participants to discriminate whether three presented stimuli match the target stimulus by pressing the right arrow key on the keyboard to respond 'no' and the left arrow key to respond 'yes'. The eyes open task requires participants to monitor the computer screen and press the spacebar when a red dot is presented every 2 seconds. The eyes closed task requires participants to close their eyes and press the spacebar when a tone is emitted every 2 seconds.

Pittsburgh Sleep Quality Index

The Pittsburgh Sleep Quality Index (PSQI; Buysse, Reynolds, Monk, Berman, & Kupfer, 1989; appendix C) measures participants' quality of sleep. The PSQI consists of 10 items that ask participants' questions regarding sleep quality over the past month. The PSQI generates seven component scores, with subscale scores ranging from zero to three: sleep quality, sleep latency, sleep duration, habitual sleep efficiency, sleep disturbance, use of sleeping medication, and daytime dysfunction. These component scores are combined to give a global score of subjective sleep quality, ranging from zero to 21, with higher scores representing poorer sleep quality.

Epworth Sleepiness Scale

The Epworth Sleepiness Scale (ESS; Johns, 1991; appendix D) is a questionnaire consisting of eight questions in which participants rate their chance of dozing off during a particular activity using a 4-point scale. The ESS measures daytime sleepiness.

Horne and Östeberg Morningness-Eveningness Questionnaire

The Horne and Östeberg Morningness-Eveningness Questionnaire (MEQ; Horne & Östeberg, 1976; appendix E) is a 19-item self-report questionnaire that asks participants to rate questions regarding preferred sleep and wake times. This questionnaire produces a score that ranges from 16 to 86, with scores of 41 and below indicating 'evening types,' scores of 59 and above indicating 'morning types,' and scores between 42 and 58 indicating 'intermediate types'.

Wechsler Adult Intelligence Scale-Revised

The Wechsler Adult Intelligence Scale-Revised (WAIS-R; Wechsler, 1981; appendix F), vocabulary section, will be used to assess verbal ability. Researchers will read a word to subjects and subjects will define or give a good synonym of the word, researchers will score answers on a 0 to 2-point scale, for a maximum score of 70.

Demographics

Participants will be requested to provide basic demographic information (appendix G), which will include age, sex, rank or year in school, any current medications taken, and caffeine consumption day of testing and weekly. Participants will also be asked to report flight hours for simulated, as pilot-in-command, flying under instrument flight rules, and flying under visual flight rules. Participants will report on total hours flown, as well as frequency within the past month.

Physiological Measures

Body temperature will be measured using a forehead thermometer. Blood pressure and heartrate will be measured using a blood pressure monitor machine.

Procedure

Participants were randomly assigned to either the morning (0800 or 0900) or afternoon (1500 or 1600) condition of the study. Participants were tested individually in the laboratory room. When participants arrived at the laboratory they gave written consent to participate in the study, and then completed the demographics questionnaire and the vocabulary subset of the Wechsler Adult Intelligence Scale (WAIS-R; Wechsler, 1981). Baseline physiological measures were taken to obtain body temperature, blood pressure, and heart rate. Following physiological measurements, participants watched a 10 minute instructional video on how to complete the MATB-II, and then partook in a 10 minute training session on the MATB-II.

After the completion of MATB-II training, the EEG headset and electrodes were placed onto the participant's scalp. Participants were asked to not chew gum or move excessively while the EEG headset was on, in order to avoid contaminating signals. Impedance tests were done once the electrodes have been applied to the scalp. Impedance values below 40kΩ were the minimum acceptable threshold to ensure optimal quality of data; however, if impedance tests are rerun and some channels remain above 40kΩ, but below 80 kΩ, they were considered acceptable (ABM B-Alert User Manual, 2014). Next, baseline measures of the participant's performance were taken. The baseline measures consist of the three-choice vigilance task, eyes open task, and eyes closed task. Impedance checks were done again prior to the beginning of data collection to ensure conductivity remains satisfactory.

During the 60 minute testing session, workload was varied throughout by altering conditions of high and low workload every 10 minutes. The variations in high and low workload were made by increasing task demands during the high workload conditions and decreasing task demands during the low workload conditions. This variation in levels of workload on the MATB-II were modeled off of several other studies that have examined workload using the MATB-II (Fairclough, Venables, & Tattersall, 2005; Wilson, Caldwell, & Russell, 2007), and were validated in a small pilot study. The high workload condition consisted of 10 systems monitoring tasks per minute, two resource monitoring failures per minute, the manual tracking task was set to high update (which increases the amount of random target movement per update cycle), and five communications tasks per minute. The low workload condition consisted of two systems monitoring tasks per minute, one resource monitoring failure per minute, the manual tracking task was set to low update, and one communications task per minute. The presentation of workload (low vs. high) was counterbalanced amongst subjects.

The workload rating scale was presented at the end of each 10 minute segment of high or low workload. After completing the testing session, the EEG headset was removed and participants completed the remaining questionnaires, and temperature, blood pressure, and heart rate will again be assessed. The total time of participation was approximately two to three hours.

CHAPTER 3

RESULTS

All data were first examined for normality. Outliers were identified through box and whisker plots. Outliers were then corrected using Winsorizing, by replacing with the upper or lower fence, depending on whether the outline fell on the lower or upper extreme. Data were also examined through Q-Q plots, and determined appropriate for parametric testing.

T-tests were also conducted to examine whether groups (morning or afternoon testing times) differed on the flight hours, WAIS, PSQI, MEQ, and ESS scores, as such group differences might impact performance. Significant differences were found between morning *(M* = 43.92, *SD* = 11.40) and afternoon *(M* = 52.93, *SD* = 10.59) groups on WAIS scores, $t = -2.98$, $p < .05$. This difference was likely due to a failure of random assignment. T-tests were also completed to examine the physiological data of heart rate, body temperature, and blood pressure as a function of the time of day groups. The pre- and post-test data from each of these measures were combined and averaged to provide one measure. Only temperature found a significant difference between groups, $t = -2.12$, $p < .05$, with temperatures lower during the morning (*M* = 97.96, *SD* = 1.13) than afternoon (*M* = 98.45, *SD* = .40). Correlations were also done for the flight hours, WAIS, PSQI, MEQ, and ESS scores with performance and EEG data. No significant correlations were found and therefore none of these were used in subsequent analyses as covariates. Separate analyses examining participants' performance, EEG, and subjective workload data are described below.

Performance Results

To examine the effects of workload conditions, phase of testing, and time of day on performance during the MATB-II tasks, separate mixed analysis of variance (ANOVAs) were completed for each of the tasks (resource monitoring, tracking, systems monitoring, and communications).

Resource Monitoring

During the resource monitoring task, measurements were taken every 30s of how far the participant maintained each tank above or below the goal of 2,500 units. These differences were then averaged across each of the ten minute workload phases to create the data to be analyzed. Resource monitoring performance was examined using a mixed ANOVA, with time of day as the between-subjects factor (two levels: morning, afternoon) and phase (three levels: beginning, middle, end) and workload (two levels: high, low) as within-subjects factors. Only the interaction of phase and workload was significant, $F(2, 96) = 3.48$, $p < .05$. Simple effects analysis of workload at each level of phase indicated that performance deviations were significantly greater in the high workload conditions than the low workload conditions during the first phase of testing, $F(1, 49) = 4.42$, $p < .05$ (see Table 1 for means).

Table 1.

Resource Monitoring Deviations by Phase and Workload.

	Phase One		Phase Two		Phase Three	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
High Workload	828.10	639.23	667.08	428.56	766.41	605.94
Low Workload	694.70	445.65	698.00	514.33	717.39	511.36

Tracking

For the tracking task, measurements were taken every 15s while the task was in "manual mode" of the root mean square deviation from the center point in pixel units to determine how close the participant was keeping the target on the center point. Tracking performance was then examined using a mixed ANOVA, with time of day as the between-subjects factor (two levels: morning, afternoon) and phase (three levels: beginning, middle, end) and workload (two levels: high, low) as within-subjects factors. The main effect for phase was significant, $F\ (2, 96) = 5.19$, $p < .05$. Pairwise comparisons using the Least Significant Differences (LSD) test show tracking deviations were significantly higher during the first phase ($M = 33.51$, $SD = 7.29$) than during the second ($M = 31.85$, $SD = 5.59$, $p < 05$) and third ($M = 31.09$, $SD = 7.50$, $p < .05$) phases. The main effect for workload was significant, $F\ (1, 48) = 186.16$, $p < .001$, participants tracking deviated more during the high workload ($M = 43.05$, $SD = 10.04$) conditions than the low workload ($M = 21.25$, $SD = 6.33$) conditions.

Systems Monitoring

The systems monitoring task records reaction times for every correct response to light or scale corrections, every missed response, and the number of false responses emitted, that is, pressing one of the buttons for the lights or scales when unnecessary. Data were analyzed for reaction times (RT) for correct responses by taking into account the number of false responses made through taking the RT and dividing it by the proportion of responses made (correct responses/opportunity for response + false alarms). Data were also separately analyzed for the number of missed responses.

Reaction time. Reaction time performance was examined using a mixed ANOVA, with time of day as the between-subjects factor (two levels: morning, afternoon) and phase (three levels: beginning, middle, end) and workload (two levels: high, low) as within-subjects factors. The main effect of phase reached significance, $F(2, 96) = 15.89$, $p < .001$. Pairwise comparisons using the LSD test show participants had significantly longer RTs during the first phase ($M = 3.47$s, $SD = .72$) than phase two ($M = 3.22$s, $SD = .80$, $p < .05$) or phase three ($M = 3.05$s, $SD = .76$, $p < .001$). Additionally, participants also had significantly shorter reaction times during phase three than during the phase two ($p < .05$). For the interaction of phase and workload Mauchly's test indicated that the assumption of sphericity had been violated, $\chi^2(2) = 13.17$, $p < .05$, therefore Greenhouse-Geisser corrected tests are reported ($\varepsilon = .80$). The interaction of phase and workload was significant, $F(1.61, 77.14) = 4.55$, $p < .05$. Simple effects analysis of workload at each level of phase indicated reaction times were significantly longer in the high workload condition than the low workload condition during the second phase, $F(1, 49) = 7.94$, $p < .05$ (see Table 2 for means).

Table 2.

Systems Monitoring Reaction Times by Phase and Workload.

	Phase One		Phase Two		Phase Three	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
High Workload	3.52	.69	3.33	.78	2.96	.64
Low Workload	3.42	.96	3.12	.89	3.13	1.00

Missed responses. The occurrence of missed responses was examined using a mixed ANOVA, with time of day as the between-subjects factor (two levels: morning, afternoon) and phase (three levels: beginning, middle, end) and workload (two levels: high, low) as within-subjects factors. The main effect of phase was found to be significant, $F(2, 96) = 13.58$, $p < .001$. Pairwise comparisons using the LSD test showed that participants missed significantly more alarms during the first phase ($M = 9.23$, $SD = 9.46$) than the second ($M = 7.66$, $SD = 8.73$, $p < .05$) and third ($M = 6.49$, $SD = 8.56$, $p < .001$) phases. Additionally, participants also missed significantly fewer responses during the third phase than the second phase ($p < .05$). The main effect of workload was significant, $F(1, 48) = 39.02$, $p < .001$, with significantly higher missed alarms during the high workload ($M = 15.91$, $SD = 16.36$) condition than during the low workload condition ($M = 4.88$, $SD = 2.98$). The interaction between phase and workload was significant, $F(1, 48) = 10$, $p < .001$. Simple effects analysis of workload at each level of phase indicated that missed responses during high workload conditions were significantly greater than during low workload conditions across all phases, $F(1, 49) = 47.30$, $p < .001$ (phase one); $F(1, 49) = 37.85$, $p < .001$ (phase two); and, $F(1, 49) = 25.62$, $p < .001$ (phase three; see Table 3 for means).

Table 3.

Systems Monitoring Missed Responses by Phase and Workload.

	Phase One		Phase Two		Phase Three	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
High Workload	15.92	16.23	13.39	15.24	11.03	14.84
Low Workload	2.54	3.22	1.94	2.57	1.96	2.61

Communication

The communications task records reaction times for every correct response to a radio frequency change and the number of false responses emitted, that is, changing a radio frequency for a call sign other than the one assigned. Data were analyzed for reaction times for correct responses by taking into account the number of false responses made through taking the RT and dividing it by the proportion of responses made (correct responses/opportunity for response + false alarms).

Communications responses were examined using a mixed ANOVA, with time of day as the between-subjects factor (two levels: morning, afternoon) and phase (three levels: beginning, middle, end) and workload (two levels: high, low) as within-subjects factors. Mauchly's test indicated that the assumption of sphericity had been violated for the main effect of phase, $\chi^2(2) = 15.26$, $p < .001$, therefore Greenhouse-Geisser corrected tests are reported ($\varepsilon = .78$). The main effect of phase was found to be significant, $F(1.57, 75.16) = 15.89$, $p < .001$. Pairwise comparisons using the LSD test showed participants' reaction times were longer during the first phase ($M = 2.51$s, $SD = 1.06$) than the second ($M = 2.10$s, $SD = .95$, $p < .05$) and third ($M = 1.87$s, $SD = .73$, $p < .05$) phases. Additionally, participants' reaction times were longer during the second phase than during the third ($p < .05$). The interaction between phase and workload approached, but did not reach significance, $F(2, 96) = 2.55$, $p = .08$.

EEG Results

Five subjects did not have EEG data due to technical difficulties with the EEG equipment during their data collection; therefore the EEG data includes 21 participants in the morning group and 24 in the afternoon group. To examine the effects of workload conditions and time of day on

EEG cognitive state classifications, separate mixed ANOVAs were completed for each cognitive state classification. Cognitive state classifications were averaged for each phase of workload to be used in analyses, and are reported below.

Cognitive State

The cognitive state metric provides a numerical value that represents the classification with greatest probability of the participant's state for each second of sampling. Cognitive states correspond with the following values: .1 = sleep onset, .3 = distraction, .6 = low engagement, and .9 = high engagement. Cognitive state was examined using a mixed ANOVA with time of day as the between-subjects factor (two levels: morning, afternoon), phase (three levels: first, second, third) and workload (two levels: high, low) as within-subjects factors. For the main effect of phase, Mauchly's test indicated that the assumption of sphericity had been violated, $\chi^2(2) = 21.05$, $p < .001$, therefore Greenhouse-Geisser corrected tests are reported ($\varepsilon = .72$). The main effect of phase was found to be significant, F (1.44, 61.68) = 10.95, $p < .001$. Pairwise comparisons using the LSD test yielded significant differences between the first phase (M = .7246, SD = .0617), and both the second (M = .7137, SD = .0585, $p < .05$) and third (M = .7097, SD = .0603, $p < .05$) phases. The main effect of workload reached significance, F (1, 43) = 6.77, $p < .05$, with participants' cognitive state value higher during the high workload (M = .7212, SD = .0631) conditions than during the low workload (M = .7108, SD = .0572) conditions.

Probability of Distraction

The probability of distraction indicates the likelihood of the participant being in a distracted state, with values closer to one indicating a higher probability. Probability of distraction was examined using a mixed ANOVA with time of day as the between-subjects factor (two levels: morning, afternoon), phase (three levels: first, second, third) and workload (two levels: high, low) as within-subjects factors. For the main effect of phase, Mauchly's test indicated that the assumption of sphericity had been violated, $\chi^2(2) = 47.37$, $p < .001$, therefore Greenhouse-Geisser corrected tests are reported ($\varepsilon = .60$). The main effect of phase was found significant, F (1.19, 51.31) = 7.79, $p < .05$. Pairwise comparisons using LSD show significant differences between phase one ($M = .0735$, $SD = .0938$) and phases two ($M = .0828$, $SD = .0927$, $p < .05$) and three ($M = .0907$, $SD = .0963$, $p < .05$). There were also marginally significant differences between phase two and phase three, $p = .05$. The main effect of workload was significant, F (1, 43) = 10.73, $p < .05$, with the probability of distraction lower during the conditions of high workload ($M = .0776$, $SD = .0923$) compared to the conditions of low workload ($M = .0870$, $SD = .0942$).

Probability of Workload

Probability of workload, where scores closer to one indicate a higher likelihood that the participant was experiencing a high workload at that given time, was examined using a mixed ANOVA, with time of day as the between-subjects factor (two levels: morning, afternoon) and phase (three levels: beginning, middle, end) and workload (two levels: high, low) as within-subjects factors. For the main effect of phase, Mauchly's test indicated that the assumption of sphericity had been violated, $\chi^2(2) = 10.25$, $p < .05$, therefore Greenhouse-Geisser corrected tests are reported ($\varepsilon = .82$). The interaction of phase and TOD was marginally significant, F (2, 86) = 3.04, $p = .05$. The main effect of workload was significant, F (1, 43) = 7.79, $p < .05$, with an increased probability of workload during the high workload ($M = .6588$, $SD = .0891$) conditions than the low workload ($M = .6472$, $SD = .0842$) conditions.

Probability of Low Engagement

Probability of low engagement, where values closer to one indicate a higher probability of low engagement during the sampling period, was examined using a mixed ANOVA, with time of day as the between-subjects factor (two levels: morning, afternoon) and phase (three levels: beginning, middle, end) and workload (two levels: high, low) as within-subjects factors. For the main effect of phase, Mauchly's test indicated that the assumption of sphericity had been violated, $\chi^2(2) = 10.73$, $p < .05$, therefore Greenhouse-Geisser corrected tests are reported (ε = .82). The main effect of phase was found significant, F (1.63, 70.18) = 3.71, $p < .05$. Pairwise comparisons using LSD show a significant difference between phase one (M = .3975, SD = .1100) and phase three (M = .4100, SD = .1170, $p < .05$). The interaction of phase and TOD was also significant, F (1.63, 70.18) = 3.54, $p < .05$. Pairwise comparisons using the LSD test show the probability of low engagement was significantly higher for the afternoon group compared to the morning group for both phases one and two ($p < .05$, see Table 4 for means). The interaction of workload and phase approached significance, F (1.66, 71.16) = 3.16, $p = .06$.

Table 4.

Probability of Low Engagement by Phase and Time of Day.

	Phase One		Phase Two		Phase Three	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Morning	.3774	.1145	.3955	.1228	.4064	.1398
Afternoon	.4141	.1077	.4208	.1075	.4134	.0989

Probability of High Engagement

Probability of high engagement, where values closer to one indicate a higher likelihood that the participant was highly engaged in the task at that given time, was examined using a mixed ANOVA, with time of day as the between-subjects factor (two levels: morning, afternoon) and phase (three levels: beginning, middle, end) and workload (two levels: high, low) as within-subjects factors. For the main effect of phase, Mauchly's test indicated that the assumption of sphericity had been violated, χ^2 (2) = 18.42, p < .001, therefore Greenhouse-Geisser corrected tests are reported (ε = .74). The main effect of phase was found significant, F (1.48, 63.47) = 7.96, p < .05. Pairwise comparisons using the LSD test show significant differences between phase one (M = .4727, SD = .1221) and both phase two (M = .4488, SD = .1157, p < .001) and phase three (M = .4471, SD = .1268, p < .05). The interaction of phase and TOD approached, but did not reach significance, F (1.48, 63.47), p = .09.

The main effect of workload was significant, F (1, 43) = 6.90, p < .05, with probability of high engagement elevated during the high workload (M = .4700, SD = .1333) conditions than during the low workload (M = .4424, SD = .1117) conditions. The interaction of phase and workload was significant, F (2, 86) = 3.38, p < .05. Simple effects analysis of workload at each level of phase indicated that probability of high engagement were significantly greater during the high workload conditions than during low workload conditions for phase two, F (1, 45) = 18.77, p < .001 and phase three, F (1, 45) = 5.82, p < .05 (see Table 5 for means).

Table 5.

Probability of High Engagement by Phase and Workload.

	Phase One		Phase Two		Phase Three	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
High Workload	.4781	.1381	.4685	.1370	.4663	.1526
Low Workload	.4693	.1242	.4307	.1018	.4293	.1214

Workload Rating Scale

The workload rating scale consists of six subscale scores (effort, frustration, mental demand, physical demand, temporal demand, and performance), and an overall mean score. Each subscale score ranges from 0 to 100, and were analyzed individually and reported below.

Effort. The effort subscale refers to how much effort the individual perceived he or she exerted during the task, with values ranging from 0 to 100, with 100 being indicative of maximum effort. The effort subscale was examined using a mixed ANOVA, with time of day as the between-subjects factor (two levels: morning, afternoon) and phase (three levels: beginning, middle, end) and workload (two levels: high, low) as within-subjects factors. For the main effect of phase, Mauchly's test indicated that the assumption of sphericity had been violated, $\chi^2 (2) = 8.42$, $p < .05$, therefore Greenhouse-Geisser corrected tests are reported ($\varepsilon = .86$). The main effect of phase was found significant, $F (1.72, 82.47) = 4.69$, $p < .05$. Pairwise comparisons using the LSD test show participants rated effort significantly higher during phase one (M = 61.27, SD = 16.60) than during phase three (M = 55.43, SD = 16.32, p < .05).

The main effect of workload was found significant, F (1, 48) = 48.43, p < .001. Participants rated effort higher during the high workload conditions (M = 64.87, SD = 14.12) than the low workload conditions (M = 51.36, SD = 17.70). The interaction of phase and workload approached, but did not reach, significance, F (2, 96) = 2.94, p = .06.

Frustration. The frustration subscale referred to participants' rating how they felt while performing the tasks, with ratings ranging from relaxed (0) to very stressed (100). The frustration subscale was examined using a mixed ANOVA, with time of day as the between-subjects factor (two levels: morning, afternoon) and phase (three levels: beginning, middle, end) and workload (two levels: high, low) as within-subjects factors. The main effect of workload was significant, F (1, 48) = 55.59, p < .001. Participants rated frustration significantly higher during the high workload conditions (M = 44.85, SD = 20.07) as compared to low workload conditions (M = 32.97, SD = 18.32).

Mental demand. The mental demand subscale refers to the amount of mental activity the participants thought the task required, ranging from low demand (0) to high demand (100). The mental demand subscale was examined using a mixed ANOVA, with time of day as the between-subjects factor (two levels: morning, afternoon) and phase (three levels: beginning, middle, end) and workload (two levels: high, low) as within-subjects factors. Mauchly's test indicated that the assumption of sphericity had been violated for the main effect of phase, $\chi^2(2) = 6.8$, p < .05, therefore Greenhouse-Geisser corrected tests are reported (ε = .88). The main effect of workload was significant, F (1, 48) = 45.90, p < .001. Participants rated mental workload demands higher during the high workload (M = 68.21, SD = 15.13) conditions than during the low workload (M = 53.70, SD = 21.92) conditions. The interaction between workload and TOD was significant, $F(1, 48) = 4.69$, $p < .05$. The interaction effect was examined using the LSD test and it was found that mental demand ratings were significantly greater in morning group than the afternoon group during the low workload conditions ($p < .05$, see Table 6 for means).

Table 6.

Mental Demand by Time of Day and Workload.

	Low Workload		High Workload	
	Mean	Std. Dev.	Mean	Std. Dev.
Morning	58.61	17.53	68.30	14.70
Afternoon	49.46	24.60	68.12	15.77

Physical demand. The physical demand subscale addressed the amount of physical activity participants thought the task required, ranging from low (0) to high (100) physical demand. Physical demand ratings were examined using a mixed ANOVA, with time of day as the between-subjects factor (two levels: morning, afternoon) and phase (three levels: beginning, middle, end) and workload (two levels: high, low) as within-subjects factors. The main effect of workload was significant, $F(1, 48) = 38.28$, $p < .001$. Participants rated physical workload higher during the high workload conditions (M = 50.35, SD = 22.72) than during the low workload conditions (M = 39.83, SD = 21.94). The interaction of workload and TOD approached, but did not reach significance, $F(1, 48) = 3.38$, $p = .07$.

Temporal demand. Temporal demand assesses the amount of time pressure the participant experienced, ranging from low or a slow pace (0) to high or a rapid pace (100). The temporal demand subscale was examined using a mixed ANOVA, with time of day as the between-subjects factor (two levels: morning, afternoon) and phase (three levels: beginning, middle, end) and workload (two levels: high, low) as within-subjects factors. For the main effect of phase, Mauchly's test indicated that the assumption of sphericity had been violated, $\chi2(2)$ = 10.19, $p < .05$, therefore Greenhouse-Geisser corrected tests are reported ($\varepsilon = .84$). The interaction of phase and TOD approached, but did not reach significance, $F(1.67, 80.35) = 2.80$, $p = .08$. The main effect of workload was significant, $F(1, 48) = 41.66$, $p < .001$, with participants' reporting a higher temporal workload during the high workload conditions (M = 62.48, SD = 13.76) than the low workload conditions (M = 55.05, SD = 10.07). The interaction of workload and TOD was significant, $F(1, 48) = 4.98$, $p < .05$. The interaction effect was explored using the LSD test, and it was found that ratings of temporal demand were significantly higher for morning group than the afternoon group during the low workload conditions ($p < .05$, see Table 7 for means).

Table 7.

Temporal Demand by Time of Day and Workload.

	Low Workload		High Workload	
	Mean	Std. Dev.	Mean	Std. Dev.
Morning	59.67	17.63	64.64	15.87
Afternoon	51.11	13.73	60.64	11.67

Performance. The performance subscale required participants to rate how well they thought they performed, ranging from poor (0) to good (100). The performance score was examined using a mixed ANOVA, with time of day as the between-subjects factor (two levels: morning, afternoon) and phase (three levels: beginning, middle, end) and workload (two levels: high, low) as within-subjects factors. The main effect of workload was significant, $F(1, 48)$ =

45.91, p < .001, with participants' rating own performance higher during the low workload conditions (M = 41.85, SD = 14.79) than the high workload conditions (M = 31.79, SD = 18.29). The three-way interaction of phase, workload, and TOD, approached, but did not reach significance, F (2, 96) = 2.65, p = .08.

Workload rating scale mean. The workload rating scale (WRS) mean consisted of the mean of all subscale scores combined. The WRS mean was examined using a mixed ANOVA, with time of day as the between-subjects factor (two levels: morning, afternoon) and phase (three levels: beginning, middle, end) and workload (two levels: high, low) as within-subjects factors. The main effect of workload was significant, F (1, 48) = 69.64, p < .001. The mean workload rating was higher for high workload conditions (M = 55.43, SD = 11.54) than low workload conditions (M = 43.32, SD = 14.64). The interaction of workload and TOD was also significant, F (1, 48) = 4.49, p < .05. The interaction was explored using the LSD test, where it was found overall workload ratings were significantly higher in the morning group compared to the afternoon group on the low workload conditions (p < .05, see Table 8 for means).

Table 8.

Workload Rating Scale Mean by Time of Day and Workload.

	Low Workload		High Workload	
	Mean	Std. Dev.	Mean	Std. Dev.
Morning	46.75	13.58	55.61	11.62
Afternoon	40.40	15.13	55.28	11.69

Moderation analyses results. The interaction with TOD was explored using moderation analyses to examine whether MEQ, PSQI, or ESS scores moderate the relationship between

TOD and workload rating scores where significant interaction effects occurred. The EEG index of low engagement was also examined using moderation analyses to explore the interaction with TOD, but did not reach or approach significance for MEQ, PSQI, or ESS scores moderating the interaction effect with TOD.

Moderation analyses did not show a significant effect of MEQ, PSQI, and ESS on moderating the interaction effect of mental demands ratings and TOD. Results of the moderation analyses for temporal demands are summarized in Table 9 below. MEQ scores approached, but did not reach, significance for moderating the effect of TOD on temporal workload ratings. The PSQI and ESS did not have a significant moderating effect on the TOD and temporal workload ratings.

Table 9.

Moderation Results for Temporal Workload Demands.

	b	*95% CI*	*SE B*	t	p
Constant	57.75	52.87, 62.63	2.43	23.81	<.001
MEQ Score	.2860	-.3871, .9591	.3344	.8553	.3968
TOD	-9.73	-19.88, .4187	5.04	-1.93	.0598
MEQ X TOD	-1.31	-2.74, .1260	.7125	-1.84	.0728

Note. $R^2 = .20$

Analyses for the moderating effect of PSQI and ESS found no moderating effects of these variables on the interaction of TOD and overall workload ratings. Results for the moderation analyses of overall workload ratings with MEQ scores as a moderating variable are summarized in Table 10 below. These results indicate that MEQ scores had a significant effect on moderating the relationship of TOD and overall workload rating scores. The moderation effect was examined by looking at the simple slopes. These found that when MEQ scores are high, there is a significant between TOD and WRS scores, *b* = -12.85, 95% CI [-23.48, -2.22], *t* = -2.43, $p < .05$, such that higher MEQ scores correspond with higher overall workload rating scores in the mornings. The simple slopes are displayed in Figure 1 below.

Table 10.

Moderation Results for Overall Workload Ratings.

	b	*95% CI*	*SE B*	t	p
Constant	50.44	46.90, 53.99	1.76	28.65	<.001
MEQ Score	.2448	-.1816, .6712	.2118	1.16	.2538
TOD	-4.70	-11.80, 2.39	3.52	-1.34	.1884
MEQ X TOD	-1.02	-1.89, -.1592	.4296	-2.38	<.05

Note. $R^2 = .13$

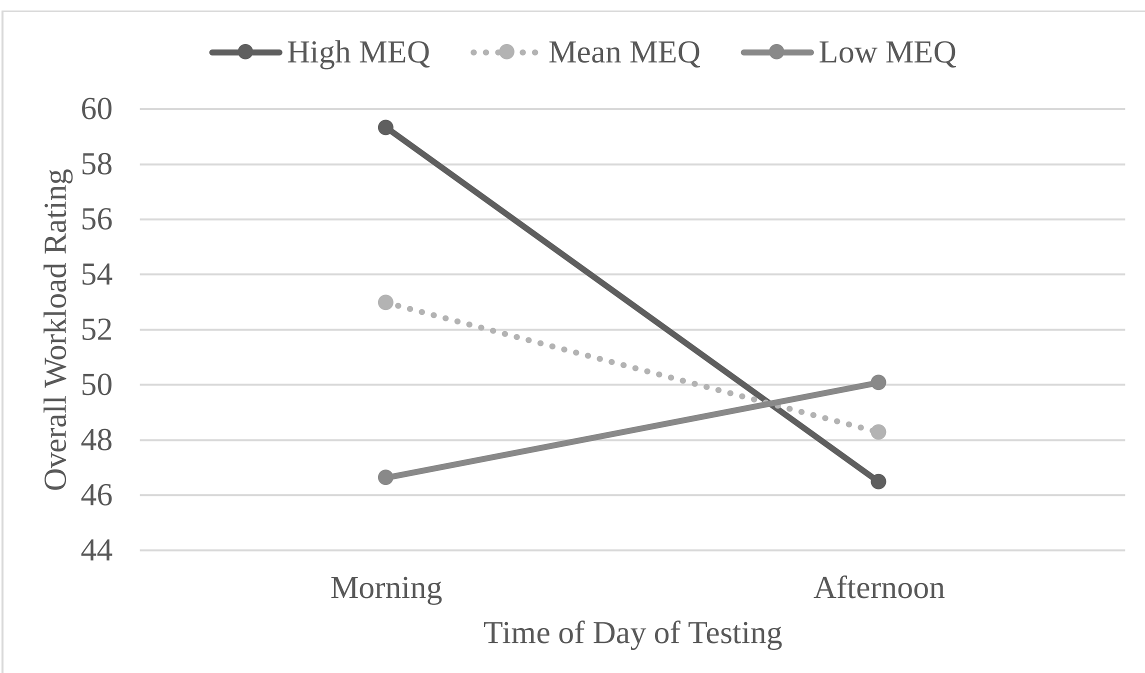


Figure 1. Simple slopes for MEQ scores on overall workload rating.

CHAPTER 4

DISCUSSION

The present study examined the influence of time of day on performance during a multitasking test, with participants' tested either in the morning (0800 or 0900) or afternoon (1500 or 1600). Participants' circadian typology was also assessed using the Morningness-Eveningness Questionnaire, with scores used as a moderating variable when time of day differences emerged on subjective workload rating scores. Overall, the study failed to find differences in performance due to time of day; however, differences in performance in regards to phase and workload did occur. Furthermore, the EEG cognitive state classifications were found to nearly mirror performance differences in terms of phase and workload, supporting the use of EEG as an objective real-time measure of workload. Findings, implications, and limitations of the present study are discussed below.

Performance Findings

Performance on the MATB-II was shown to vary throughout the testing period, partially supporting the first hypothesis regarding performance changes throughout testing. Specifically, both the tracking and systems monitoring task performance were influenced by workload conditions, supporting the hypothesis that performance would worsen during the high workload conditions. Decreased performance during the high workload conditions was expected, as participants' cognitive resources are depleted due to the increased task demands (Wickens, 2002, 2008). The lack of a significant difference due to workload manipulation alone on the resource monitoring and communication task has been found in other studies using the MATB-II, as well (Prinzel, Freeman, & Prinzel, 2005). While these two tasks did not yield significant differences due to workload manipulation, the resource monitoring task found a significant interaction with workload and phase, such performance varied throughout the phases in the high and low workload conditions. Examination of the mean tank deviations for each phase also showed a change in performance such that deviations lessened during phase two and increased during phase three. Furthermore, the changes were greater during the high workload conditions than low workload conditions across the phases. This change is likely related to time on task and a depletion of cognitive resources available to attend to the task (Wickens, 2002, 2008). Performance during the communications task was significantly affected by phase, with the workload and phase interaction approaching significance. Reaction times to the communications tasks significantly improved throughout each of the three phases.

Performance on nearly all of the tasks improved with time on task, rather than show the usual performance decrement that is known to occur with increased time on task (Lim et al., 2010). The resource monitoring task was the only task to show any significant decline in performance during the third phase of the testing period. Examination of overall task prioritizing found that prioritization of tasks were relatively equal, where approximately one-third of participants rated the systems monitoring tasks as a top priority and one-quarter of participants rated the resource monitoring task as a top priority. Thus, it is unlikely that performance differences amongst tasks were related to how participants were prioritizing attention to tasks. However, task prioritization was collected at the end of the entire testing session, so it is possible that participants prioritized tasks to attend to differently depending on workload levels. Previous research has shown participants will often develop a strategy to manage performance on simultaneous tasks by prioritizing responses (Kurzban, Duckworth, Kable, & Myers, 2014). The variable workload levels in the present task may have allowed participants the opportunity to regain cognitive resources in order to develop and utilize a strategy to maintain performance throughout the remainder of the testing period.

The hypothesis regarding performance decreasing during the latter half of the testing period was not supported. This result is similar to that of Fairclough and Venables (2006) who demonstrated in a previous study using a “demanding” version of the MATB no significant differences in performance across time, where participants completed four consecutive 20 minute blocks on the MATB. However, their study did not manipulate workload conditions, which as noted in the interaction between workload and phase, likely had an effect on the performance differences in the current study. The Malleable Attentional Resource theory (Young & Stanton, 2002) states that cognitive resources are reduced during periods of low workload as fewer resources are required to maintain performance. Thus, it is plausible that the temporary reduction in resources required to maintain performance gave participants enough of a “break” to then recruit the additional resources required to maintain performance in the high workload conditions.

TOD and Performance

The present study found no significant differences in performance or EEG classifications based on time of day, thus not supporting the third hypothesis. This supports the recent findings of Clegg and colleagues (2015), who also examined the effects of time of day in multitasking performance on the MATB-II and an additional task, and found no significant differences in performance based on time of day. The lack of performance differences occurring due to time of day may be attributed to several factors. A recent meta-analysis by Wickens and colleagues (2015) examined the effects of total sleep deprivation, partial sleep deprivation, and circadian cycle on complex task (e.g., multitasking) performance. It was found that while performance decrements occurred within each of these types of studies, they were not as severe when compared to their simple task counterparts (e.g., psychomotor vigilance task). Previous studies manipulating time of day on performance found significant differences between morning and afternoon testing times. However, the tasks used in such studies were often simple tasks that did not require multitasking. For example, Hourihan and Benjamin (2014) found time of day differences in memory recall using morning testing times of 8 or 9 am and afternoon testing times of 3 or 4pm, just as the present study used. Additionally, Knight (2013) found a significant different in alerting based on time of day, when also using similar testing times. Thus, it is likely the current study did not find significant performance differences between time of day of testing due to task complexity and not the testing times that were chosen.

The ability to maintain performance in complex tasks compared to simple tasks has been attributed to the notion that complex tasks requiring a participant to engage in multitasking are often more engaging than simple tasks. For example, a study by Wilson, Caldwell, and Russell (2007) had participants complete a sustained attention task while sleep deprived. They found that participants' performance during the last testing session was not as degraded as expected, which they attributed to the participants being engaged in a high workload task during this period, after previously being engaged in a low workload task. The high workload task was thought to be more engaging and complex, causing the participants to pay closer attention despite being in a fatigued state.

Sleep Quality and Daytime Sleepiness

There were no significant differences in reported sleep quality or daytime sleepiness between the times of day of participation, thus the second hypothesis was not supported. However, given that participants were not required to adjust their schedules, nor pre-screened based on circadian typology, a difference on these measures in the current population would be unlikely, particularly given that the majority participants fell within the intermediate type on the MEQ. Individuals who fall closer to the E-type end of MEQ scores tend to have higher rates of poor sleep quality (Buysse et al., 1989; Roeser et al., 2012; Wittman et al., 2006), thus a population consisting mainly of intermediate types would likely no result in any significant differences on this measure. Furthermore, when these measures were examined to determine whether any correlations with sleep quality and daytime sleepiness existed, no significant correlations were found, and as a result were not used as covariates. The moderation analyses also did not find any significant impact with these measures entered as potential moderating variables where interactions with TOD occurred.

Electroencephalogram

The hypothesis of differences in EEG state classifications based on workload (high versus low) was supported in the current study, supporting prior research that EEG state classifications can be used as an objective measure for quantifying workload (Young, Brookhuis, Wickens, & Hancock, 2014). Specifically, participants' classification probabilities were elevated for cognitive state and workload during the high workload conditions as compared to the low workload conditions. Additionally, the probability of distraction was significantly lower during the low workload conditions compared to high workload conditions. This finding supports similar findings in previous studies where EEG has been used in detecting differences in workload (e.g., Smith et al., 2001; Gevins et al., 1998; Hankins & Wilson, 1998).. The differences in EEG classifications based on workload varied in accordance with the performance measures, such that both the EEG classification and objective performance were significantly worse during high workload conditions for the tracking and systems monitoring tasks; thus lending credence and support to the idea that EEG state classifications can be used as an objective measure of workload.

Additionally, phase differences were noted in the EEG classifications that matched those with the performance changes. Performance was shown to significantly improve across phases, which is supported in the EEG cognitive state classifications and the probability of high engagement changing across phases, such that they each decreased across each phase, whereas probabilities of distraction and low engagement increased across phases. The changes in these EEG indices are indicative of participants needing fewer cognitive resources while still managing to maintain and improve performance. Additionally, these changes in EEG indices that mirrored those of performance changes support the findings of Kamzanova, Kustubayeva, and Matthews (2014), where they were able to demonstrate changes in EEG indices of engagement in response to a vigilance task.

The hypothesis regarding differences in sleep quality and daytime sleepiness in EEG classifications was not explored since no differences in time of day were noted for these two subjective measures. However, for the probability of low engagement and high engagement there was a significant interaction effect between phase and time of day. The moderation analyses revealed no moderating effect of MEQ scores, daytime sleepiness, or sleep quality on low engagement or high engagement probabilities, suggesting the interaction is simply related to TOD. Additionally, the hypothesis regarding differences in physiological measures of blood pressure and body temperature resulted in no TOD differences only in temperature, which was likely due to circadian phase (Rogers et al., 2003).

Subjective Workload

While there were no hypotheses regarding differences in subjective measures of workload, a significant interaction was found for two workload rating subscale scores and time of day. Further analyses found that circadian typology may have had a moderating role on subscale scores for physical and temporal workload, although significance was not reached. Differences in overall WRS scores based on time of day of testing were found to be moderated by circadian typology. Examination of the simple slopes indicated a significant effect for higher MEQ scores, which indicate a tendency towards morningness, during morning sessions where they rated overall workload as higher. This finding supports previous research which has identified circadian typology as a potential modifying factor when examining an individual's stress or perceived workload during a given task (Oginska et al., 2010).

The finding of MEQ scores as a moderating factor on subjective workload is likely due to a higher production of cortisol in the morning in individuals who tend towards morningness-type. Previous studies have examined cortisol level differences in morning and evening type, and have consistently found that morning types have higher cortisol levels in the mornings compared to evening types (Bailey & Heitkemper, 2001; Kudielka et al., 2006, 2007). For example, Bailey and Heitkemper (2001) compared blood cortisol levels and temperature between M-type and E-type individuals, and found that M-types reached their acrophase, or peak time of circadian rhythm, one hour earlier than E-types. Additionally, it has been shown that M-type individuals have higher daytime levels of cortisol compared to E-types (Kudielka et al., 2007). Cortisol is known to play a role in activating the hypothalamic-pituitary-adrenal gland axis, which plays an integral role in the body's reaction to stress (Lovallo & Thomas, 2000).

While the current study did not measure cortisol levels, the known association between M-types and cortisol production in the morning provides a basis for understanding why those tending toward M-type would rate workload significantly higher than those who fall toward the intermediate and E-type end of the continuum. Moreover, while previous research has identified the perception of workload to be influenced by factors such as event rates, type of task, need to multitask, and individual's skill or experience level with the task (Borghini et al., 2014; Parasuraman, 1979), little attention has been given to individual differences such as circadian typology in respect to time of day. The findings of the present study highlight the need to consider time of day and circadian typology as factors when examining subjective workload.

CHAPTER 5

CONCLUSION

The goal of the present study was to determine whether performance differences occurred in response to multitasking test with varying levels of workload at different times of the day. While the present study did not find performance differences in regards to time of day, several noteworthy findings emerged, including the finding that performance on complex tasks may be less susceptible to time of day effects.

Performance was found to vary in response to the workload manipulations on two of the four tasks, supporting the multiple resource theory of attention (Wickens, 2002, 2008) where increases in task demands result in competition for resources, which result in a decrease in performance when resources are not available to address all task demands. However, the usual decrease in performance resulting from time on task was not seen in the current study. Instead, performance was noted to improve with time on task. While this finding may be attributed to learning effects, it is also possible that the study design influenced participants' ability to learn and develop strategies to utilize throughout the duration of the test. Previous research has demonstrated that when faced with the need to multitask, individuals will often develop a strategy for meeting task demands (Kurzban, Duckworth, Kable, & Myers, 2014). The present study used equal-length periods of high and low workload demands, whereas decreases in performance during periods of low workload are typically noted longer periods of time, such as a cruise phase during flight (Cabon et al., 1993). However, it is also possible that participants were able to utilize the periods of low workload as an opportunity to further develop strategies in managing workload.

While time of day did not result in any differences in performance, it does beg the question of whether current fatigue models, which are based on simple tasks that are known to be affected by time of day (Hursh et al., 2004; Lamond et al., 2005) are accurate indicators of an individual's performance ability. The current study did not manipulate fatigue levels, but instead only examined task performance in relation to circadian variability based on time of day. Therefore, while the current study is unable to address the fatigue aspect of such tasks being used to create prediction models, the finding that performance on a complex task is not affected by time of day as are simple tasks, warrants further research on whether or not performance on simple tasks translates into operationally-relevant performance.

Perhaps most noteworthy is the interactions of TOD with workload ratings on the workload rating scale. Subjective measures of workload have historically been used to determine quantify an individual's experience of workload when completing a given task (Wickens & Tsang, 2015). The differences in WRS ratings based on TOD and MEQ scores points to the need for further research on the characteristics that participants bring with them to laboratory, or state-based characteristics (Hourihan & Benjamin, 2014). That is, simply relying on subjective measures of workload without consideration for the state-based characteristics the participant brings, may result in inaccurate interpretations of the findings. The present study found individuals tending toward M-type on the MEQ rated workload higher, despite no significant differences in performance measures. Thus, studies relying solely on subjective and performance measures to determine workload classification may be divergent. The current study showed a basically mirrored relationship between EEG classifications and performance measures, suggesting that the use of psychophysiological measures in identifying workload differences and changes in the participants' state might be more robust.

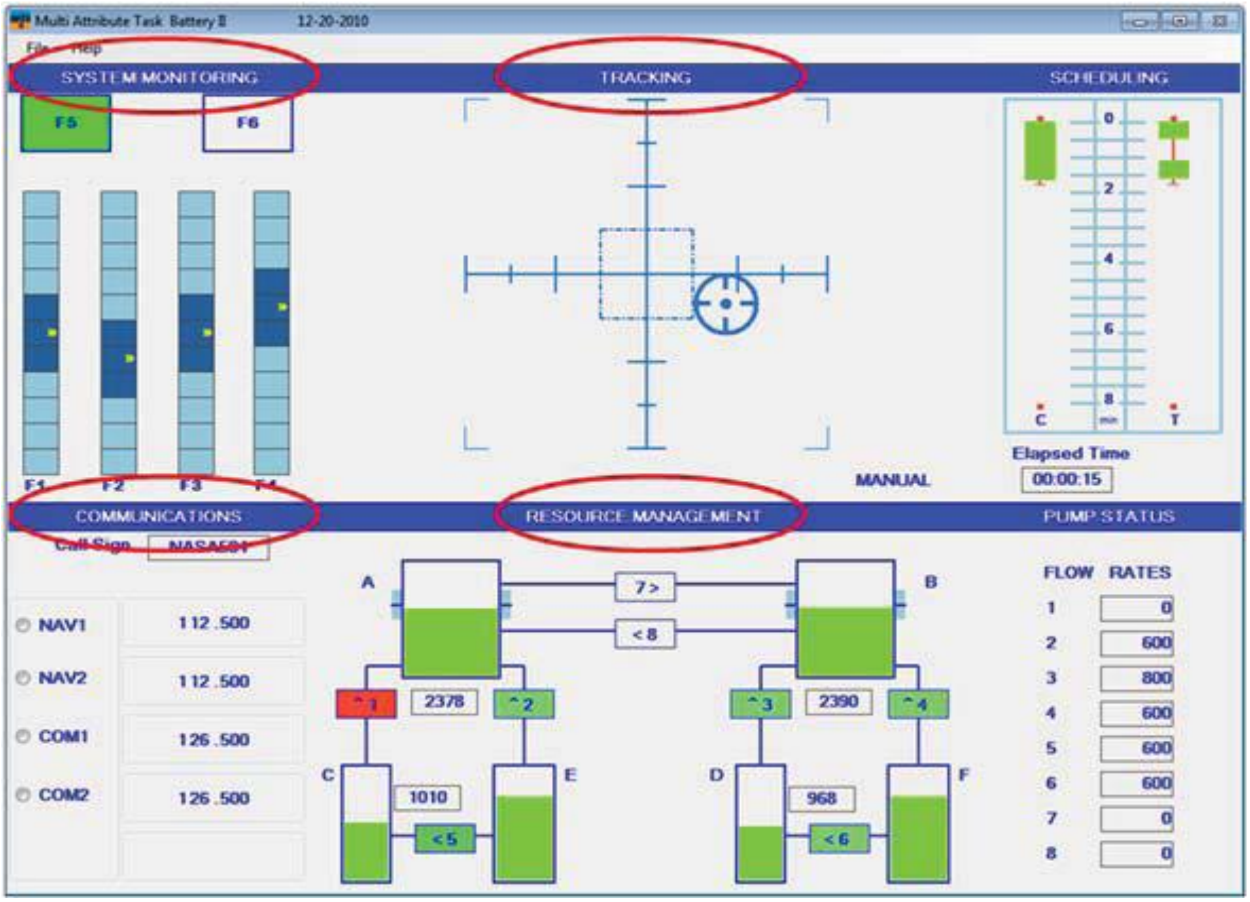
Limitations

The present study was limited in scope by a number of factors. One limitation was the variable levels of workload used in the study. Given that current research is still unclear in regards to the effects of exposure to different levels of workload in reference to both time of day and circadian typology, future research could benefit by examining these variations individually. That is, a comparison of tasks during each a high and low workload condition, rather than a mixture of the two throughout the testing duration, could provide additional insight regarding performance changes. Along similar lines, research comparing simple task performance, for example, psychomotor vigilance tasks, to more operationally-relevant tasks (e.g., a simulated flight with a high level of workload), would also provide insight on whether these simple task measures are viable candidates for basing fatigue models off of.

Another limiting factor of the study was that participants were not pre-selected based on circadian typology. By not prescreening for typology, the current study was unable to examine extreme typologies in terms of performance. Future work should examine performance differences using individuals who fall strictly in the morning or evening-type categories to gain a clearer picture of the influence these have on performance. Similarly, it would be beneficial to also assess cortisol levels and other physiological measures to gain a broader understanding of differences between groups.

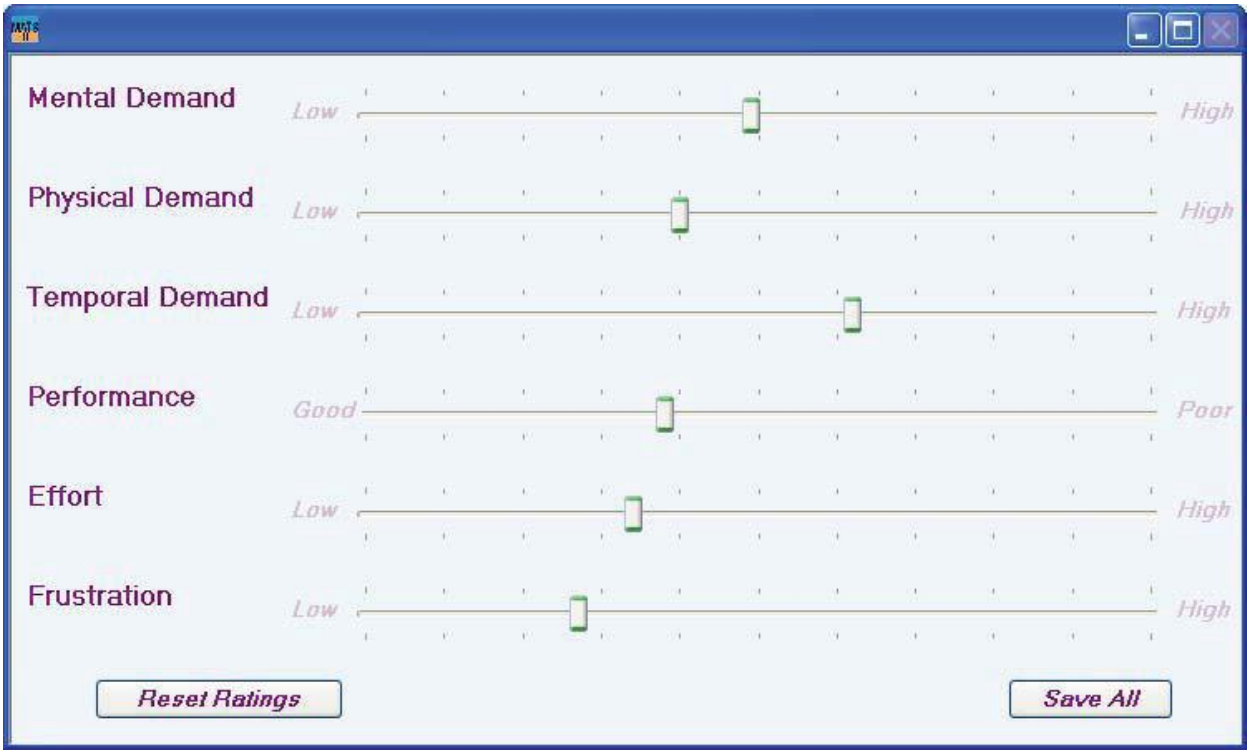
APPENDICES

Appendix A
Multi-Attribute Task Battery-II (MATB-II)



Appendix B

Workload Rating Scale



Appendix C

Pittsburgh Sleep Quality Index

Instructions:

The following questions relate to your usual sleep habits during the past month only. Your answers should indicate the most accurate reply for the majority of days and nights in the past month. Please answer all questions.

1. During the past month, what time have you usually gone to bed at night?

BED TIME _____________

2. During the past month, how long (in minutes) has it usually taken you to fall asleep each night?

NUMBER OF MINUTES___________

3. During the past month, what time have you usually gotten up in the morning?

GETTING UP TIME _____________

4. During the past month, how many hours of actual sleep did you get at night? (This may be different than the number of hours you spent in bed).

HOURS OF SLEEP PER NIGHT ____________

For each of the remaining questions, check the one best response. Please answer all questions.

5. During the past month, how often have you had trouble sleeping because you . . .

a) Cannot get to sleep within 30 minutes

Not during the past month_______	Less than once a week______	Once or twice a week_________	Three or more times a week_________

b) Wake up in the middle of the night or early morning

Not during the past month_______	Less than once a week______	Once or twice a week_________	Three or more times a week_________

c) Have to get up to use the bathroom

Not during the past month_______	Less than once a week______	Once or twice a week_________	Three or more times a week_________

d) Cannot breathe comfortably

Not during the past month_______	Less than once a week______	Once or twice a week_________	Three or more times a week_________

e) Cough or snore loudly

Not during the past month_______	Less than once a week______	Once or twice a week_________	Three or more times a week_________

f) Feel too cold

Not during the past month_______	Less than once a week______	Once or twice a week_________	Three or more times a week_________

g) Feel too hot

Not during the past month_______	Less than once a week______	Once or twice a week_________	Three or more times a week_________

h) Had bad dreams

Not during the past month_______	Less than once a week______	Once or twice a week_________	Three or more times a week_________

i) Have pain

Not during the past month_______	Less than once a week______	Once or twice a week_________	Three or more times a week_________

j) Other reason(s), please describe ____
____

How often during the past month have you had trouble sleeping because of this?

Not during the past month_______	Less than once a week______	Once or twice a week_________	Three or more times a week_________

6) During the past month, how would you rate your sleep quality overall?

Very good _________________
Fairly good____________________

Fairly bad______________________
Very bad ______________________

7) During the past month, how often have you taken medicine to help you sleep (prescribed or "over the counter")?

Not during the past month_______	Less than once a week______	Once or twice a week_________	Three or more times a week_________

8) During the past month, how often have you had trouble staying awake while driving, eating meals, or engaging in social activity?

Not during the past month_______	Less than once a week______	Once or twice a week_________	Three or more times a week_________

9) During the past month, how much of a problem has it been for you to keep up enough enthusiasm to get things done?

No problem at all ___________
Only a very slight problem ___________
Somewhat of a problem ____________
A very big problem _____________

10) Do you have a bed partner or roommate?

No bed partner or roommate ___________
Partner/roommate in other room ___________
Partner in same room, but not same bed ___________
Partner in same bed ___________

If you have a roommate or bed partner, how often has he/she said in the past month you have had …

a) Loud snoring

Not during the past month_______	Less than once a week______	Once or twice a week_________	Three or more times a week_________

b) Long pauses between breaths while asleep

Not during the past month_______	Less than once a week______	Once or twice a week_________	Three or more times a week_________

c) Legs twitching or jerking while you sleep

Not during the past month_______	Less than once a week______	Once or twice a week_________	Three or more times a week_________

d) Episodes of disorientation or confusion during sleep

Not during the past month_______	Less than once a week______	Once or twice a week_________	Three or more times a week_________

e) Other restlessness while you sleep; please describe ____________________________________
____

Not during the past month_______	Less than once a week______	Once or twice a week_________	Three or more times a week_________

Appendix D

Epworth Sleepiness Scale

How likely are you to doze off or fall asleep in the following situations, in contrast to feeling just tired? This refers to your usual way of life in recent times. Even if you haven't done some these things recently try to work out how they would have affected you.

Use the following scale to choose the **most appropriate number** for each situation:

0 = would **never** doze
1 = **slight chance** of dozing
2 = **moderate chance** of dozing
3 = **high chance** of dozing

It is important that you answer each question as best you can.

Situation	**Chance of Dozing (0-3)**
Sitting and reading	_______________
Watching TV	________________
Sitting, inactive in a public place (e.g., a theater or a meeting)	________________
As a passenger in car for an hour without a break	________________
Lying down to rest in the afternoon when circumstances permit	________________
Sitting and talking to someone	________________
Sitting quietly after a lunch without alcohol	________________
In a car, while stopped for a few minutes in the traffic	________________

Appendix E

Horne & Osteberg Questionnaire

Instructions:
Please read each question carefully before answering. Answer ALL questions. Answer questions in numerical order. Each question should be answered independently of others. Do NOT go back and check your answers. All questions have a selection of answers. For each question place a cross (X) alongside ONE answer only. Some questions have a scale instead of a selection of answers. Place an X at the appropriate point along the scale. Please answer each question as honestly as possible. Both your answers and the results will be kept in strict confidence.

1. Considering only your own "feeling best" rhythm, at what time would you get up if you were entirely free to plan your day?

AM **5** I I I **6** I I I **7** I I I **8** I I I **9** I I I **10** I I I **11** I I I **12** PM

2. Considering only your own "feeling best" rhythm, at what time would you go to bed if you were entirely free to plan your evening?

PM **8** I I I **9** I I I **10** I I I **11** I I I **12**AM I I I **1** I I I **2** I I I **3** AM

3. If there is a specific time at which you have to get up in the morning, to what extent are you dependent on being woken up by an alarm clock?

Not at all dependent ____
Slightly dependent ____
Fairly dependent ____
Very dependent ____

4. Assuming adequate environmental conditions, how easy do you find getting up in the morning?

Not at all easy ____
Not very easy ____
Fairly easy ____
Very easy ____

5. How alert do you feel during the first half-hour after having woken up in the morning?

Not at all alert ____
Slightly alert ____

Fairly alert ____
Very alert ____

6. How is your appetite during the first half-hour after having woken up in the morning?

Very poor ____
Fairly poor ____
Fairly good ____
Very good ____

7. During the first half-hour after waking up in the morning, how tired do you feel?

Very tired ____
Fairly tired ____
Fairly refreshed ____
Very refreshed ____

8. When you have no commitments the next day, at what time do you go to bed compared to your usual bedtime?

Seldom or never late ____
Less than 1 hour later ____
1-2 hours later ____
More than 2 hours later ____

9. You have decided to engage in some physical exercise. A friend suggests that you do this 1 hour twice a week and the best time for him is between 7:00 and 8:00 AM. Bearing in mind nothing else but your own "feeling best" rhythm, how do you think you would perform?

Would be on good form ____
Would be on reasonable form____
Would find it difficult ____
Would find it very difficult ____

10. At what time in the evening do you feel tired and as a result in need of sleep?

PM **8** I I I **9** I I I **10** I I I **11** I I I **12**AM I I I **1** I I I **2** I I I **3** AM

11. You wish to be at your peak performance for a test which you know is going to be mentally exhausting and lasting for 2 hours. You are entirely free to plan your day and considering only you own "feeling best" rhythm, which ONE of the four testing times would you choose?

8:00-10:00 AM ____
11:00 AM-1:00 PM ____
3:00-5:00 PM ____
7:00-9:00 PM ____

12. If you went to bed at 11:00 PM, at what level of tiredness would you be?

Not at all tired ____
A little tired ____
Fairly tired ____
Very tired ____

13. For some reason you have gone to bed several hours later than usual, but there is no need to get up at any particular time the next morning. Which ONE of the following events are you most likely to experience?

Will wake up at usual time and will NOT fall asleep ____
Will wake up at usual time and will doze thereafter ____
Will wake up at usual time and will fall asleep again ____
Will NOT wake up until later than usual ____

14. One night you have to remain awake between 4:00-6:00 AM in order to carry out a night watch. You have no commitments the next day. Which ONE of the following alternatives will suit you best?

Would NOT go to bed until the watch was over ____
Would take a nap before and sleep after ____
Would take a good sleep before and nap after ____
Would take ALL sleep before watch ____

15. You have to do 2 hours of hard, physical work. You are entirely free to plan your day and considering only your own "feeling best" rhythm, which ONE of the following times would you choose?

8:00-10:00 AM ____
11:00AM-1:00 PM ____
3:00-5:00 PM ____
7:00-9:00 PM ____

16. You have decided to engage in some physical exercise. A friend suggests that you do this 1 hour twice a week and the best time for him or her is between 10:00-11:00 PM. Bearing in mind nothing else but only your own "feeling best" rhythm, how do you think you would perform?

Would be on good form ____
Would be on reasonable form ____
Would find it difficult ____
Would find it very difficult ____

17. Suppose that you can choose your own work hours. Assume that you worked a FIVE hour day (including breaks) and that your job was interesting and paid by results. Which FIVE CONSECUTIVE HOURS would you select?

12 1 2 3 4 5 6 7 8 9 10 11 12 1 2 3 4 5 6 7 8 9 10 11 12
Midnight Noon

18. At what time of the day do you think that you reach your "feeling best" peak?

12 1 2 3 4 5 6 7 8 9 10 11 12 1 2 3 4 5 6 7 8 9 10 11 12
Midnight Noon

19. One hears about "morning" and "evening" types of people. Which ONE of these types do you consider yourself to be?

Definitely a "morning" type ____
Rather more a "morning" than an "evening" type ____
Rather more an "evening" than a "morning" type ____
Definitely an "evening" type ____

Appendix F

<u>WAIS-R Vocabulary Subset</u>

1) Bed
2) Ship
3) Penny
4) Winter
5) Breakfast
6) Repair
7) Fabric
8) Assemble
9) Enormous
10) Conceal
11) Sentence
12) Consume
13) Regulate
14) Terminate
15) Commence
16) Domestic
17) Tranquil
18) Ponder
19) Designate
20) Reluctant
21) Obstruct
22) Sanctuary
23) Compassion
24) Evasive
25) Remorse
26) Perimeter
27) Generate
28) Matchless
29) Fortitude
30) Tangible
31) Plagiarize
32) Ominous
33) Encumber
34) Audacious
35) Tirade

Appendix G

Demographic Questionnaire

Date: ______________________________

Before we begin, I would like you to answer the following questions. Thank you.

1. Sex: _____ Male ______ Female
2. Age: _____________________
3. Education History:

   A. High School Graduate Year: ___________________ Degree: _________________

   B. College Graduation Year: _____________________ Degree: _________________

   If currently in college, circle class: FR SO JR SR

4. Using the following scale, please circle the number which corresponds to your current health level in comparison to others your age.

1	2	3	4	5
Excellent	Above Average	Average	Below Average	Poor

5. If you are currently taking any medication(s), would you please describe the type(s) and quantity(s) below.

   ____

   ____

   ____

6. Have you been diagnosed with attention-deficit hyper-activity disorder (ADHD), depression, or received a traumatic brain injury? Yes_______ No______

7. If you answered yes to question 6, what is your diagnosis: ________________________

8. Have you participated in a previous study that utilized the Multi-Attribute Task Battery?

   _________ Yes ________ No

9. Do you have 20/20 uncorrected vision? __________Yes ____________No

10. If you answered NO to number 6, do you wear:

    _______Glasses _______Contacts _______Both _______Neither

11. Do you have any other visual impairments, such as color blindness?

    ________Yes ________No

    If Yes was selected, please state the impairment(s) below:

    ____

    ____

12. Do you regularly drink caffeinated beverages? ________ Yes ________No

13. If you answered yes to question 12, approximately how many caffeinated beverages (such as pop, coffee, or tea) do you drink per day? __________________________

14. Did you drink any caffeinated beverages before coming in to complete this study?

    _______ Yes ________ No

15. If you answered yes to question 14, please answer the following:

    How many caffeinated beverages did you drink? ______________

    What type of beverage(s) did you drink? _____________________

    How long ago did you drink the beverage(s)? __________________

Please respond to the following questions regarding your flight training and experience

1. Total hours of flight time: ____________________

2. Total hours as pilot in command on cross-country flights: _________________________

3. Total hours of instrument flight (actual and simulated)

   A = _______________________

   S = _______________________

4. Total hours of simulated flight: ________________________

REFERENCES

Baber, C. (1991, March). Human factors aspects of automatic speech recognition in control room environments. In *Systems and Applications of Man-Machine Interaction Using Speech I/O, IEE Colloquium on IET* (pp. 10-11).

Bailey, S. L., & Heitkemper, M. M. (2001). Circadian rhythmicity of cortisol and body temperature: Morningness-eveningness effects. *Chronobiology International, 18*(2), 249-261.

Balkin, T. J., Rupp, T., Picchioni, D., & Wesensten, N. J. (2008). Sleep loss and sleepiness: Current issues. *Chest*, *134*, 653-660.

Bennett, C.L., Petros, T.V., Johnson, M., & Ferraro, F.R. (2008). Individual differences in the influence of time of day on executive functions. *American Journal of Psychology, 121*(3), 349-361.

Berka, C., Levendowski, D. J., Lumicao, M. N., Yau, A., Davis, G., …& Craven, P. L. (2007). EEG correlates of task engagement and mental workload in vigilance, learning, and memory tasks. *Aviation, Space, and Environmental Medicine, 78*(5), B231-B244.

Borghini, G., Astolfi, L., Vecchiato, Mattia, D., & Babiloni, F. (2014). Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of workload, fatigue, and drowsiness. *Neuroscience and Biobehavioral Reviews, 44*, 58-75.

Buschkens, J., Graham, D., & Cottrell, D. (2010). Well-being under chronic stress: Is morningness an advantage? *Stress and Health, 26,* 330-340.

Buysse, D.J., Reynolds, C.F. 3rd, Monk, T.H., Berman, S.R., & Kupfer, D.J. (1989). The Pittsburgh Sleep Quality Index: A new instrument for psychiatric practice and research. *Psychiatry Research, 28(2)*, 193-213.

Cabon, P. H., Coblentz, A., Mollard, R., & Fouillot, J. P. (1993). Human vigilance in railway and long-haul flight operation. *Ergonomics, 36*(9), 1019-1033.

Caldwell, J. (2001). Work and sleep hours of U.S. Army aviation personnel working reverse cycle. *Military Medicine, 166*(2), 159-166.

Caldwell, J. A. (2004). Fatigue in aviation. *Travel Medicine and Infectious Disease, 3*, 85-96.

Caldwell, J. A. (2012). Crew schedules, sleep deprivation, and aviation performance. *Current Directions in Psychological Science, 21*(2), 85-89.

Caldwell, J.A., & Ramspott, S. (1998). Effects of task duration on sensitivity to sleep deprivation using the multi-attribute task battery. *Behavior Research Methods, Instruments, & Computers, 30 (4)*, 651-660.

Clegg, B. A., Wickens, C. D., Vieane, A. Z., Gutzwiller, R. S., & Sebok, A. L. (2015, September). Circadian effects on simple components of complex task performance. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 59, No. 1, pp. 627-631). SAGE Publications.

Coca, A. (1994). Circadian rhythm and blood pressure control: Physiological and pathophysiological factors. *Journal of Hypertension, 12*(5), 13-21.

Comstock, J. R., Jr., & Arnegard, R. J. (1992). The Multi-Attribute Task Battery for human operator workload and strategic behavior research (NASA TM-104174). Hampton, Virginia: NASA Langley Research Center.

De Gennaro, L., Ferrara, M., Curcio, G., & Bertini, M. (2001). Visual search performance across 40 h of continuous wakefulness: Measures of speed and accuracy and relation with oculomotor performance. *Physiology and Behavior, 74*, 197-204.

De Gennaro, L., Marzano, C., Veneiro, D., Moroni, F., Fratello, F., Giuseppe, C. … & Rossini, P.M. (2007). Neurophysiological correlates of sleepiness: A combined TMS and EEG study. *NeuroImage, 33*, 1277-1287.

Di Nocera, F., Camilli, M., & Terenzi, M. (2007). A random glance at the flight deck: Pilots' scanning strategies and the real-time assessment of mental workload. *Journal of Cognitive Engineering and Decision Making, 1*(3), 271-285.

Dinges, D. F. (2004). Critical research issues in development of biomathematical models of fatigue and performance. *Aviation, Space, and Environmental Medicine, 75*(1), A181-A191.

Duguay, D., & Cermakian, N. (2009). The crosstalk between physiology and circadian clock proteins. *Chronobiology International, 26*(8), 1479-1513.

Durmer, J.S., & Dinges, D.F. (2005). Neurocognitive consequences of sleep deprivation. *Seminars in Neurology, 25*(1), 117-129.

Ebbinghaus, H. (1885). 1964. *Memory: A contribution to experimental psychology*.

Fairclough, S.H., Venables, L., & Tattersall, A. (2005). The influence of task demand and learning on the psychophysiological response. *International Journal of Psychophysiology, 56*, 171-184.

Gevins, A., Smith, M. E., Leong, H., McEvoy, L.,Whitfield, S., Du, R., & Rush, G. (1998). Monitoring working memory load during computer-based tasks with EEG pattern recognition methods. *Human Factors, 40*(1), 79-91.

Griefahn, B. (2002). The validity of the temporal parameters of the daily rhythm of melatonin levels as indicator of morningness. *Chronobiology International, 19*(3), 561-577.

Gunn, D.V., Warm, J.S., Nelson, W.T., & Bolia, R.S., et al. (2005). Target acquisition with UAVs: Vigilance displays and advanced cuing interfaces. *Human Factors, 47*, (3), 488-497.

Hankins, T. C., & Wilson, G. F. (1998). A comparison of heart rate, eye activity, EEG, and subjective measures of pilot mental workload during flight. *Aviation, Space, and Environmental Medicine, 69*(4), 360-367.

Hart, S.G. & Staveland, L.E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P.A. Hancock & N. Meshkati (Eds.), *Human mental workload* (pp. 139-183). Amsterdam: North-Holland.

Horne, J.A., & Ostberg, O. (1976). Self-assessment questionnaire to determine morning-eveningness in human circadian rhythms. *International Journal of Chronobiology*, *4,* 97-110.

Hourihan, K. L., & Benjamin, A. S. (2014). State-based metacognition: How time of day affects the accuracy of metamemory. *Memory*, *22*(5), 553-558.

Huikuri, H.V., Kessler, K.M., Terracall, E., Castellanos, A., Linnaluto, M.K., & Myerburg, R.J. (1990). Reproducibility and circadian rhythm of heart rate variability in healthy subjects. *The American Journal of Cardiology*, *65*(5), 391-393.

Hursh, S. R., Balkin, T. J., Miller, J. C., & Eddy, D. R. (2004). *The fatigue avoidance scheduling tool: Modeling to minimize the effects of fatigue on cognitive performance* (No. 2004-01-2151). SAE Technical Paper.

Jasper, I., Roenneberg, T., Häußler, A., Zierdt, A., Marquardt, C., & Hermsdörfer, J. (2010). Circadian rhythm in force tracking and in dual task costs. *Chronobiology International, 27* (3), 653-673. doi: 10.3109/07420521003663793.

Johns, M.W. (1991). A new method for measuring daytime sleepiness: The Epworth Sleepiness Scale. *Sleep*, *14*(6), 540-545.

Johnson, R. R., Popovic, D. P., Olmstead, R. E., Stikic, M., Levendowski, D. J., & Berka, C. (2011). Drowsiness/alertness algorithm development and validation using synchronized EEG and cognitive performance to individualize a generalized model. *Biological psychology*, *87*(2), 241-250.

Johnson, A., & Proctor, R. W. (2004). *Attention: Theory and Practice*. Sage Publications.

Kamzanova, A. T., Kustubayeva, A. M., & Matthews, G. (2014). Use of EEG workload indices for diagnostic monitoring of vigilance decrement. *Human Factors, 56*(6), 1136-1149.

Klyde, D. H., Lampton, A. K, Schulze, P. C., Alvarez, D. J., Johnson, R., & Rowe, L. (2013). The real-flight approach to assess flight simulator force cueing fidelity. *In AIAA Atmospheric Flight Mechanics Conference Proceedings*, Boston, M.A.

Knight, M. (2013). Look out-It's your off-peak time of day! Time of day matters more for alerting than for orienting of executive attention. *Experimental Aging Research, 39*, 305-321.

Kudielka, B. M., Federenko, I. S., Dirk, Hellhammer, D. H., & Wüst, S. (2006). Morningness and eveningness: The free cortisol rise after awakening in "early birds" and "night owls". *Biological Psychology, 72*, 141-146.

Kudielka, B. M., Bellingrath, S., & Hellhammer, D. H. (2007). Further support for higher salivary cortisol levels in "morning" compared to "evening" persons. *Journal of Psychosomatic Research, 62,* 595-596.

Kurzban, R., Duckworth, A., Kable, J. W., & Myers, J. (2013). An opportunity cost model of subjective effort and task performance. *Behavioral and Brain Sciences, 36*(6), 661-679.

Lal, S. K. I., & Craig, A. (2000). Psychophysiological effects associated with drowsiness: Driver fatigue and electroencephalography. *International Journal of Physiology, 35*(1), 39-47.

Lal, S. K. I., & Craig, A. (2002). Driver fatigue electroencephalography and psychological assessment. *Psychophysiology, 39*, 1-9.

Lamond, N., Dawson, D., & Roach, G. D. (2005). Fatigue assessment in the field: Validation of a hand-held electronic psychomotor vigilance task. *Aviation, Space, and Environmental Medicine*, *76*(5), 486-489.

Lovallo, W. R., & Thomas, T. L. (2000). Stress hormones in psychophysiological research. In J. T. Cacioppo, L. G., Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology* (2nd ed., pp. 342-367). New York, NY: Cambridge University Press.

Maire, M., Reichert, C.F., & Schmidt, C. (2013). Sleep-wake rhythms and cognition. *Journal of Cognitive and Behavioral Psychotherapies*, *13*(1a), 133-170.

Makeig, S., & Jung, T.P. (1995). Changes in alertness are a principal component of variance in the EEG spectrum. *NeuroReport, 7*, 213-217.

Marcotte, T. D., Meyer, R. A., Hendrix, T., & Johnson, R. (2013). The relationship between realtime EEG engagement, distraction and workload estimates and simulator-based driving performance. In *Proceedings of the Seventh International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design, Bolton Landing, NY.*

Massin, M.M., Maeyns, K., Withofs, N., Ravet, F., & Gérard, P. (2000). Circadian rhythm of heart rate variability. *Archives of Disease in Childhood, 83*(2), 179-182.

Matthews, G., & Desmond, P. A. (2002). Task-induced fatigue states and simulated driving performance. *The Quarterly Journal of Experimental Psychology, 55A*(2), 659-686.

Miller, C.A., & Parasuraman, R. (2007). Designing for flexible interaction between humans and automation: Delegation interfaces for supervisory control. *Human Factors, 49*(1), 57-75.

Molloy, R., & Parasuraman, R. (1996). Monitoring an automated system for a single failure: Vigilance and task complexity effects. *Human Factors, 38*(2), 311-322.

Morris, C. H., & Leung, Y. K. (2006). Pilot mental workload: How well do pilots really perform? *Ergnomics, 49*(15), 1581-1596.

Mouloua, M., Gilson, R., & Hancock, P.A. (2003). Designing controls for future unmanned aerial vehicles. *Ergonomics in Design, 11,* (6), 6-11.

Mouloua, M., Gilson, R., Kring, J., & Hancock, P. (2001). Workload, situation awareness, and teaming issues for UAV/UCAV operations. *Proceedings of the Human Factors and Ergonomics Society 45th Annual Meeting.*

Natale, V., & Cicogna, P.C. (2002). Morningness-eveningness dimension: Is it really a continuum? *Personality and Individual Differences, 32,* 809-816.

Oginska, H., Fafrowicz, M., Golonka, K., Marek, T., Mojsa-Kaja, J., & Tucholska, K. (2010). Chronotype, sleep loss, and diurnal pattern of salivary cortisol in a simulated daylong driving. *Chronobiology International, 27*(5), 959-974.

Parasuraman, R. (1979). Memory load and event rate control sensitivity decrements in sustained attention. *Science, 205*, 924-927.

Parasuraman, R. (2000). Designing automation for human use: Empirical studies and quantitative models. *Ergonomics, 43*(7), 931-951.

Parasuraman, R., Bahri, T., & Molloy, R. (1992). *Adaptive automation and human performance: I. Multi-task performance characteristics* (NAW-CAD-WAR 92035-60). Warminster, PA: Naval Air Warfare Center.

Petros, T.V., Beckwith, B.E., & Anderson, M. (1990). Individual differences in the effects of time of day and passage difficulty on prose memory in adults. *British Journal of Psychology, 81*, 63-72.

Previc, F.H. (1998). The neuropsychology of 3-D space. *Psychological Bulletin, 124*(2), 123-164.

Prinzel, L. J., Freeman, F. G., & Prinzel, H. D. (2005). Individual differences in complacency and monitoring for automation failures. *Individual Differences Research, 3*(1), 27-49.

Rabinowitz, Y.G., Breitbach, J.E., & Warner, C.H. (2009). Managing aviator fatigue in a deployed environment: The relationship between fatigue and neurocognitive functioning. *Military Medicine, 174*(4), 358-362.

Reis, C., Mestre, C., & Canhão, H. (2013). Prevalence of fatigue in a group of airline pilots. *Aviation, Space, and Environmental Medicine*, *84*(8), 828-833.

Roach, G. D., Sargent, C., Darwent, D., & Dawson, D. (2012). Duty periods with early start times restrict the amount of sleep obtained by short-haul airline pilots. *Accident Analysis and Prevention, 455*, 22-26.

Roeser, K., Meule, A., Schwerdtle, B., Kübler, A., & Schlarb, A.A. (2012). Subjective sleep quality exclusively mediates the relationship between morningness-eveningness preference and self-perceived stress response. *Chronobiology International, 29*(7), 955-960.

Rogers, N.L., Dorrian, J., & Dinges, D.F. (2003). Sleep, waking, and neurobehavioural performance. *Frontiers in Bioscience, 8*, 1056-1067.

Salas, E., Jentsch, F., & Maurino, D. (2010). *Human factors in aviation.* Academic Press.

Santiago-Espada, Y. (2014). MATB-II: Revised multi-attribute task battery. Retrieved from http://matb.larc.nasa.gov/

Santiago-Espada, Y., Myer, R. R., Latorella, K. A., & Comstock, J. R. (2011). The multi-attribute task battery II (MATB-II) software for human performance and workload research: A user's guide (NASA/TM–2011–217164). *Hampton, Virginia: NASA Langley Research Center*.

Schmidt, C., Collette, F., Cajochen, C., & Peigneux, P. (2007). A time to think: Circadian rhythm in human cognition. *Cognitive Neuropsychology, 24*(7), 755-789.

Smith, M. E., Gevins, A., Brown, H., Karnik, A., & Du, R. (2001). Monitoring task loading with multivariate EEG measures during complex forms of human-computer interaction. *Human Factors, 43*(3), 366-380.

Stanton, N. A., Young, M., & McCaulder, B. (1997). Drive-by-wire: The case of driver workload and reclaiming control with adaptive cruise control. *Safety Science, 27*(2-3), 149-159.

Stark, J. M., Scerbo, M. W., Freeman, F. G., & Mikulka, P. J. (2000, July). Mental fatigue and workload: Effort allocation during multiple task performance. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 44, No. 22, pp. 863-866). SAGE Publications.

Stikic, M., Johnson, R. R., Levendowski, D. J., Popovic, D. P., Olmstead, R. E., & Berka, C. (2011). EEG-derived estimators of present and future cognitive performance. *Frontiers in Human Neuroscience, 5*(70), 1-13.

Taillard, J., Philip, P., & Bioulac, B. (1998). Morningness/eveningness and the need for sleep. *Journal of Sleep Research, 8*, 291-295.

Van Dongen, H. P. A., Maislin, G., Mullington, J. M., & Dinges, D. F. (2003). The cumulative cost of additional wakefulness: Dose-response effects on neurobehavioral functions and sleep physiology from chronic sleep restriction and total sleep deprivation. *Sleep, 26*(2), 117-126.

Warburton, D. M. (1986). A state model for mental effort. In *Energetics and human information processing* (pp. 217-232). Springer Netherlands.

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

Warm, J.S., Parasuraman, R., & Matthews, G. (2008). Vigilance requires hard mental work and is stressful. *Human Factors, 50*(3), 433-441. doi: 10.1518/00187200X31252

Wechsler, D. 1981. *Wechsler Adult Intelligence Scale-Revised*. San Antonio: Psychological Corporation.

West, R., Murphy, K.J., Armilio, M.L., Craik, F.I.M., & Stuss, D.T. (2002). Effects of time of day on age differences in working memory. *The Journal of Gerontology, 57B*(1), 3-10.

Wickens, C. D. (1980). The structure of attentional resources. In *Attention and performance* III. Psychology Press.

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

Wickens, C.D. (2008). Multiple resources and mental workload. *Human Factors, 50*, 449-455.

Wickens, C. D., Hutchins, S. D., Laux, L., & Sebok, A. (2015). The impact of sleep deprivation on complex cognitive tasks: A meta-analysis. *Human Factors, 57*(6), 930-946.

Wickens, C.D., Sandry, D., & Vidulich, M. (1983). Compatibility and resource competition between modalities of input, output, and central processing. *Human Factors, 25*, 227-248.

Wickens, C. D., & Tsang, P. S. (2015). Workload. In Boehm-Davis, D. A., Durso, F. T., & Lee, J. D. (Eds.), *Handbook of human systems integration* (pp. 277-292). Washington, D.C. :American Psychological Association.

Wiggins, M. W. (2011). Vigilance decrement during a simulated general aviation flight. *Applied Cognitive Psychology, 25*(2), 229-235.

Wilson, G. F., Caldwell, J. A., & Russell, C. A. (2007). Performance and psychophysiological measures of fatigue effects on aviation related tasks of varying difficulty. *The International Journal of Aviation Psychology*, *17*(2), 219-247.

Wilson, G. F. (2002). An analysis of mental workload in pilots during flight using multiple psychophysiological measures. *The International Journal of Aviation Psychology, 12*(1), 3-18.

Winget, C. M., DeRoshia, C. W., Markley, C. L., & Holley, D. C. (1984). A review of human physiological and performance changes associated with desynchronosis of biological rhythms. *Aviation, Space, and Environmental Medicine*, *55*(12), 1085-1096.

Wittman, M., Dinich, J., Merrow, M., & Roenneberg, T. (2006), Social jetlag: Misalignment of biological and social time. *Chronobiology International, 23*(1&2), 497-509. doi: 10.1080/07420500545979

Wright, N., & McGown, A. (2004). Involuntary sleep during civil air operations: Wrist activity and the prevention of sleep. *Aviation, Space, and Environmental Medicine, 75*(1), 37-45.

Young, M. S., Brookhuis, K. A., Wickens, C. D., & Hancock, P. A. (2015). State of science: Mental workload in ergonomics. *Ergonomics*, *58*(1), 1-17.

Young, M. S., & Stanton, N. A. (2001). Size matters: The role of attentional capacity in explaining the effects of mental underload on performance. In D. Harris (Ed.), *Engineering psychology and cognitive ergonomics: Volume 5. Aerospace and transportation systems* (pp. 357-364). Aldershot, UK: Ashgate.