

Fall 2002

The Performance Effects of a Low Dose of Caffeine on a Cognitive Vigilance Task

Suzanne K. Robinson

Embry-Riddle Aeronautical University - Daytona Beach

Follow this and additional works at: <https://commons.erau.edu/db-theses>



Part of the [Cognitive Psychology Commons](#)

Scholarly Commons Citation

Robinson, Suzanne K., "The Performance Effects of a Low Dose of Caffeine on a Cognitive Vigilance Task" (2002). *Theses - Daytona Beach*. 176.

<https://commons.erau.edu/db-theses/176>

This thesis is brought to you for free and open access by Embry-Riddle Aeronautical University – Daytona Beach at ERAU Scholarly Commons. It has been accepted for inclusion in the Theses - Daytona Beach collection by an authorized administrator of ERAU Scholarly Commons. For more information, please contact commons@erau.edu.

THE PERFORMANCE EFFECTS OF A LOW DOSE OF
CAFFEINE ON A COGNITIVE VIGILANCE TASK

by

SUZANNE K. ROBINSON
B.S., Embry-Riddle Aeronautical University, 2000

A Thesis Submitted to the
Department of Human Factors & Systems
in Partial Fulfillment of the Requirements for the Degree of
Master of Science in Human Factors & Systems

Embry-Riddle Aeronautical University
Daytona Beach, Florida
Fall, 2002

UMI Number: EP31899

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

UMI[®]

UMI Microform EP31899
Copyright 2011 by ProQuest LLC
All rights reserved. This microform edition is protected against
unauthorized copying under Title 17, United States Code.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Copyright by Suzanne K. Robinson, 2002
All Rights Reserved


THE PERFORMANCE EFFECTS OF A LOW DOSE OF
CAFFEINE ON A COGNITIVE VIGILANCE TASK


by


Suzanne K. Robinson

This thesis was prepared under the direction of the candidate's thesis committee chair, Steven Hall, Ph.D., Department of Human Factors & Systems, and has been approved by the members of the thesis committee. It was submitted to the Department of Human Factors & Systems and has been accepted in partial fulfillment of the requirements for the degree of Master of Science in Human Factors & Systems.

THESIS COMMITTEE:


Steven M. Hall, Ph.D., Chair


Jon French, Ph.D., Member


Dennis A. Vincenzi, Ph.D., Member


MS HFS Program Coordinator


Department Chair, Department of Human Factors & Systems

ACKNOWLEDGEMENTS

The author wishes to express sincere gratitude to the thesis chair, Dr. Hall, for his continual support and guidance throughout the entire thesis process. A special thanks is also due to the committee members, Dr. French and Dr. Vincenzi, for lending their experience and enthusiasm to this project.

Deepest appreciation is due to Eric, Dale, Mike, and Lynn Robinson for their unconditional support, understanding, and encouragement. A thank you is also due to Cheryl Urquhart and Angela Wendell for their selfless commitment of time and energy.

Finally, my deepest thanks to Stephen Warren for the late nights and early mornings, the emotional highs and lows, and for sticking by my side from the beginning to the conclusion of this work.

ABSTRACT

The purpose of this study was to analyze the performance effects of a low and high dose of caffeine on a Bakan cognitive vigilance task. 69 student volunteers participated in the experiment. Participants were randomly distributed among caffeine dosage levels of 0, 20, and 200 mg. The correct response score, which was chosen as the dependent variable, was collected by the vigilance program, however reaction time and false alarm data was also evaluated. These scores were analyzed over time blocks (first, second, third, or fourth ten minute period of the forty minute task). A 3 x 4 mixed design ANOVA was performed on each of these data sets to determine if significant mean differences were present. The Stanford Sleepiness Scale was used to evaluate arousal levels before and after the task among caffeine conditions. The NASA TLX was also implemented post-task to evaluate task difficulty between caffeine conditions.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	iii
ABSTRACT	iv
LIST OF TABLES	ix
LIST OF FIGURES	x
LIST OF ABBREVIATIONS	xi
INTRODUCTION	1
Statement of the Problem	2
Review of the Literature	3
Vigilance	3
Vigilance Theories	3
Individual State Theories	4
Vigilance Tasks	5
Vigilance Performance Measures	6
Caffeine	7
Bioavailability of Caffeine	8
Caffeine and Arousal	8
Theories	8
Circadian Rhythms	9
Event-Related Potentials	10
Sleepiness Scale	11

Caffeine and Information Processing	13
Caffeine and Performance	14
Caffeine Dose-Response Relationship	15
Individual Differences	19
Age	19
Gender	19
Obesity	19
Smoking	20
Tolerance	20
Summary	20
Statement of the Hypothesis	22
METHOD	24
Participants and Design	24
Test Battery	25
Vision Test	25
Health Survey	25
Stanford Sleepiness Scale	25
Bakan Cognitive Vigilance Task	25
Introductory Screen	26
Practice Task	26
Trial Session	26
Second Stanford Sleepiness Scale	27
NASA TLX Rating Scale	27

Apparatus	27
Performance Measurement	28
Design	28
Procedure	29
RESULTS	33
DISCUSSION	47
CONCLUSIONS & RECOMMENDATIONS	54
REFERENCES	56
APPENDICES	
Appendix A. Recruitment Script	64
Appendix B. Consent Form/Experiment Briefing	66
Appendix C. Health Survey	69
Appendix D. Stanford Sleepiness Scale	71
Appendix E. NASA TLX	73
Appendix F. Tables of Mean Differences	77
Appendix G. Tables of Subsequent Tests	83

LIST OF TABLES

Table 1.	Caffeine Content of Selected Beverages and Food.	18
Table 2.	Results of Analysis of Variance for Detection Rate	36
Table 3.	Detection Rate Pairwise Comparisons	37
Table 4.	Results of Analysis of Variance for False Alarms	39
Table 5.	False Alarm Pairwise Comparisons	39
Table 6.	Results of Analysis of Variance for Response Times	41
Table 7.	Response Times Pairwise Comparisons	41
Table 8.	Results of Analysis of Variance for Pre- and Post-Task Stanford Sleepiness Scale Scores	43
Table 9.	Stanford Sleepiness Scale Pairwise Comparisons	44
Table 10.	Results of Analysis of Variance for NASA TLX scores	45
Table 11.	NASA TLX Pairwise Comparisons	46

LIST OF FIGURES

Figure 1.	The Main Effects of Caffeine on the Bakan Vigilance Task	37
Figure 2.	The Effects of Time and Caffeine on False Alarm Rates	40
Figure 3.	The Effects of Time and Caffeine on Response Times	42
Figure 4.	The Effects of a Vigilance Task on Stanford Sleepiness Scale Scores	44

LIST OF ABBREVIATIONS

ANOVA	Analysis of Variance
CM	Centimeter
ERP	Event-Related Potentials
GG	Greenhouse-Geisser
LSD	Least Significant Difference
MG	Milligrams
MSLT	Multiple Sleep Latency Test
NASA	National Aeronautics and Space Administration
OZ	Ounce
PIN	Participant Identification Number
SA	Sphericity Assumed
SCN	Suprachiasmatic Nucleus
SSS	Stanford Sleepiness Scale
TBSP	Tablespoon
TLX	Task Load Index

INTRODUCTION

The ability to maintain alertness is often a determining factor between professional failure and success. Alertness refers to a person's receptivity to external stimuli. The term vigilance refers to a state of maximum efficiency, or maximum receptivity (Davies & Parasuraman, 1982). However, when individuals encounter a situation requiring vigilance their performance tends to decrease over time. This deterioration is referred to as the vigilance decrement (Davies & Parasuraman, 1982). The demonstration of this decrement through the considerable amount of published studies has been tremendously consistent (Mackie, 1987). The vigilance decrement is a growing concern as an increasing number of professions require 24-hour operations with long hours of monotonous shift work.

Caffeine is a stimulant that is commonly found in popular foods and beverages. Research into the performance-enhancing effects of this stimulant has demonstrated its potential for restoring alertness (Lieberman, 1992). However, a criticism of caffeine research is that doses administered are often many times greater than what is regularly consumed through a single dietary serving (refer to table 1) (Stelt & Snel, 1998; Durlach, 1998). Therefore, research concerning the effects of relatively low doses of caffeine on performance could enhance the external validity of similar studies since the laboratory results would be comparable to effects generated through daily dietary or medical caffeine consumption.

Low caffeine dose studies are rare. Lieberman, Wurtman, Emde, Roberts, & Coviella (1987) examined the impact of low-doses of caffeine on a Wilkinson (1970) auditory vigilance task. Surprisingly, their results indicated that maximum performance occurred in the lowest caffeine condition (32 mg). However, auditory vigilance tasks are typically considered sensory not cognitive tasks (Davies & Parasuraman, 1982). Yet sensory vigilance tasks constitute the majority of tasks used in vigilance research. It is arguable that cognitive vigilance tasks contain greater external relevance than sensory based tasks because they contain a greater similarity to that of operational monitoring tasks like air traffic control or nuclear power facility monitors or computer watch keeping tasks in general. Although research exists on the relationship between caffeine and cognitive vigilance task performance (for example, Frewer & Lader, 1991), the vast majority of these involve large doses.

Statement of the Problem

This investigation proposed to explore the effects of low doses of caffeine on a cognitive vigilance task. The application of the results have the potential to address numerous safety and job efficiency aspects of our 24 hour society as related to performance on vigilance tasks. The purpose of this study is to determine if a low dose of caffeine will reduce the vigilance decrement on a cognitive vigilance task.

This research will also evaluate the theoretical mechanisms of arousal and workload which may be responsible for the vigilance decrement, and the performance enhancing properties of caffeine on this task.

Review of the Literature

Vigilance

Vigilance tasks are defined as those that require directed attention over continuous, long periods, with the purpose of detecting small changes. Such tasks are commonly referred to as monitoring or watch keeping tasks (Davies & Parasuraman, 1982).

Mackworth (1950) was among the first to formally explore the phenomenon that has become known as the operational vigilance decrement. Mackworth designed a task to measure vigilance performance, known as the clock task. Originally created to mimic a radar operator's workstation, the task consisted of a circular-rotating pointer that sequentially paused, similar to a clock's second hand. Participants were instructed to trigger a switch whenever the pointer missed a pause, making a 'double jump'. The duration of the task was two hours. The results of this experiment identified that performance degrades over the entire session, with the majority of the decline occurring between the first thirty minutes, and the second thirty minutes.

In one attempt to reduce the vigilance decrement, Mackworth administered 10 mg of Benzedrine to participants before testing. Benzedrine is a form of amphetamine, a powerful central nervous stimulant. Mackworth found that this stimulant reduced the vigilance decrement.

Vigilance Theories

Since Mackworth, the study of vigilance has greatly developed. Researchers have created various theories to explain vigilance decrements, and to predict how decrements will affect an individual's performance.

One of the first vigilance theories is the inhibition theory, developed by Mackworth (1950). His explanation for the vigilance decrement was that when a response was not reinforced, it eventually disappeared. To counteract the vigilance decrement Mackworth introduced knowledge of results and rest pauses; both interventions eliminated the vigilance decrement. Therefore, inhibition builds like fatigue with every stimuli occurrence that is not reinforced and when few to zero rest pauses are permitted. Eventually, inhibition builds to an extent where the conditioned response fails to exist. However, the main argument against this theory is that increasing signal frequency reduces the vigilance decrement, rather than impairing performance.

The filter theory was developed by Broadbent (1957). This theory suggests that all humans have an internal filter, which intermittently fails to register information due to 'internal blinks'. The rate of these failures increases with time on task. This results in decreasing performance as a function of increasing signal presentation rate, as the likelihood of a signal presented during an 'internal blink' increases.

The expectancy theory explains vigilance task performance through the observer estimating the probability of signal presentation, based on past signals (Baker, 1959). Therefore, detection increases if signals are presented to the participant at regular intervals that allow accurate prediction.

Individual State Theories

The motivation theory, introduced by Smith (1966), explains the vigilance decrement as a result of the lack of motivation of individuals. Smith argued that all persons are capable of maintaining vigilance for a few hours with no mistakes, however

they are not motivated to do so in professional or laboratory environments due to the monotony of vigilance tasks and the lack of intrinsic motivation.

The habituation theory occurs as a result of repeated exposures to a once new stimulus producing progressively smaller behavioral responses. Habituation explains the reduction in performance resulting from repeated stimulation, generally worsening with presentation rate. Jerison and Pickett (1964) found that rapid signal rates produced lower performance on vigilance tasks. This finding supports habituation as when stimuli are repeated more frequently, their shock value decreases and is eventually eliminated.

The arousal theory states that the monotonous nature of vigilance tasks causes the alertness level of the central nervous system to diminish. This causes a decrease in responsiveness and efficiency, which results in a performance decrement (Davies & Parasuraman, 1982). This theory was evaluated by administering the Stanford Sleepiness Scale (Appendix D), a measure of alertness, both before and after the 40 minute cognitive vigilance task. Theoretically, participants were predicted to be at a greater level of sleepiness after the task than beforehand. This theory also suggests that caffeine, a central nervous stimulant, has potential to reduce the vigilance decrement. This was evaluated by comparing the Stanford Sleepiness Scores of the placebo condition to the caffeine conditions.

Vigilance Tasks

Although there are a number of vigilance tasks used in research, most can be classified as either a sensory or a cognitive task. The majority of vigilance tasks are sensory tasks, as they use sensory signals such as auditory tones. Cognitive vigilance

tasks are less common, and typically use numbers or letters as presentation stimuli (Davies & Parasuraman, 1982).

The Bakan task is the most common cognitive vigilance task. This difficult task requires information to be continually held in working memory. This task presents a series of seemingly random numbers over an extended period. Participants are asked to signal the occurrence of three sequential even or odd digits (Bakan, 1959).

This task is considered cognitive, as it requires participants to make several discriminations from the presentation of seemingly random digits. These discriminations include the successiveness of digits, oddness-evenness of digits, identity of digits, and memory for previous digits while watching for current digits (Bakan, 1959). This task has been shown to produce reliable vigilance decrements over time (Harkins, Nowlin, Ramm, & Schroeder, 1974).

Vigilance Performance Measures

Davies and Parasuraman (1982) stated that there are three measures used by researchers to assess subject proficiency on vigilance tasks. The first measure is the number of signals correctly indicated by the participant, usually called the 'detection rate' or 'correct response'. The second is the amount of time between the presentation of the signal and the participant reaction input, known as 'detection latency' or 'response times'. The final measure is the number signals falsely reported by the participant, termed 'false positives' or 'false alarms'.

The third measure mentioned represents a newer view in vigilance research, suggesting that the vigilance decrement is caused by a steady increase in fatigue, which may result in an increased number of false positives. Bakan (1959) explains that fatigue

builds when participants stop attempting to make discriminations within the stimuli. When this happens, he or she instead watches a sequence of undifferentiated numbers. This produces monotony in the stimulation, which is conducive to drowsiness or sleep. This replaces the traditional view that the decrement is caused by deterioration in participant sensitivity to signals, resulting in a reduced detection rate (Davies & Parasuraman, 1982).

In fact, Bakan (1959) reports that feelings of drowsiness are very common among subjects in vigilance tasks. This view is interesting as it implies the question of whether implementing anti-fatigue strategies, such as caffeine, will improve performance on vigilance tasks.

Although the correct response score will be used to analyze performance on this task, all measures will be collected to create a thorough perspective of the vigilance decrement.

Caffeine

Caffeine is in a class of natural occurring substances termed methylxanthines. Two other methylxanthines, theobromine and theophylline, occur naturally in cocoa and tea respectively. Caffeine is naturally found in many plants including tealeaves, cocoa nuts, and coffee beans (Lieberman, 1992).

Caffeine is a chemical stimulant that is present in many popular foods and beverages. According to Nehlig, Daval, and Debry (1992), caffeine is considered the central nervous stimulant most widely consumed by humankind. In fact, hundreds of millions of people consume behaviorally active amounts of caffeine daily through various forms (Lieberman, 1992).

Shohet and Landrum (2001) determined that the mean daily intake of caffeine for College students was 228 mg/day. The participants consumed their caffeine in its various forms three times a day. However, age is positively correlated to caffeine intake, and it has been determined in the general population that men consume 349 mg/day and women 394 mg/day (Jacobson & Bouher, 1991).

Bioavailability of Caffeine

Caffeine is absorbed quickly and easily, diffusing throughout the entire human being, having a volume of distribution similar to body water and quickly penetrating into the brain (Nehlig et al., 1992). Caffeine levels in human plasma generally peak 15-45 minutes after oral ingestion (Bonati, Latini, Galletti, Young, Tognoni, & Garattini, 1982). In addition, Blanchard and Sawers (1983) found that gastrointestinal absorption of caffeine is 99% complete in about 45 minutes. Arnaud (1998) concluded that only a small percentage of caffeine dosage, 0.5-2%, is recovered in the urine. Therefore, practically all of administered dosages of caffeine are absorbed and utilized by the human body.

Caffeine and Arousal

Theories

The effects of caffeine on basal levels of arousal form an inverted U relationship between arousal and level of performance in cognitive tasks (Yerkes & Dodson, 1908). Lorist, Snel, and Kok (1994) concluded that the possible explanation for six of 30 subjects showing no caffeine related improvement is due to the Yerkes-Dodson theory. Thus, the absence of the arousal elevating effect of caffeine can be explained by

suggesting that caffeine increased arousal beyond an optimal level and therefore impaired performance in these subjects.

Broadbent (1971) suggested a compensatory system where lowered performance, due to low arousal, could be counteracted by increased subjective effort. Using 150 mg of caffeine, Linde (1995) determined that subjective tiredness increased significantly in subjects given a placebo versus caffeine, at midnight and 4 am. In addition to performance measures, a subjective rating of effort was collected using a magnitude estimation scale. It was determined that effort was significantly higher in the placebo condition. This evidence suggests a compensatory arousal mechanism. This theory was evaluated in the present study by administering the NASA-TLX rating scale to measure the difficulty level experienced by each participant after the 40 minute vigilance task. It was predicted that participants in the caffeine condition would experience less task difficulty than those in the placebo condition.

Circadian Rhythms

Circadian rhythms are physiological processes that cycle regularly (circa 24 hours) between a peak and a trough by internal biological “clocks”. These rhythms are exhibited by most organisms and determine optimal sleep wake cycles. Since it has been determined that circadian variations persist without natural light, these rhythms have been attributed to endogenous neural generators, specifically the suprachiasmatic nucleus (SCN) of the hypothalamus (Manly, Lewis, Robertson, Watson, & Datta, 2002). The SCN projects to the ventricular nucleus and other structures of the hypothalamus (Manly et al., 2002).

The circadian cycle is subjectively observable as a varying sense of tiredness, resulting in distractibility and reduced mental alertness. Daan, Beersma, and Borbely (1984) have determined that circadian related arousal is somewhat independent of sleep, as under some conditions of sleep deprivation the alertness cycle is maintained.

Klein, Herrmann, Kuklinski, and Wegmann (1977) evaluated performance based on normal circadian rhythms using scores on psychomotor, cancellation, digit summation, and two simulator tasks. It was apparent that scores rise during the day to a plateau between 1200 and 2100 hours, and decline to a minimum between 0300 and 0600 hours. The range of circadian oscillation showed a magnitude of between 10% and 30% of the 24-hour mean. It was noted that the shape of the performance curve and the range of oscillations are in good agreement with similar round the clock studies. Due to these findings, this study will conduct all testing sessions between 1200 and 2100 hours.

As arousal varies as a function of the time of day, Smith (1998) has determined that the performance enhancing effects of low-doses of caffeine are most pronounced in the early afternoon. However, Miller, Lombardo, and Fowler (1995) concluded that caffeine significantly increased arousal throughout the entire day.

Event-Related Potentials

Lorist, Snel, and Kok (1994) performed a study using event-related potentials (ERPs) to assess caffeine's effects on non-fatigued individuals. Caffeine was found to affect two ERP components, N1 and P3. The N1 and P3 are two components which occur 100 and 300 milliseconds after the stimulus, respectively. After caffeine treatment, an increasingly negative going N1 in combination with a shorter latency was produced. This result suggests that caffeine increases receptivity to external stimuli and accelerates

information processing. Concerning the second component, caffeine produced a more positive going P3. This increase in amplitude at the posterior electrode site represents an increase in phasic cortical arousal. It is also noted that both N1 and P3 are related to signal detection, however P3 reflects recognition and identification of stimuli while N1 reflects early information processing. Therefore, the P3 effects indicate an enhancement in the intensity of encoding compared to the placebo condition.

The authors also described a well-known finding which states that larger the amplitude of the P3 component, the easier the task. Alternatively, the increased alertness and vigor reported by participants in the caffeine condition, combined with the increased P3 amplitude, suggest that the actual task complexity is perceived as being lower (Lorist, Snel, & Kok, 1994). This theory was evaluated in the current study by implementing the NASA-TLX scale to evaluate the workload experienced on the vigilance task (refer to Appendix E).

Sleepiness Scale

To minimize error in this study a standard measure of alertness is necessary to reduce variance among participants. A number of processes for measuring daytime sleepiness have been developed. These measures may be categorized into four general types: behavioral observation, laboratory test performance, subjective feelings of sleepiness, and physiological parameters (Carskadon, 1993). Several tests have been developed from these measures, however the scope of this evaluation will expand upon the latter two: subjective and physiological.

A common physiological measure is the Multiple Sleep Latency Test (MSLT). This test measures the time for a participant to fall asleep, and has been found to be

extremely reliable in both normal and sleep-disturbed individuals (Carskadon, & Dement, 1979; Chervin, Aldrich, Pickett, & Guilleminault, 1997; Lichstein, Wilson, Noe, Aguillard, & Bellur, 1994). Although this measure is often considered the “gold standard” of sleepiness tests, it is extremely costly and nearly impossible to administer outside of a laboratory setting (Pilcher, Schoeling, & Prosansky, 2000).

Subjective scales are an alternative measure of sleepiness. Such scales have been developed to measure the feeling of sleepiness that individuals subjectively experience. A well-known measure of subjective sleepiness is the Stanford Sleepiness Scale (SSS) (Hoddes, Dement, & Zarcone, 1972). This scale categorizes sleepiness from a level one to seven. A level one participant subjectively rates themselves as “feeling active, vital, alert, or wide awake” where a level seven is “no longer fighting sleep, sleep onset soon; having dream-like thoughts.” Refer to Appendix D for an example of the SSS.

Danker-Hopfe, Kraemer, Dorn, Schmidt, Ehlert, and Herrmann (2001) reported a confirmed difference between subjective sleepiness (SSS) and physiological sleepiness (MSLT). The MSLT physiologically reflects the event of falling asleep, whereas the SSS reflects the participant’s subjective estimation of their actual capacity to perform, yet these aspects are different in nature. The physiological test measures the likelihood of filling the biological sleep need, where the subjective test reflects a complex cognitive state influenced by attention level, motivation, anxiety, and so forth. Therefore, the SSS reflects high cognitive and emotional functions (Danker-Hopfe, Kraemer, Dorn, Schmidt, Ehlert, & Herrmann, 2001). This may explain why little correlation is often found between physiological and subjective tests of sleepiness (Johnson, Freeman, Spinweber, & Gomez, 1991).

Manly et al. (2002) performed an experiment assessing the time-of-day effects on a sustained attention task and two subjective measures of sleepiness (the SSS, and the visual analogue scale of sleepiness). It was determined that the sleepiness scales were highly correlated with significant variation between the periods of 1 am, 7 am, 1 pm, and 7 pm. Sleepiness levels peak at 7 am and were lowest in the afternoon and evening sessions, rising again at 1 am. Accuracy on the sustained attention task was greatest in the afternoon and evening, corresponding to the responses on the sleepiness scales. These data support the findings of Klein, Herrmann, Kuklinski, and Wegmann (1977), and lend evidence to the accuracy of sleepiness scales as a predictor of performance.

Due to the time-consuming and costly nature of the MSLT, it is ill suited for this study. However, the combination of the SSS's ease of administration, low cost, and performance prediction capabilities nicely fulfill the requirements of a sleepiness scale within this experiment.

Caffeine and Information Processing

Caffeine produces diverse and complex effects, even when administered in small quantities (Lorist, 1998). Since caffeine is a potent adenosine antagonist at adenosine receptor sites, an understanding of the function of adenosine molecules is essential when exploring the effects of caffeine. Adenosine is produced as a byproduct of cellular activity. When adenosine binds to adenosine receptors it slows the activity of the neurons, causing a decrease in arousal. Several electrophysiological, behavioral, and biochemical studies have found adenosine to have potent inhibitory actions (Hirsh, 1984). Caffeine's beneficial effects on cognitive functioning are the result of blocking (in effect taking the place of) adenosine receptors at multiple sites throughout the brain (Phillis,

1991). Caffeine inhibits the binding of adenosine to its receptor sites causing an increase in central nervous system activity (Lieberman, 1992).

Until the blocking of adenosine receptors was identified as the cause of caffeine's behavioral effects, several other mechanisms were hypothesized. The most common were: calcium mobilization, phosphodiesterase inhibition, and prostaglandin antagonism. However, the adenosine receptor effect occurs at a much lower concentration of caffeine than the other mechanisms. In fact, these other mechanisms may account for some toxic caffeine effects at high doses (Lieberman, 1992).

Through its effects on adenosine receptors, caffeine indirectly affects the noradrenaline, acetylcholine, and dopamine neurotransmitter systems. These systems have different functions that might affect alertness and hence information processing activities (Lorist, 1998).

Overall, it is evident that performance benefits produced by caffeine are the result of a complex combination of factors. Initially caffeine blocks adenosine receptors, which facilitates excitatory neurotransmitters that indirectly effect noradrenaline, acetylcholine and dopamine. These neurotransmitter systems are linked to the energetical mechanisms of effort, arousal, and activation; which are responsible for changes in efficiency. These mechanisms then influence performance on human information processing activities (Lorist, 1998). This sequence of events describes the relationship between caffeine and the resulting improvement in human information processing.

Caffeine and Performance

Hollingworth (1912) was one of the first researchers to study the performance effects of caffeine. His groundbreaking work assessed the effects of caffeine on several

mental performance measures. Hollingworth's overall conclusion was that caffeine facilitates performance.

Since that time a great deal of research has been performed on many aspects of human performance. After reviewing eighty-five studies on the effects of caffeine and mental performance, Stelt and Snel (1998) identified several general features. The studies were evenly distributed between within-subjects and between-subjects methodologies. The advantage of the former being increased power, and of the latter being the elimination of carry-over effects. The studies generally limited participants to low-to-moderate daily caffeine consumers (about 200-300 mg), and to those in good health. Another predominant feature was the administration of a caffeine dosage that is significantly larger (200-300 mg and greater) than that contained in a standard single serving of food or a beverage. The advantage of a high dose of caffeine is the exaggeration of potential effects, the disadvantage being reduced external validity, as equivalent doses may be rare in normal dietary sources.

The present study used a log caffeine dose progression of 0, 20, and 200 mg. The high dose of 200 mg was used as a positive control to demonstrate the test conditions are sensitive to caffeine.

Caffeine Dose-Response Relationship

Few studies utilize ranges of caffeine doses contained in a typical serving of soda, tea, or a mug of coffee (Smit & Rogers, 2000; Durlach, 1998; Lieberman et al., 1987). Lieberman et al. (1987) stated that they were not aware of any previous study where a dose below 75 mg was administered to assess performance effects.

Hasenfratz and Battig (1994) investigated the dose-effect relationship of caffeine on mental performance using a rapid information processing task. Participants received approximately 0, 100, 200, or 400 mg of caffeine. The resulting performance curves were surprisingly heterogeneous. It was found that increasing dosages caused increased effects for alpha- and beta- EEG frequencies, anxiety, and wakefulness. However, increasing dosages decreased performance on the rapid information processing rate and blood pressure. No apparent relationship was observed for response time. The authors concluded that the doses with beneficial effects are at the lower end of the tested dose range, comparable to those found in dietary sources.

Smit and Rogers (2000) explored the effects of low doses of caffeine on performance using 0, 12.5, 25, 50, and 100 mg doses. It was found that caffeine could significantly improve cognitive performance at dose ranges even lower than those found in a single serving of a popular beverage. Using a rapid information processing task it was demonstrated that a very flat dose-response relationship was formed. Surprisingly, all doses of caffeine improved performance to an almost equal extent.

Conducting a naturalistic observation of the effects of day-long consumption of tea and coffee, it was concluded that tea consumption produces similar alerting effects to coffee, despite lower caffeine levels, and is less likely to produce side-effects such as sleep disruption (Hindmarch, Rigney, Stanley, Quinlan, Rycroft, & Lane, 2000).

However, absolute agreement does not exist on the effects of low-doses of caffeine. In fact, Kamimori, Penetar, Headley, Thorne, Otterstetter, and Belenky (2000) found no alerting affects associated with their low dose of 150 mg using the Stanford Sleepiness Scale and a response time measure, suggesting that this dose had no affect on

alertness. Several possible explanations were offered for this phenomenon. The first being the occurrence of a threshold effect, which is known to be linked with caffeine. This effect requires a minimum blood concentration of a drug to produce a significant alerting affect. A second explanation may be related to individual caffeine tolerance. Subjects self-classified themselves as moderate to low caffeine users, however participants may have actually been high caffeine consumers. Chronic caffeine use results in an increase in tolerance to physiological and psychological effects. The final explanation offered for the lack of effects in their low dose group may be due to high individual variability.

Although minimal research exists on the effects of caffeine in low doses, the research that has been performed demonstrates significant results if adequate controls are employed. These results demonstrate a need for continued low dose research, to extend the results to other cognitive tasks associated with a vigilance decrement like the Bakan task.

Table 1. Caffeine Content of Selected Beverages and Food.

Item	Caffeine content (mg)
Coffee (5 oz cup)	
Drip method	90-150
Percolated	64-124
Instant	40-108
Decaffeinated	2
Tea, loose or bags (5 oz cup)	
1-minute brew	9-33
3-minute brew	20-46
5-minute brew	20-50
Tea products	
Instant (5 oz cup)	12-28
Iced tea (12 oz can)	22-36
Chocolate products	
Hot cocoa (6 oz)	2-8
Milk chocolate (1 oz)	1-15
Baking chocolate (1 oz)	35
Sweet dark chocolate (1 oz)	5-35
Chocolate-flavored syrup (2 tbsp)	4
Cola beverages (12 oz)	
Coca-Cola Classic	46
Pepsi	38
Diet Pepsi	36
Diet Coke	46
TAB	46
Other soft drinks (12 oz)	
Dr Pepper	41
Mountain Dew	54
Mellow Yellow	52
Mr. Pibb	40
Other products	
Midol	32
Excedrine	65
NoDoz	100
Vivarin	200

(Institute of Food Technologists' Expert Panel, 1987;
Shohet & Landrum, 2001)

Individual Differences

Age

Research has determined that no significant differences exist in caffeine metabolism between age groups (Grant, Tang, & Kalow, 1983; Campbell, Speilberg, & Kalow, 1987). Based on these findings, participants were recruited without age restrictions.

Gender

Caffeine effects vary between the genders, according to the hormonal state of the female. Balogh, Irmisch, Klinger, Splinter, and Hoffmann (1987) found that there is a 25% increase in caffeine elimination found in the luteal phase of the menstrual cycle compared to the follicular phase. Arnaud (1993) found that oral contraceptives double the half-life of caffeine, and that the half-life was prolonged during the last trimester in pregnant women. For these reasons, females who are pregnant or taking oral contraceptives were be eliminated from this study.

Obesity

Physical composition is a factor in the absorption of caffeine. It is common sense that 100 mg of caffeine will have a stronger affect on a 90 pound woman than a 250 pound man as the stimulant is in greater proportion to the total body mass. In addition, research has shown that obesity affects caffeine distribution. Kaminori, Somani, Knowlton, and Perkins (1987) found that in obese participants with more than 30% body fat, a larger caffeine distribution volume was observed. Significantly higher absorption rate constants, lower elimination rate constants and a longer mean serum half-life were also reported in obese vs. non-obese participants. It was concluded that caffeine

distribution was incomplete into excess body fat. For this reason, obese individuals were asked not to participate in this study.

Smoking

Research has demonstrated that caffeine clearance is stimulated by smoking (Caraco, Zylber-Katz, Barry, & Levy, 1995; Kotake, Schoeller, Lambert, Baker, Schaffer, & Josephs, 1982; May, Jarboe, Van Bakel, & Williams, 1982). Therefore, only non-smokers were recruited as participants.

Tolerance

Tolerance is experienced when individuals continuously consume and become adapted to the stimulant effects of a drug, and increased dosage is required to produce similar physiological effects. Nehlig et al. (1992) have shown that regular consumption of 6-11 cups of coffee a day is likely to produce effects that are not entirely counterbalanced by tolerance. Therefore, the central nervous system is only slightly tolerant to the stimulant effects of caffeine. To standardize tolerance levels among individual participants, all persons who wished to take part in this experiment must have self-classified their daily caffeine intake (based on Table 1). If their daily intake was moderate (200-400 mg) they were asked to participate, all others were excluded.

Summary

This literature review has explored the phenomenon of the vigilance decrement, and the performance enhancing properties of caffeine. Very little research has utilized caffeine doses small enough to be equivalent to those from dietary sources. Previous research has documented the reduction of the vigilance decrement through low-doses of caffeine on a sensory vigilance task. However, as vigilance tasks are classified as either

sensory or cognitive it is important to question if similar benefits exist on a cognitive task. The Bakan task is a cognitive vigilance task that has been found to produce reliable vigilance decrements over time. Therefore, the combination of low doses of caffeine (0, 20, and 200 mg) and a Bakan cognitive vigilance task offered a greater understanding of the performance enhancing potential of caffeine.

The arousal theory states that the monotonous nature of vigilance tasks causes the alertness level of the central nervous system to diminish. This causes a decrease in responsiveness and efficiency, resulting in the vigilance decrement. This theory suggests that caffeine, being a central nervous stimulant, has great potential to reduce or even eliminate this decrease in arousal. To evaluate this theory the Stanford Sleepiness Scale was implemented before and after the vigilance task in placebo and caffeine conditions.

The compensatory system suggests that low arousal is counteracted by increased subjective effort (Broadbent, 1971). It has been demonstrated that the arousal enhancing effects of caffeine reduced the subjective effort required to complete a task, compared to the placebo condition (Linde, 1995). This is supported by ERP findings which demonstrated that increased alertness and vigor reported by participants in caffeine conditions, combined with increased P3 amplitudes, suggested that actual task complexity was perceived as being lower (Lorist, Snel, & Kok, 1994). This theory was evaluated in the current study by implementing the NASA-TLX scale to evaluate the workload experienced on the vigilance task between caffeine and placebo conditions.

The NASA-TLX is a subjective task load assessment index. This index allows researchers to carry out subjective workload assessments on research participants performing a task. NASA-TLX is a multi-dimensional rating system, deriving an overall

score based on a weighted average of six subscale ratings. These subscales include Mental Demands, Physical Demands, Temporal Demands, Performance, Effort, and Frustration.

Statement of the Hypothesis

Research has demonstrated that a drop in performance, caused by a vigilance decrement, occurs within the first 30 minutes of a vigilance task (Mackworth, 1950). Based on knowledge of the vigilance decrement, the following hypothesis was formed:

Hypothesis: It was hypothesized that participant performance on a Bakan cognitive vigilance task would decrease as a function of time on task. It was also hypothesized that after receiving a dose of caffeine (20, 200 mg) participants would demonstrate improved performance over the placebo condition. It was hypothesized that participants would be more sleepy post-task than pre-task, and that participants in caffeine conditions will be less sleepy than those in the placebo condition, as measured by the Stanford Sleepiness Scale. It was also hypothesized that participants in the placebo condition will experience greater task difficulty than those in caffeine conditions.

Prediction One: It was predicted that the performance of all treatment levels would reduce as a function of time on task.

Prediction Two: It was predicted that the 20 and 200 mg doses of caffeine would reduce the vigilance decrement, thereby increasing performance over time compared to the placebo condition as measured by the detection rate.

Prediction Three: It was predicted that sleepiness levels would be significantly higher after the task in all conditions, when compared against pre-task scores.

Prediction Four: It was predicted that post-task Stanford Sleepiness Scale scores within the caffeine conditions (20 and 200 mg) would be significantly less sleepy than scores in the placebo condition.

Prediction Five: It was predicted that perceived task difficulty would be lower in the caffeine conditions (20 and 200 mg) compared to the placebo condition, as measured by the NASA-TLX.

METHOD

Participants and Design

Seventy-eight male and female participants were recruited from the student population at Embry-Riddle Aeronautical University to participate in this experiment. However, only seventy three participants completed the task. Out of the remaining five participants three refused participation after reading the consent form and were offered an alternative assignment for extra credit. One participant was asked to return at a later time due their low response on the Stanford Sleepiness Score and one was removed due to their caffeine ingestion that morning.

After completing the task, 4 data sets were excluded from the analysis due to individual characteristics of each participant. Refer to the results section for further details.

The remaining 69 participants consisted of 21 females and 48 males. The mean age of participants was 20.9 years of age, the youngest participant being 17 and the oldest participant being 33 years old.

The mean weight of male participants was 168.9 pounds, the lightest participant weighing 118 and the heaviest weighing 250 pounds. The mean weight of female participants was 137.8 pounds, with the lightest weighing 100 and the heaviest weighing 190 pounds.

The experimenter recruited subjects by reading the 'recruitment script' (refer to Appendix A) in classroom settings. Participants were offered extra credit for

participation. To reduce error variance, all participants self-classified themselves as being in good health, low-to-moderate daily caffeine consumers (200-400 mg a day), sleeping at least 8 hours a night, non-smokers, and not obese. In addition, female subjects must not have been consuming oral contraceptives, and must not have been pregnant. Participants must also have abstained from alcohol during the previous 24 hours, caffeine for 12 hours, and food for 3 hours before participation. All of these characteristics were identified as affecting an individual's reaction to caffeine. All testing sessions occurred within the period between 1200 and 2100 h to reduce time-of-day effects based on circadian rhythms.

Eligible participants were randomly placed into one of the three dosage conditions (0, 20, and 200 mg). A maximum of ten participants performed the vigilance task simultaneously. All participants wore ear plugs to limit distractions.

Detailed instructions were provided, and an informed written consent form was completed by participants before participation (refer to Appendix B). The University's ethical review board approved the study.

Test Battery

Vision Test

A standard vision test was utilized to ensure 20/20 natural or corrected vision. All participants possessed 20/20 natural or corrected vision.

Health Survey

All participants completed a health survey (refer to Appendix C). The survey covered physiological areas that affect caffeine absorption levels or performance. If

participants answered 'no' to any of the questions, they were removed from the study. One participant was removed from the study based on their survey response as they had recently ingested caffeine. This screening process was designed to minimize the effects of individual differences within the results.

Stanford Sleepiness Scale

This scale allowed a subjective assessment of alertness (refer to Appendix D). This information allowed participants to rate their alertness. If any participant rated themselves as a level four or lower, they were removed from the study to minimize individual sleepiness error. One participant was removed from the study based on this screening.

Bakan Cognitive Vigilance Task

This task measured cognitive vigilance through rapid information processing. Single numerical digits were presented on a 15" color computer monitor. Digits were presented every second throughout the practice session, and every 800 ms throughout the trial session. Stimulus size was 2 cm, at 90-degree visual angle at a 60 cm distance. Participants were required to press the space bar as quickly as possible when target sequences of three consecutive odd, or three consecutive even digits were detected. Sixty of these sequences were presented every 10 minutes; sequences were separated by a minimum of five and a maximum of 30 random digits. Identical number sequences were repeated for each 10 minute block to ensure that differences in scores were due to caffeine or time effects alone. From the instant the third digit in an odd or even sequence was presented, participants were given 1500 ms to register a correct response by pressing the space bar. A measure of response time (ms) was recorded for each correct response.

The response time was the number of milliseconds from the instant the third digit in a sequence was presented to the moment the space bar was pressed. The maximum response time was limited to 1500 ms. Any space bar responses made outside of this 1500 ms period were classified as a false alarm.

Introductory Screen

Participants were given instructions and asked to enter their participant identification number.

Practice Task

Participants then conducted a practice task to become familiar with the apparatus. The practice session was 2 minutes in duration. The signal presentation rate was 1 s.

Trial Session

Participants conducted a 40-minute session throughout which performance was recorded. Upon completion of the session, the computer monitor alerted participants that the task was finished. The program then returned to the introductory screen, ready for the next participant.

Second Stanford Sleepiness Scale

The participants then completed a second Stanford Sleepiness Scale. This measure allowed comparison of alertness levels before and after the Bakan cognitive vigilance task.

NASA-TLX Rating Scale

A NASA-TLX task difficulty scale was given to participants upon conclusion of the vigilance task (refer to Appendix E). This scale allowed a subjective rating of the

task difficulty. This rating provided additional information to clarify the mechanisms impacting caffeine's effects on performance.

Apparatus

The apparatus consisted of a 15-inch computer monitor with a standard keyboard. The spacebar of the keyboard was used by the participants to signal stimuli sequences.

The computer program utilized was specifically developed for this study. In Bakan's (1959) original study numbers were played audibly from a tape recorder and participants were asked to signal the sequences. However, to allow for increased precision in data collection the computer program was created. The specifics of this particular program have been adapted from a Bakan vigilance program developed by Frewer and Lader (1991).

The caffeine pills were created by a local pharmacist. The caffeine was pharmaceutical grade. Doses and placebo pills were diluted with calcium carbonate. All dosages were contained within a gelatin caplet, and all were of equal size and appearance.

Performance Measurement

Throughout the cognitive vigilance task's 40-minute trial session three dependent variables were recorded: the number of correct responses, the number of false alarms, and the detection latency (response time). A response was classified as a correct response if the space bar was pressed within 1500 ms of the presentation of the third digit in a sequence. The detection latency was the time from the presentation of the third digit in a sequence until the space bar was pressed, the maximum latency being 1500 ms. The number of false alarms was determined as any response made outside of the 1500 ms window. The correct response (detection rate) score was utilized in the statistical

analysis; however the false alarm and detection latency measures were also evaluated to permit a thorough review of performance. These measures were time coded to allow for statistical analysis of performance between the first, second, third, and fourth ten-minute segments of the forty minute session. Through experimentation this computer program was found to be reliable, and precise in collecting data.

Design

A 3 x 4 mixed design ANOVA was performed with the between subject factor being the caffeine dosage level (0, 20, or 200 mg), and the within subject factor being the time block (first, second, third, or fourth ten minute period of the forty minute task). The detection rate was used as the dependent variable. If significance was found, subsequent tests using a modified Bonferroni Type I error correction were utilized to determine the significance of group mean differences. An alpha level of .05 was utilized.

A 3 x 2 mixed design ANOVA was performed with the between subject factor being the caffeine dosage level (0, 20, or 200 mg) and the within subject factor being the Stanford Sleepiness Scale scores (before and after the Bakan task). This reduction would be evident in the placebo condition as a significantly higher mean sleepiness score after the task, compared to before the task. This analysis also allowed a comparison of alertness levels between the caffeine conditions and the placebo condition. To determine this measure the 'after' sleepiness scale score in the caffeine conditions was compared to the 'after' sleepiness score in the placebo condition. An alpha level of .05 was utilized.

A one way ANOVA was performed between the treatment conditions of caffeine dosage level (0, 20, or 200 mg) and scores on the NASA-TLX rating scale. This analysis was used to determine if participants in the placebo condition experienced a significantly

greater task difficulty compared to the caffeine conditions. An alpha level of .05 was utilized.

Participants were randomly assigned to their caffeine treatment condition. Participant identification number (PIN) badges were placed in a bucket and upon entering the testing room participants drew their number. The first number of their PIN badge represented their treatment condition. The second and third number of their PIN represented their participant number within that specific level.

In an effort to control confounding variables, the number of individuals participating in the study simultaneously was, whenever possible, a function of three. This allowed for one participant from each caffeine condition to be subject to the exact same conditions: such as the experimenter's briefing, time-of-day, temperature, lighting, and unpredictable distractions.

Procedure

Upon arrival at the testing laboratory, participants were thanked for their participation. They were immediately asked to take a seat at a computer station. At each station was a handout containing the consent form, Stanford Sleepiness Scale, Health Survey, second Stanford Sleepiness Scale, and the NASA-TLX rating scale.

They were then asked to read the consent form and experimental briefing (Appendix B). They were encouraged to ask questions. If participants agreed to continue, they were asked to sign the consent form. A copy was offered for their personal records.

Participants then completed the health survey (Appendix C). This survey restates requirements that participants were instructed to follow before the test day. If any

violations occurred, this survey provided a means to detail any deviations from instructions. If participants did not comply with instructions, they were asked to return at a later time. Next, participants completed the Stanford Sleepiness Scale (Appendix D). This scale allowed for a rapid subjective measure of alertness. If participants scored themselves as a four or below, they were asked to return to another session.

The researcher then asked for the participant's name and student identification number, which was entered into a database. A participant identification number (PIN) was randomly assigned to each participant to be utilized throughout the remainder of the study. The identification number consisted of three digits, the first representing one of three treatment levels, the second and third digits were a number from 1 to 20 representing the number of the participant in that treatment level. The number was written on a 3 X 5 inch piece of adhesive paper, which the participants were asked to wear for the remainder of their involvement.

Participants then commenced a vision test. Participants stood on an indicated spot and were asked to read the bottom line of a wall-mounted eye chart. If successful, this determined that participants possessed 20/20 vision.

Next, the participants ingested a dose of caffeine or a placebo. The gelatin caplet was taken with a small cup of water. A timer was started. Ten minutes after caffeine ingestion participants began the cognitive vigilance task. Until that time a thorough explanation of the Bakan vigilance task was given and participants were encouraged to ask questions and discuss concerns. The practice session was also completed in this period.

Participants were then asked to remove their wristwatch to limit distractions, and place it in their pocket. They were informed to insert their ear plugs when their computer display instructed them to do so. Participants were then asked to adjust their chair to a proper height. The researcher then instructed the participants to follow the directions on their computer monitors. The monitors guided the participants through the trial session, and informed the participants when their forty-minute session was finished. While remaining seated, participants completed the second Stanford Sleepiness Scale and the NASA-TLX task difficulty rating.

Once all participants had finished their task and completed the required forms, the researcher instructed all participants to remove their ear plugs. They were then told to remove their PIN labels, and affix their labels on the front page of their handouts. Once their handouts were returned to the researcher, participants were given envelopes with complimentary cinema tickets inside. Participants were then free to leave; however, they were encouraged to remain and discuss any remaining questions with the researcher.

The computer program automatically saved the participant's performance data and reset to the introductory screen, ready for the next participant.

RESULTS

Descriptive Statistics

A total of 73 individuals participated in the experiment. Out of 73 participants, 4 data sets were removed from the analysis due to individual characteristics which affected performance.

Participant number 204's unusual behavior was noticed at the conclusion of the task when a series of auditory beeps were heard by the researcher. This participant had fallen asleep with their hand on the spacebar. When the computer task ended and reset to the introductory computer screen the space bar is used to cycle through screens and restarts the task. However when the program asked for his PIN number holding down the spacebar caused an entry with many more digits than the standard 3 digit PIN. This caused the series of beeps. The participant was awoken and was extremely disoriented. This participant's data set has been removed from the analysis.

The second data set removed was Participant number 306. Upon completion of the task the program resets to the introductory screen. This participant continued through the introductory computer screens intended for the next participant and reentered their PIN number. This action caused his data file to write over itself and upon analysis their data file came up blank.

Participant number 219's data was also removed. This participant's unusual behavior was noticed when they continuously turned their head to look back at the

researcher. It was originally thought that this individual might have a question, but then they would turn their head again and look out the window, or at the person sitting at the next computer. This behavior became increasingly obvious towards the end of the task. There were periods when this participant would put their hands behind their head and stare at the ceiling for a period of a few minutes.

The final data set that was removed from the analysis was that of Participant number 222. English was not this student's first language. When the directions were read to the class, this individual engaged the participant at the next station in conversation. That participant alerted the researcher that participant number 222 was unsure of what the task was asking them to do. When the researcher noticed this, the task was explained in lengthy detail although the participant remained confused. Eventually the participant said that they understood, although they seemed quite unsure and embarrassed. Since nine other people were waiting to start the task (and the ten minute waiting period had almost elapsed) the trial session began. This participant sat directly in front of the researcher's station. Throughout their session the researcher observed this individual to determine if any correct responses were being signaled. It became obvious that this individual was not making any correct responses. Afterwards, this individual explained that since English is his second language he did not clearly understand what 'odd' and 'even' numbers represented.

Data Analysis

Overall tables of group means and standard deviations are listed in Appendix F. Significant subsequent comparisons are listed in the body of the text, however tables of all preplanned comparisons are listed in Appendix G. A Bonferroni Type I error

correction has been implemented (Keppel, 1991). Refer to the text of each analysis for information regarding that specific pairwise alpha. It is also important to note that the stated confidence intervals contain no Type I error correction.

The Performance Effects of Caffeine on a Cognitive Vigilance Task

A 3 X 4 (Caffeine dosage condition X Time block) repeated measure design was implemented to assess performance on the Bakan cognitive vigilance task. The dependent measure for this task was the number of signals correctly indicated by the participant, called the detection rate or correct response rate. The additional dependent measures of false alarms and response times were also analyzed. The repeated measures statistical analysis function of SPSS® was used for the comparison of performance.

Mauchly's test of sphericity indicated that non-sphericity was present in the detection rate results. The Greenhouse-Geisser correction was utilized in the analysis to correct for non-sphericity (Keppel, 1991).

The ANOVA source table for the analysis pertaining to the detection rate performance analysis is presented in Table 2. Prediction one stated that the performance of all treatment levels would reduce as a function of time. The analysis did not support this hypothesis as no effect was observed within the Time factor, $F_{(3, 198)} = 1.207, p = .321$.

Prediction two stated that the 20 and 200 mg caffeine doses would reduce the vigilance decrement, thereby increasing performance over time compared to the placebo condition. This prediction was partially supported. Significant interaction effects were found for the Caffeine by Time interaction, $F_{(5, 191, 171, 306)} = 2.294, p = .045$. A between-

subjects significant effect was also found between the Caffeine conditions, $F_{(2,66)} = 3.286$, $p < .044$. Refer to Figure 1 for a graphical representation of these findings.

A Bonferroni correction was implemented to reduce Type I error inflation (Keppel, 1991). Pairwise alpha was determined to be .017 for the vigilance task data. This value was determined by dividing the family wise alpha of .05 by the number of comparisons performed to produce a pairwise alpha value of .017.

Results for the subsequent comparisons using the modified Bonferroni Type I error correction are presented in Table 3. As can be seen, the 200 mg Caffeine condition demonstrated significantly higher detection rates in the final ten minutes of the session compared to the 20 mg condition. The 20 mg condition was not significantly different from the placebo condition.

Table 2. Source Table for the Detection Rate Analysis of Variance

Source	Sphericity	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>p</i>	Power	η^2
Caffeine		2	1104.2	552.1	3.29	.044	.61	.091
Error		66	11088. 7	168.0				
Time	SA	3	40.0	13.3	1.21	.308	.32	.018
	GG	2.60	40.0	15.4	1.21	.307	.30	.018
Time * Caffeine	SA	6	151.9	25.3	2.29	.037	.79	.065
	GG	5.19	151.9	29.3	2.29	.045	.74	.065
Error	SA	198	2184.8	11.0				
	GG	171.31	2184.8	12.8				

SA = Sphericity Assumed
GG = Greenhouse-Geisser

Table 3. Detection Rate Subsequent Tests

Time Block	Caffeine	Caffeine	Mean Difference	Standard Error	<i>p</i>	95% Confidence Interval	
						Lower Bound	Upper Bound
4	0	20	2.402	2.207	.280	-2.004	6.808
	0	200	-4.897	2.113	.024	-9.115	-.680
	20	200	-7.299	2.164	.001	-11.620	-2.978

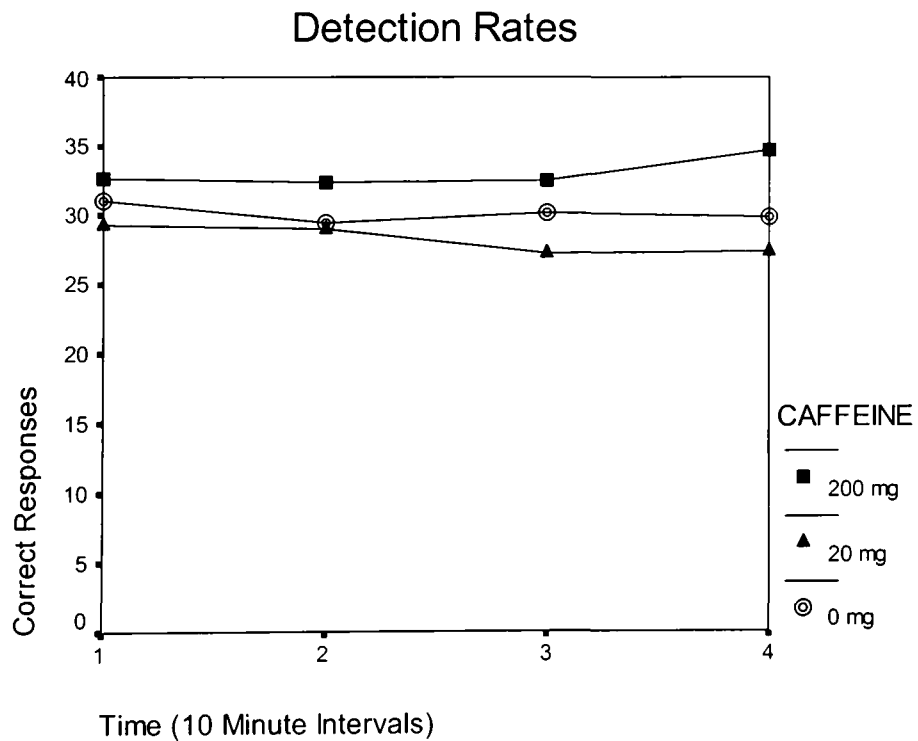


Figure 1. The Main Effects of Caffeine on the Bakan Vigilance Task

To permit a thorough perspective of vigilance task performance the number of false alarms (space bar presses outside of correct response period) and response times (time from the presentation of the third digit in a sequence until the space bar response was made) were collected by the computer program, and analyzed.

The false alarm results were also non-spherical and the Greenhouse-Geisser correction was implemented. The ANOVA source table for the following analyses is presented in Table 4. Again, prediction one was not supported as significance was not found within the Time condition, $F_{(2\ 426, 160\ 101)} = .879, p = .453$.

Prediction two was also not supported. Significant effects were found for the Caffeine by Time interaction, $F_{(4\ 806, 157\ 954)} = 2.427, p = .038$ (refer to Table 3). However, subsequent tests determined that in the final ten minutes of the task, the 200 mg caffeine condition had significantly more false alarms than in the 20 mg condition (Table 5). Refer to Figure 2 for a graphical representation.

Table 4. Results of Analysis of Variance for False Alarms

Source	Sphericity	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>p</i>	Power	η^2
Caffeine		2	363.9	182.0	.98	.380	.21	.029
Error		66	12231.0	185.3				
Time	SA	3	90.3	30.1	.88	.453	.24	.013
	GG	2.43	90.3	37.2	.88	.435	.22	.013
Time * Caffeine	SA	6	502.9	83.8	2.45	.026	.82	.069
	GG	4.85	502.9	103.7	2.45	.038	.75	.069
Error	SA	198	6776.8	34.2				
	GG	160.10	6776.8	42.3				

SA = Sphericity Assumed

GG = Greenhouse-Geisser correction

Table 5. False Alarm Subsequent Tests

Time Block	Caffeine	Caffeine	Mean Difference	Standard Error	<i>p</i>	95% Confidence Interval	
						Lower Bound	Upper Bound
4	0	20	2.455	2.719	.370	-2.972	7.883
	0	200	-4.743	2.602	.073	-9.939	.453
	20	200	-7.198	2.666	.009	-12.521	-1.875

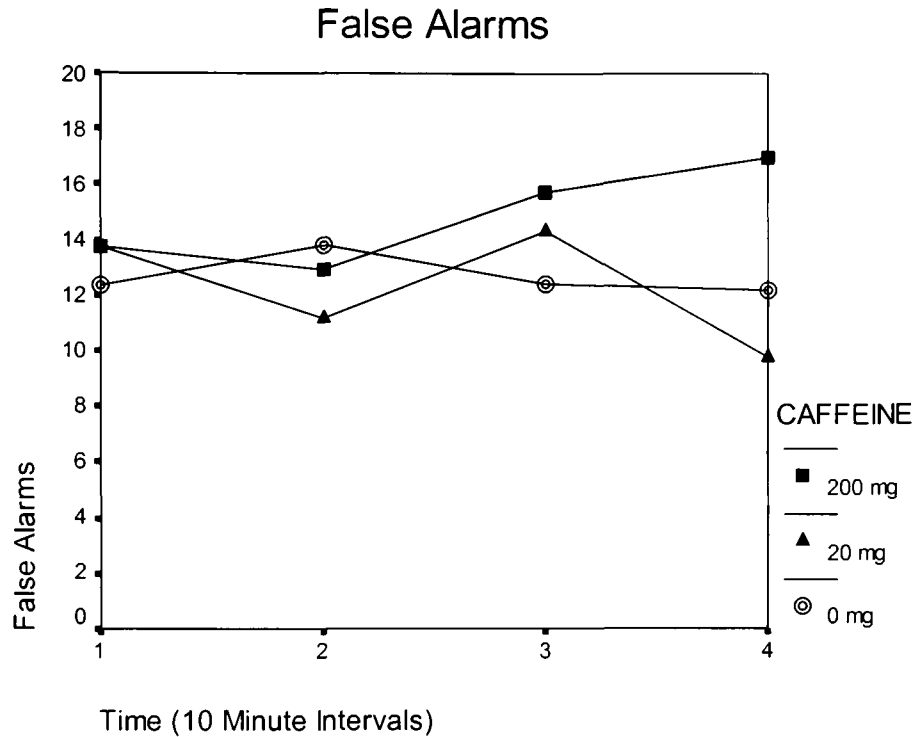


Figure 2. The Effects of Time and Caffeine on False Alarm Rates

A final analysis was performed using response times between Caffeine conditions. These data were determined to have sphericity. The ANOVA source table for this analysis is presented in Table 6.

Prediction one was not supported in this data, as significant results for Time were not found, $F_{(3, 198)} = 1.555, p = .202$. Prediction two was partially supported. The analysis determined that significant effects were present within the Time by Caffeine interaction, $F_{(6, 195)} = 2.212, p = .034$. Subsequent tests determined that the 20 mg condition displayed significantly faster response times than the placebo condition. However, the 20 mg group also displayed significantly faster reaction times than the 200

mg condition (see Table 7). A graphical representation of this data is presented in Figure 3.

Table 6. Results of Analysis of Variance for Response Times

Source	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>p</i>	Power	η^2
Caffeine	2	26352.5	13176.2	1.32	.275	.28	.038
Error	66	659992.0	9999.9				
Time	3	30294.0	3370.0	1.56	.202	.41	.023
Time * Caffeine	6	30294.0	5049.0	2.33	.034	.80	.066
Error	198	429157.8	2167.5				

Table 7. Response Times Subsequent Tests

Time Block	Caffeine	Caffeine	Mean Difference	Standard Error	<i>p</i>	95% Confidence Interval	
						Lower Bound	Upper Bound
4	0	20	48.752	18.936	.012	10.944	86.559
	0	200	-3.071	18.127	.866	-39.263	33.121
	20	200	-51.823	18.571	.007	-88.902	-14.744

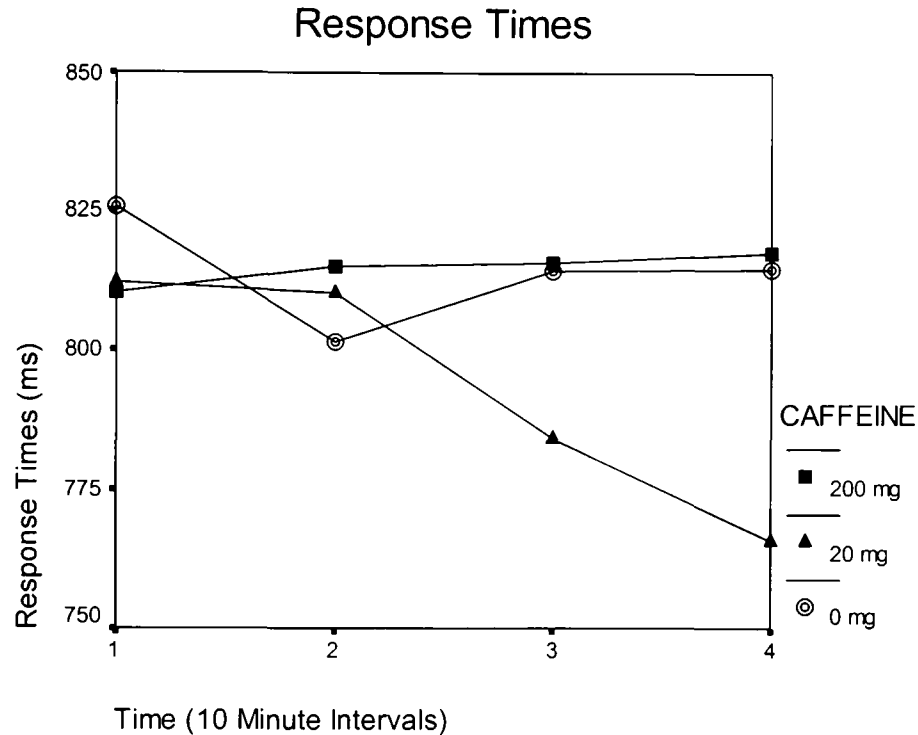


Figure 3. The Effects of Time and Caffeine on Response times

Pre- and Post-Task Stanford Sleepiness Scale Scores

A 3 X 2 (Caffeine dosage condition by Pre- or Post-test) mixed design was implemented to assess differences in alertness before and after the vigilance task. The dependent measures were scores on Stanford Sleepiness Scales, before and after the vigilance task. The ANOVA source table for this analysis is presented in Table 8.

A Bonferroni correction was implemented to reduce type one error inflation. Pair wise alpha was determined to be .017. This value was determined by dividing the overall alpha value of .05 by the total number of comparisons made (3).

Prediction three stated that sleepiness levels would be significantly higher after the vigilance task in all conditions, when compared against pre-task scores. This

prediction was supported as significant effects were found between the Stanford Sleepiness Scales administered before and after the task, $F_{(1,66)} = 55.284, p < .001$. Subsequent tests determined that all conditions were significantly sleepier after the task, $p < .001$, refer to Table 9.

Prediction four stated that post-task Stanford Sleepiness Scale scores within the caffeine conditions would be significantly less sleepy than those in the placebo condition. This prediction was not supported as no significant interaction effects were found, $F_{(2,66)} = .430, p = .652$. Neither were any significant differences found between the Caffeine conditions, $F_{(2,66)} = .636, p = .533$, refer to Table 5. Refer to Figure 4 for a graphical representation of these results.

Table 8. Results of Analysis of Variance for Pre- and Post-Task Stanford Sleepiness Scale Scores

Source	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>p</i>	Power	η^2
Scores	1	60.8	60.8	55.28	.000	1.00	.456
Error	66	72.5	1.1				
Scores* Caffeine	2	.95	.5	.43	.652	.12	.013
Error	66	72.6	1.1				
Caffeine	2	1097.3	1.0	.64	.533	.15	.019
Error	66	105.0	1.6				

Table 9. Stanford Sleepiness Scale Subsequent Tests

Caffeine	SSS	SSS	Mean Difference	Standard Error	<i>p</i>	95% Confidence Interval	
						Lower Bound	Upper Bound
200 mg	1	2	-1.120	.297	.000	-1.712	.528
20 mg	1	2	-1.524	.324	.000	-2.170	.878
0 mg	1	2	-1.348	.309	.000	-1.965	-.731

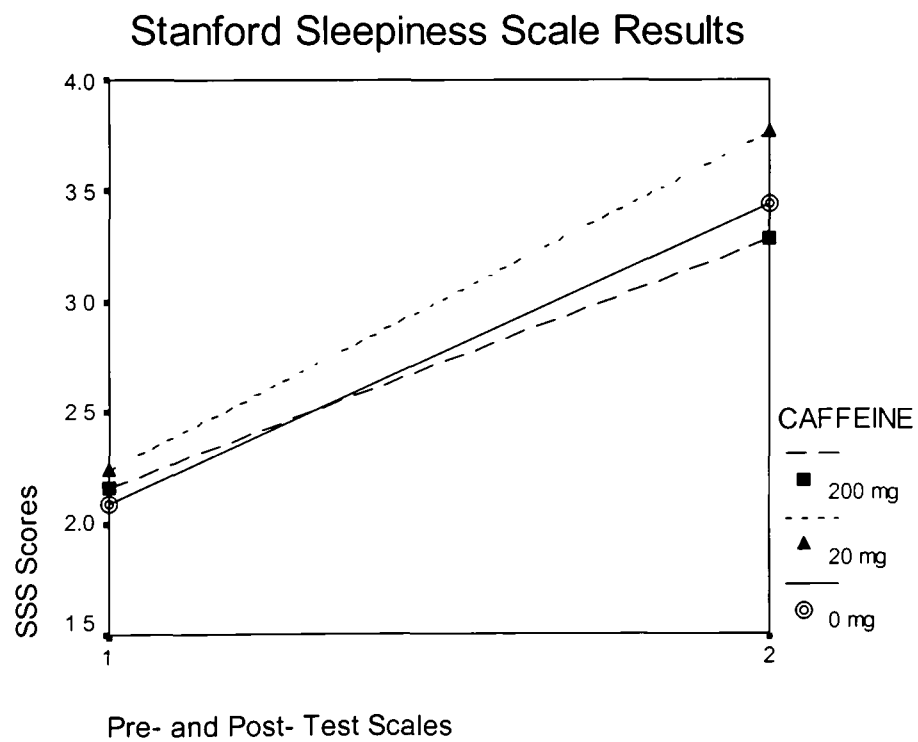


Figure 4. The Effects of a Vigilance Task on Stanford Sleepiness Scale Scores

NASA-TLX Scores

A 3 X 1 (Caffeine dosage X NASA-TLX Score) between-subjects design was implemented to assess differences in task difficulty between dosage conditions. The dependent measures were scores on NASA-TLX rating scales. The univariate statistical analysis function of SPSS® was used for the comparison of performance. A Bonferroni correction was used for the subsequent comparisons. For this analysis .05 was divided by 2 comparisons to create a family wise alpha value of .025.

Prediction five stated that perceived task difficulty would be lower in the caffeine conditions compared to the placebo condition. This prediction was partially supported. A significant main effect for Caffeine was found on the task load index scales, $F_{(2,66)} = 6.065, p = .004$ (see Table 10). Workload scores in the 200 mg condition were significantly lower than workload scores in the placebo condition and 20 mg groups, indicating that the participants in the 200 mg group perceived the task as easier than participants in the other groups. However, no differences in task difficulty scores were found between the 20 mg and placebo conditions (see Table 11).

Table 10. Results of Analysis of Variance for NASA-TLX scores

Source	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>p</i>	Power	η^2
Caffeine	2	652.9	326.4	6.07	.004	.87	.155
Error	66	3552.6	53.8				

Table 11. NASA-TLX Subsequent Tests

		Mean Difference	Standard Error	<i>p</i>	95% Confidence Interval	
Caffeine	Caffeine				Lower Bound	Upper Bound
0 mg	20 mg	.4762	2.21439	.830	-4.8974	3.9450
	200 mg	6.1600	2.11977	.005	1.9278	10.3922

DISCUSSION

Previous research has demonstrated that low doses of caffeine produce enhanced performance similar to high doses (Smit & Rogers, 2000), and that the Bakan cognitive vigilance task is extremely reliable in producing vigilance decrements (Harkins et al., 1974). This study sought to combine these findings to determine if low doses of caffeine will improve performance on a cognitive vigilance task. This research also explored the arousal and compensatory theories by implementing Stanford Sleepiness Scales and the NASA-TLX respectively. The results of this study supported past results, with a few exceptions.

Prediction One

This study failed to produce a significant vigilance decrement. All performance curves (detection rate, false alarms, and response times) in the placebo condition were relatively flat, showing no time related decrement. However it is important to note that this decrement is unlikely to have been caused by all participants mastering the task, as several factors suggest otherwise. For example, on average participants missed half of all sequences presented. Also, in post-task discussion participants often commented that the task was very cognitively demanding and frustrating. Task difficulty was confirmed as being high by the results of the NASA-TLX. Why then did the task fail to produce a vigilance decrement? There are several possibilities.

In fact, it is very likely that a vigilance decrement did occur. However, it was not evident in the results because the decrement was balanced by learning effects. It seems illogical to assume that participant performance would neither improve as a result of learning nor decrease as a result of a vigilance decrement throughout a forty minute session. Especially considering that arousal levels were significantly lower after the task. Instead it is reasonable to assume that a learning effect did take place, causing an increase in performance. However, this increase was counterbalanced by the performance degrading effects of a vigilance decrement, thereby causing a plateau in performance. Based on this reasoning, any factors that facilitated learning negatively impacted the study's ability to produce a vigilance decrement.

The vigilance task that was utilized was forty minutes in duration, segregated into ten minute blocks. The number and identity of digits in each ten minute block were identical. In post-task discussion several participants indicated that near the conclusion of the task they began to recognize the sequences. These statements support the expectancy theory developed by Baker (1959). The expectancy theory explains vigilance task performance through the observer estimating the probability of signal presentation, based on past signals. Therefore, detection increases if signals are presented to the participant at regular intervals that allow accurate prediction. Although it was hypothesized that using four identical number sets within each ten minute trial would not allow for prediction, post-task discussion and results suggest otherwise. In fact, this factor may have significantly impacted task learning.

Another factor that may have improved task learning deals with Smith's (1966) motivation theory. This theory describes the vigilance decrement as the result of a lack

of individual motivation. Smith argues that all individuals are capable of maintaining vigilance for a few hours with no mistakes, however they are not motivated to do so in professional or laboratory environments due to the monotony of vigilance tasks and the lack of intrinsic motivation. When conducting the research participants sat side-by-side and the researchers sat at the back of the room in full view of all computer stations. It is possible that participants were motivated to maintain a high level of performance due to peripherally observing their neighbor signaling responses. In post-task discussion, one participant mentioned that they were silently keeping track of their performance and comparing to the responses signaled by their neighbor. Participants may have also felt pressure to perform to a high level due to the researcher continually observing their performance. It would be interesting to conduct this research again with participants in individual rooms, and observe any variability that occurs as a result of being alone.

On separate track, it is also possible that the vigilance decrement was not demonstrated because the duration of the task was insufficient. Mackworth (1950) concluded that the majority of the vigilance decrement occurs between the first and second thirty minute segments of a task. Even though this task was 40 minutes long, it is possible that a significant decrement may have been clearly demonstrated if the task was extended for an additional 20 minutes.

In any case, although a significant vigilance decrement was not produced, the task was a successful means to demonstrate the performance enhancing qualities of caffeine over time.

Prediction Two

The performance effects of caffeine truly are diverse and complex. When viewing the performance curves presented in Figure 1 it is tempting to conclude that the 20 mg group had no performance effects as significant differences do not exist between it and the placebo condition. It is also tempting to attribute the lower mean scores of the 20 mg group to individual variability or chance alone. However, when the false alarm and response time curves are taken into consideration it becomes obvious that these conclusions would have been made in haste.

When a decrease in detection rate is demonstrated, it may reflect either a decrease in sensitivity to signals, or a shift to a more conservative criterion for responding. A researcher can determine which caused the decrease in detection rate by evaluating the number of false alarms. If the false alarm rate improved along with the detection rate degrading, a more conservative response criterion has been adopted by the participants (Proctor & Van Zandt, 1994). It is evident by comparing Figure 1 to Figures 2 and 3 that although the 20 mg group's detection rate slightly deteriorated over time, their number of false alarms and response times significantly improved. These results indicate that their responding criterion became more conservative. This demonstrates that the 20 mg condition did significantly impact performance, although not on the correct response measure alone which was originally chosen for data analysis. Although the 200 mg condition demonstrated the highest detection rate in the final ten minutes of the task, their false alarm rate and response times were significantly higher.

These findings suggest that both the low and high dose of caffeine (20 and 200 mg) significantly impact performance, although in different ways. The low dose of

caffeine produces a more conservative response criterion, which creates a greater efficiency with a lower detection rate overall. The high dose of caffeine produces a liberal response criterion, which produces a higher overall detection rate at the cost of a slower response time and an increased number of false alarms.

Evaluating the ERP effects of caffeine Lorist, Snel, and Kok (1994) demonstrated that after caffeine treatment, an increasingly negative going N1 in combination with a shorter latency was produced. This result suggests that caffeine increases receptivity to external stimuli and accelerates information processing. Caffeine also produced a more positive going P3. This increase in amplitude at the posterior electrode site represents an increase in phasic cortical arousal which affects perceived task difficulty.

The N1 findings of this research may further explain the distinction between low and high caffeine dose results. It is possible that the high dose of caffeine significantly increased receptivity to a point where participants were on the verge of being over stimulated. This effect was demonstrated through a detection rate increase, at the expense of an increased number of false alarms. External stimuli receptivity was also likely to have increased in the low caffeine condition, however to a lesser extent which instead facilitated more accurate predictions and faster response times.

It is also important to note that even though this data suggests significant effects of caffeine, it is impossible to be certain that maximum absorption took place within the tested time period without a physiological measure. Therefore implementing a saliva or plasma method of measuring caffeine absorption levels would reduce the Type 1 error potential.

Prediction Three

Using the Stanford Sleepiness Scale to measure alertness this study determined that arousal decreased throughout the vigilance task. This finding supports the arousal theory which states that the monotonous nature of vigilance tasks causes the alertness level of the central nervous system to diminish. Thereby causing a decrease in responsiveness and efficiency, resulting in a performance decrement (Davies & Parasuraman, 1982).

Prediction Four

This analysis also hypothesized that significantly greater alertness levels would be found post-task in the caffeine conditions compared to the placebo condition, as caffeine is a central nervous system stimulant which counteracts the decrease described by the arousal theory. However, no significant differences in alertness levels were found between caffeine conditions.

A possible explanation for this is that since the Stanford Sleepiness Scale is a subjective measure it was not sensitive enough to detect differences in alertness levels. An objective alertness measure, such as heart rate or blood pressure monitoring, would potentially gather alertness data with greater precision. This may allow for detection of differences in arousal levels between caffeine and placebo conditions, if they exist.

Prediction Five

The NASA-TLX produces a subjective measure of task load. This measure was implemented after the vigilance task, and results were analyzed between caffeine treatment conditions.

This analysis is based on the compensatory system which suggests that low arousal, in this case caused by the vigilance task, is combated by increased subjective effort (Broadbent, 1971). It has been demonstrated that the arousal enhancing effects of caffeine reduced the subjective effort required to complete a task, compared to the placebo condition (Linde, 1995). This is supported by ERP findings where increased alertness and vigor reported by participants in caffeine conditions, combined with increased P3 amplitudes, suggested that actual task complexity is perceived as being lower (Lorist, Snel, & Kok, 1994).

The results of the analysis indicated that subjective difficulty was significantly lower in the high caffeine condition. However, the low caffeine and placebo conditions were not significantly different. This finding supports the compensatory system.

Perhaps this measure also explains why signal detection criterion becomes increasing liberal in the 200 mg condition. Since individuals under a high dose of caffeine perceive the vigilance task as less difficult it is logical that they will be less concerned with the potential of signaling false positives.

On the other hand, participants in the low caffeine dose condition perceive the same task difficulty as those in the placebo condition. However, their significantly better results on response times and false alarms indicate that the low dose of caffeine is improving their receptivity.

CONCLUSIONS & RECOMMENDATIONS

The real question lies in the overall meaning of the data summarized in this study. What are the practical applications of this research for individuals whose professions incorporate vigilance tasks? The broad conclusion is that the impact of caffeine on performance is diverse and complex.

When choosing to consume caffeine one must first consider the task they plan on performing. If the task places a high importance on accuracy and millisecond response times, such as the task of air traffic controllers monitoring radar displays, a low dose of caffeine has the potential to significantly improve performance. However, if the task places utmost importance on the overall number of correct responses regardless of false alarms or millisecond response times, such as the task of a mall security guard monitoring closed circuit video displays, then a high dose of caffeine has potential to significantly improve on the job performance.

This recommendation implies the question of precisely what dose of caffeine is considered high or low. The present study used doses of 200 mg for the high condition and 20 mg for the low condition. Future research is required to build a dose-response curve detailing the precise caffeine dosage at which the shift from a conservative to a liberal response criterion occurs. It is recommended that future researchers avoid using identical number sequences throughout each ten minute period, thus avoiding any effects associated with the expectancy theory. Additionally, it is recommended that future

studies extend the duration of their vigilance tasks to 60 minutes, place all participants in solitary locations when conducting the task, and evaluate plasma caffeine levels to be certain of absorption rates.

Looking at the pre- and post-test scores on the Stanford Sleepiness Scales it is apparent that a drop in arousal occurs. The use of this scale did not distinguish significant arousal differences between caffeine and placebo conditions. However, further research using a precise measurement such as heart rate or blood pressure monitoring is required to produce an arousal curve demonstrating precisely when arousal begins to decline and the rate of this decline, and differences between placebo and caffeine conditions.

The results of the NASA-TLX provide an interesting insight into perceived vigilance task difficulty under the influence of caffeine. Task difficulty was found to be significantly less in the high caffeine dose condition. Additional research is required to determine if this finding is applicable to other tasks.

REFERENCES

- Arnaud, M.J. (1993). Metabolism of caffeine and other components of coffee. In: S. Garattini (Ed.), *Caffeine, Coffee and Health*. New York: Raven Press, Ltd., 43-95.
- Arnaud, M.J. (1998). Pharmacokinetics and metabolism of caffeine. In: Snel, J., and Lorist, M. (Eds.), *Nicotine, Caffeine and Social Drinking Behavior and Brain Function*. Amsterdam: Harwood Academic Publishers, 153-165.
- Bakan, P. (1959). Extroversion-introversion and improvement in an auditory vigilance task. *British Journal of Psychology*, 50, 325-332.
- Baker, C.H. (1959). Towards a theory of vigilance. *Canadian Journal of Psychology*, 13, 35-42.
- Balogh, A., Irmisch, E., Klinger, G., Splinter, F.K., and Hoffman, A. (1987). Elimination of caffeine and metamxole in the menstrual cycle of the fertile female. *Zentralblatt Fur Gynakologie*, 109 (18), 1135-1142.
- Blanchard, J., and Sawers, S.J. (1983). The absolute bioavailability of caffeine in man. *European Journal of Clinical Pharmacology*, 24 (1), 93-98.
- Bonati, M., Latini, R., Galletti, F., Young, J.F., Tognoni, G., and Garattini, S. (1982). Caffeine disposition after oral doses. *Clinical Pharmacology and Therapeutics*, 32, 98-106.

- Boulenger, J.P., Uhde, T., Wolff, E.A., and Post, R.M. (1984). Increased sensitivity to caffeine in patients with panic disorders. *Archives of General Psychiatry*, 41: 1067-1071.
- Broadbent, D.E. (1957). A mechanical model for human attention and immediate memory. *Psychological Review*, 64, 205-215.
- Broadbent, D.E. (1971). *Decision and stress*. London: Academic Press.
- Campbell, M.E., Spielberg, S.P., and Kalow, W. (1987). A urinary metabolic ratio that reflects systemic caffeine clearance. *Clinical Pharmacology and Therapeutics*, 42, 157-165.
- Caraco, Y., Zylber-Katz, E., Berry, E.M., and Levy, M. (1995). Caffeine pharmacokinetics in obesity and following significant weight reduction. *International Journal of Obesity*, 19, 234-239.
- Carskadon, M.A. (1993). Evaluation of excessive daytime sleepiness. *Neurophysiologie Clinique*, 23, 91-100.
- Carskadon, M.A., and Dement, W.C. (1979). Effects of total sleep loss on sleep tendency. *Perceptual Motor Skills*, 18, 495-506.
- Chervin, R.D., Aldrich, M.S., Pickett, R., and Guilleminault, C. (1997). Comparison of the results of the Epworth Sleepiness Scale and the Multiple Sleep Latency Test. *Journal of Psychosomatic Research*, 42, 145-155.
- Daan, S., Beersma, D.G.M, and Borbely, A.A. (1984). Timing of human sleep: Recovery process gated by a circadian pacemaker. *American Journal of Physiology*, 246, 161-178.

- Danker-Hopfe, H., Kraemer, S., Dorn, H., Schmidt, A., Ehlert, I., and Herrmann, W.M. (2001). Time-of-day variations in different measures of sleepiness (MSLT, pupillography, and SSS) and their interrelations. *Psychophysiology*, 38, 828-835.
- Davies, D.R. and Parasuraman, R. (1982). *The psychology of vigilance*. London: Academic Press, Inc.
- Durlach, P.J. (1998). The effects of a low dose of caffeine on cognitive performance. *Psychopharmacology*, 1, 116-119.
- Frewer, L.J., and Lader, M. (1991). The effects of caffeine on two computerized tests of attention and vigilance. *Human Psychopharmacology*, 6, 119-128.
- Grant, D.M., Tang, B.K., and Kalow, W. (1983). Variability in caffeine metabolism. *Clinical Pharmacology and Therapeutics*, 33: 591-602.
- Harkins, S.W., Nowlin, J.B., Ramm, D., and Schroeder, S. (1974). Effects of age, sex and time-on-watch on a brief continuous performance task. In E. Palmore (Ed.) *Normal Aging*. Vol.2. Durham, North Carolina: Duke University Press.
- Hasenfratz, M., and Battig, K. (1994). Acute dose-effect relationships of caffeine and mental performance, EEG, cardiovascular and subjective parameters. *Psychopharmacologia*, 114 (2), 281-287.
- Hindmarch, I., Rigney, U., Stanley, N., Quinlan, P., Rycroft, J., and Lane, J. (2000). A naturalistic investigation of the effects of day-long consumption of tea, coffee and water on alertness, sleep onset and sleep quality. *Psychopharmacology*, 149 (3), 203-216.

- Hirsh, K. (1984). Central nervous system pharmacology of the dietary methylxanthines. In: Spiller, G.A. (Ed.), *The Methylxanthine Beverages and Foods: Chemistry, Consumption, and Health Effects*. New York: Allan R. Liss, Inc.
- Hoddes, E., Dement, W., and Zarcone, V. (1972). The history and use of the Stanford Sleepiness Scale. *Psychophysiology*, 9:150.
- Hollingworth, H.L. (1912). The influence of caffeine on mental and motor efficiency. *Archives of Psychology*, 22 (1), 1-166.
- Institute of Food Technologists' Expert Panel on Food Safety and Nutrition (1987). *Evaluation of Caffeine Safety*. Chicago: Institute of Food Technologists.
- Jacobson, B.H., and Bouher, B.J. (1991). Caffeine consumption by selected demographic variables. *Health Values*, 15, 49-55.
- Jerison, H.J., and Pickett, R.M. (1964). Vigilance: The importance of the elicited observing rate. *Science*, 143, 970-971.
- Johnson, L.C., Freeman, C.R., Spinweber, C.L., and Gomez, S.A. (1991). Subjective and objective measures of sleepiness: Effect of benzodiazepine and caffeine on their relationship. *Psychophysiology*, 28 (1), 65-71.
- Kamimori, G.H., Penetar, D.M., Headley, D.B., Thorne, D.R., Otterstetter, R., and Belenky, G. (2000). Effect of three caffeine doses on plasma catecholamines and alertness during prolonged wakefulness. *European Journal of Clinical Pharmacology*, 56, 537-544.
- Kamimori, G.H., Somani, S.M., Knowlton, R.G. and Perkins, R.M. (1987). The effects of obesity and exercise on the pharmacokinetics of caffeine in lean and obese volunteers. *European Journal of Clinical Pharmacology*, 31, 595-600.

- Keppel, G. (1991). *Design and Analysis: A researcher's handbook*. New Jersey: Prentice Hall, Inc.
- Klein, K.E., Herrmann, R., Kuklinski, P., and Wegmann, H-M. (1977). Circadian performance rhythms: Experimental studies in air operations. In: Mackie, R.R. (Ed.), *Vigilance: Theory, operational performance and physiological correlates*. New York: Plenum Press.
- Kotake, A.N., Schoeller, D.A., Lambert, G.H., Baker, A.L., Schaffer, D.D., and Josephs, H. (1982). The caffeine CO₂ breath test: Dose response and route of N-demethylation in smokers and nonsmokers. *Clinical Pharmacology and Therapeutics*, 32, 261-269.
- Levy, M., and Zylber-Katz, E. (1983). Caffeine metabolism and coffee-attributed sleep disturbances. *Clinical Pharmacology and Therapeutics*, 33 (6), 770-775.
- Lichstein, K.L., Wilson, N.M., Noe, S.L., Aguillard, R.N., and Bellur, S.N. (1994). Daytime sleepiness in insomnia: Behavioral, biological, and subjective indices. *Sleep*, 17, 693-702.
- Lieberman, H.R. (1992). Caffeine. In: A.P. Smith and D.M. Jones (Eds.), *Handbook of Human Performance*. Vol. 2. Health and Performance. London: Academic Press, 49-72.
- Lieberman, H.R., Wurtman, R.J., Emde, G.G., Roberts, C., and Coviella, I.L.G. (1987). The effects of low doses of caffeine on human performance and mood. *Psychopharmacology*, 92, 308-312.
- Linde, L. (1995). Mental effects of caffeine in fatigued and non-fatigued female and male subjects. *Ergonomics*, 38 (5), 864-885.

- Lorist, M.M. (1998). Caffeine and information processing in man. In: Snel, J., and Lorist, M. (Eds.), *Nicotine, caffeine and social drinking Behavior and brain function*. Amsterdam: Harwood Academic Publishers, 167-183.
- Lorist, M.M., Snel, J., and Kok, A. (1994). Influence of caffeine on information processing stages in well rested and fatigued subjects. *Psychopharmacology*, 113, 411-421.
- Mackie, R.R. (1987). Vigilance research – Are we ready for countermeasures? *Human Factors*, 29 (6), 707-723.
- Mackworth, N.H. (1950). *Research on the measurement of human performance* (Medical Research Council Special Report No. 268). London: H.M. Stationary Office.
- Manly, T., Lewis, G.H., Robertson, I.H., Watson, P.C., and Datta, A.K. (2002). Coffee in the cornflakes: Time-of-day as a modulator of executive response control. *Neuropsychologia*, 40, 1-6.
- May, D.C., Jarboe, C.H., Van Bakel, A.B., and Williams, W.M. (1982). Effects of cimetidine on caffeine disposition in smokers and nonsmokers. *Clinical Pharmacology and Therapeutics*, 31, 656-661.
- Miller, L.S., Lombardo, T.W., and Fowler, S.C. (1995). Caffeine and time of day effects on a force discrimination task in humans. *Physiological Behavior*, 57 (6), 1117-1125.
- Nehlig, A., Daval, J.L., and Debry, G. (1992). Caffeine and the central nervous system: Mechanisms of action, biochemical, metabolic and psycho stimulant effects. *Brain Research Reviews*, 17, 139-170.

- Phillis, J.W. (1991). *Adenosine and adenine nucleotides as regulators of cellular function*. Boca Raton, Florida: CRC Press.
- Pilcher, J.J., Schoeling, S.E., and Prosansky, C.M. (2000). Self-report sleep habits as predictors of subjective sleepiness. *Behavioral Medicine*, 24 (4), 161-168.
- Proctor, R.W., and Van Zandt, T. (1994). *Human Factors in Simple and Complex Systems*. Massachusetts, Allyn and Bacon.
- Shohet, K.L., and Landrum, R.E. (2001). Caffeine consumption questionnaire: A standardized measure for caffeine consumption in undergraduate students. *Psychological Reports*, 89, 521-526.
- Smit, H.J., and Rogers, P.J. (2000). Effects of low doses of caffeine on cognitive performance, mood and thirst in low and higher caffeine consumers. *Psychopharmacology*, 2, 167-173.
- Smith, A.P. (1998). Effects of caffeine on attention: Low levels of arousal. In: Snel, J., and Lorist, M. (Eds.), *Nicotine, caffeine and social drinking Behavior and brain function*. Amsterdam: Harwood Academic Publishers, 167-183.
- Smith, R.L. (1966). *Monotony and Motivation: A Theory of Vigilance*. Los Angeles, California: Dunlop and Associates, Inc.
- Stelt, O., and Snel, J. (1998). Caffeine and human performance. In: Snel, J., and Lorist, M. (Eds.), *Nicotine, caffeine and social drinking behavior and brain function*. Amsterdam: Harwood Academic Publishers, 167-183.
- Wilkinson, R.T. (1970). Methods for research on sleep deprivation and sleep function. In: Hartmann, E. (Ed.), *Sleep and Dreaming*. Boston: Little Brown.

Yerkes, R., and Dodson, J. (1908). The relationship of strength of stimulus to rapidity of habit formation. *Journal of Comparative Neurology and Psychology*, 18, 459-482.

Appendix A
Recruitment Script

Recruitment Script

A study is currently being conducted by a graduate student in the Human Factors and Systems degree program concerning the effects of a low-dose of caffeine on vigilance performance. Vigilance refers to a person's ability to perform a task over a period of time. An example of a real-world vigilance task would be air traffic controllers monitoring displays.

This investigation is being developed to further explore the effects of caffeine on vigilance, at varying levels of dosage. Eligible participants will be asked to perform a computer test called a cognitive vigilance task for forty minutes while under the influence of caffeine, which will be given by the experimenter before starting the task.

Participants are being recruited across campus that meet the following criterion:

- In good health
- Low-to-moderate daily caffeine intake (200-400 mg) *Refer to table 1*
- On average, sleep at least 8 hours a night
- Ability to abstain from
 - Alcohol for 24 hours before participation
 - Caffeine products for 12 hours before participation
 - Food for 3 hours before participation
- Non-smoker
- Not obese
- If female
 - Not taking oral contraceptives
 - Not pregnant

Three dosage levels of caffeine exist within the study, 0 mg, 20 mg, and 200 mg. The highest dosage level is roughly equivalent to a strong mug of coffee. The caffeine pills were received over the counter, and contain caffeine in the same form found in popular food and beverages. However, there are risks involved with consuming caffeine. Caffeine may increase anxiety (Boulenger, J P, Uhde, T, Wolff, E A, and Post, R M (1984) Increased sensitivity to caffeine in patients with panic disorders *Archives of General Psychiatry*, 41: 1067-1071), and interfere with sleep when consumed by certain individuals before bed (Levy, M and Zylber-Katz, E (1983) Caffeine metabolism and coffee-attributed sleep disturbances *Clinical Pharmacology and Therapeutics*, 33 (6): 770-775).

The duration of the study for each participant will be approximately one hour. During the first twenty minutes the caffeine will be given to the participant, instructions will be read, and a demonstration of the task will be performed. The final forty minutes will be the duration of the task.

To compensate participants, extra credit will be offered in class by your professor. The amount of extra credit is determined by your professor and is based on how much time is spent in the experiment.

Appendix B

Experiment Briefing / Consent Form

Experiment Briefing

Please understand that you have the right to refuse participation in this experiment at any time without penalty. If you refuse participation before signing the consent form, no extra credit for your class will be rewarded. If participation is later refused, extra credit will be awarded based on the amount of time spent participating, rounded to the nearest half hour.

The experiment you are about to participate in was created with the purpose of determining the benefits of caffeine on a vigilance task. With agreement to participate, and signature on the consent form you agree to receive a dose of caffeine that is equal to either 0, 20, or 200 mg. Caffeine will be administered in a gelatin caplet. Neither yourself nor the experimenter will be aware of which dose you receive at that time.

In order for your body to absorb the caffeine, you will not begin the computer task until 20 minutes after taking the caffeine. Throughout this time you will be given an eye test, thorough instructions of the computer task, perform a short computer-based demonstration of the task, complete a physiological health survey, complete the Stanford sleepiness scale, and be able to ask any questions that you may have.

After 20 minutes, you will commence the cognitive vigilance computer task. The computer task is 40 minutes in duration. Throughout this time, you are asked to perform to the very best of your abilities. The task consists of numbers presented on a computer monitor. When three odd or even numbers are presented in a row, you will be asked to press the space bar. After 40 minutes, the computer will stop presenting numbers and instruct you to return to the briefing room.

Earplugs will be worn to reduce noise distractions while completing the task.

The computer will keep track of the number of correct responses, incorrect responses, and misses. This data will allow the researcher to analyze the effects of caffeine on a vigilance task.

All information recorded will be held strictly confidential, meaning that names and data can be matched, but only members of the research team will have access to that information. Publication of the data will not include names.

An estimated duration of your time within the study is one hour.

There are risks involved with the study. The 200 mg dose is roughly equivalent to a strong mug of coffee, and risks are similar. Individuals respond differently to caffeine and you will be the best judge of how your body will react, based on your past experience with caffeine. If you would not be comfortable drinking a strong cup of coffee, you are urged not to participate in this experiment. Caffeine consumption at this dosage risks increased anxiety, impaired motor coordination, and disturbed sleep.

If you wish to know the results of the study, they will be published and available in the library under the author's name, Suzanne K. Robinson. The study should be published during the Spring semester of 2002.

CONSENT FORM

Department of Human Factors and Systems Embry-Riddle Aeronautical University

I consent to participating in the research project entitled:

The Performance Effects of a Low Dose of Caffeine on a Cognitive Vigilance Task.

The principle investigator of the study is:

Suzanne K. Robinson

The individual above, or their research assistants, have explained the purpose of the study, the procedures to be followed, and the expected duration of my participation. Possible benefits of the study have been described as have alternative procedures, if such procedures are applicable and available.

I acknowledge that I have had the opportunity to obtain additional information regarding the study and that any questions I have raised have been answered to my full satisfaction. Furthermore, I understand that I am free to withdraw consent at any time and to discontinue participation in the study without prejudice to me.

Finally, I acknowledge that I have read and fully understand the consent form. I sign it freely and voluntarily. A copy has been given to me.

Date: _____

Name (please print): _____
(Participant)

Signed: _____
(Participant)

Signed: _____
(Researcher/Assistant)

Appendix C
Health Survey

Health Survey

Participant ID Number _____

Age: _____

Height: _____

Weight: _____

Gender: _____

Please answer the following questions as accurately as possible:

	Yes	No
Are you in good general health?		
Is your daily caffeine intake usually between 200-400 mg? (Refer to table 1)		
Have you abstained from caffeine for the last 10 hours?		
Have you abstained from alcohol for the last 12 hours?		
Have you abstained from food for the last 3 hours?		
Are you a non-smoker?		
Do you regularly sleep 8 hours each night?		
If female, are you not taking oral contraceptives?		
If female, are you not pregnant?		

If you answered 'no' to any of the above questions, please describe your situation below.

Appendix D
Stanford Sleepiness Scale

Stanford Sleepiness Scale

Participant ID Number _____

Date: _____ Time: _____

This is a quick way to assess your alertness. Please circle the number that best

Degree of Sleepiness	Scale Rating
Feeling active, vital, alert, or wide awake	1
Functioning at high levels, but not at peak; able to concentrate	2
Awake, but relaxed; responsive but not fully alert	3
Somewhat foggy, let down	4
Foggy; losing interest in remaining awake; slowed down	5
Sleepy, woozy, fighting sleep; prefer to lie down	6
No longer fighting sleep, sleep onset soon; having dream-like thoughts	7

describes your degree of sleepiness.

(Hoddes, Dement, and Zarcone, 1972)

Appendix E

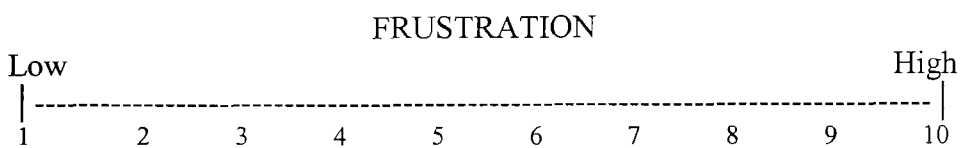
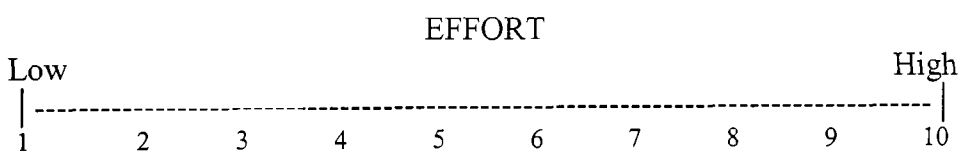
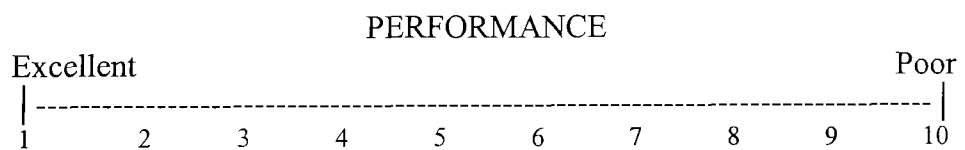
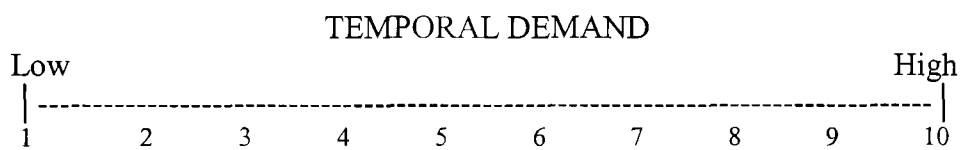
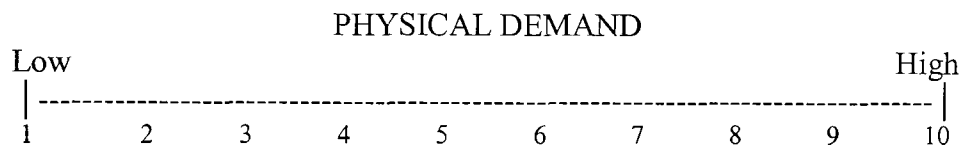
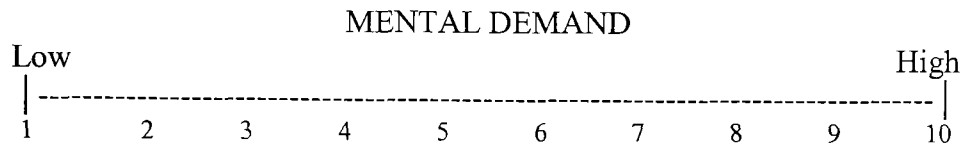
NASA-TLX

NASA-TLX Rating Scale Definitions

Title	Endpoints	Descriptions
Mental Demand	Low/High	How much mental and perceptual activity was required (e.g. thinking, calculating, remembering, and searching)? Was the task easy or demanding, simple or complex?
Physical Demand	Low/High	How much physical activity was required (e.g. pushing, pulling, turning, controlling, activating)? Was the task easy or demanding, slow or brisk?
Temporal Demand	Low/High	How much time pressure did you feel due to the rate or pace at which the task or task elements occurred? Was the pace slow and leisurely or rapid and frantic?
Effort	Low/High	How hard did you have to work (mentally and physically) to accomplish your level of performance?
Performance	Excellent/Poor	How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals?
Frustration Level	Low/High	How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content and relaxed did you feel during the task?

NASA-TLX Rating Sheet

Instructions: On each scale, place a mark that represents the magnitude of that factor in the task you just performed.



Pairwise Comparison of Factors

Instructions: Circle the member of each pair that provided the most significant source of variation in the task(s) that you just performed.

PHYSICAL DEMAND / MENTAL DEMAND

TEMPORAL DEMAND / MENTAL DEMAND

PERFORMANCE / MENTAL DEMAND

FRUSTRATION / MENTAL DEMAND

EFFORT / MENTAL DEMAND

TEMPORAL DEMAND / PHYSICAL DEMAND

PERFORMANCE / PHYSICAL DEMAND

FRUSTRATION / PHYSICAL DEMAND

EFFORT / PHYSICAL DEMAND

TEMPORAL DEMAND / PERFORMANCE

TEMPORAL DEMAND / FRUSTRATION

TEMPORAL DEMAND / EFFORT

PERFORMANCE / FRUSTRATION

PERFORMANCE / EFFORT

EFFORT / FRUSTRATION

Appendix F
Tables of Mean Differences

Detection Rate Mean Differences

Time (minutes)	Source Caffeine Dose (mg)	Mean	SD	N
0-10	200	32.6	6.1	25
	20	29.2	7.5	21
	0	31	5.2	23
	Total	31.0	6.3	69
10-20	200	32.4	7.1	25
	20	28.9	7.9	21
	0	29.3	5.5	23
	Total	30.3	7.0	69
20-30	200	32.5	8.1	25
	20	27.2	9.0	21
	0	30.1	6.1	23
	Total	30.1	8.0	69
30-40	200	34.7	6.9	25
	20	27.4	9.5	21
	0	29.8	5.1	23
	Total	30.1	7.8	69

False Alarm Mean Differences

Time (minutes)	Source Caffeine Dose (mg)	Mean	SD	N
0-10	200	13.7	6.3	25
	20	13.8	10.6	21
	0	12.3	5.0	23
	Total	13.3	7.4	69
10-20	200	12.9	4.5	25
	20	11.2	7.9	21
	0	13.8	10.2	23
	Total	12.7	7.8	69
20-30	200	15.8	8.7	25
	20	14.8	9.6	21
	0	12.4	10.5	23
	Total	14.3	9.6	69
30-40	200	17.0	9.7	25
	20	9.9	5.3	21
	0	12.2	10.8	23
	Total	13.3	9.4	69

Response Time Mean Differences

Time (minutes)	Source Caffeine Dose (mg)	Mean	SD	N
0-10	200	810.4	60.1	25
	20	812.2	58.7	21
	0	825.6	60.6	23
	Total	816.0	59.4	69
10-20	200	814.9	57.5	25
	20	811.3	61.6	21
	0	801.3	58.2	23
	Total	809.3	58.4	69
20-30	200	815.8	62.3	25
	20	785.6	102.6	21
	0	814.5	58.8	23
	Total	806.5	75.5	69
30-40	200	817.7	59.1	25
	20	766.5	69.2	21
	0	814.6	62.0	23
	Total	801.6	66.3	69

Stanford Sleepiness Scale Mean Differences

	Source	Mean	SD	N
1	Before or After Task (1 or 2)			
	Caffeine Dose (mg)			
	200	2.16	0.687	25
	20	2.24	0.539	21
2	0	2.09	0.793	23
	Total	2.16	0.678	69
	200	3.28	1.487	25
	20	3.76	1.261	21
2	0	3.43	1.674	23
	Total	3.48	1.481	69

NASA-TLX Mean Differences

Source	Mean	SD	N
Caffeine Dose (mg)			
200	72.8	7.73	25
20	79.5	7.05	21
0	79.0	7.15	23
Total	76.9	7.86	69

Appendix G

Tables of Subsequent Tests

Subsequent Tests for Detection Rates

Caffeine	Time Block	Time Block	Mean Difference	Standard Error	Significance	95% Confidence Interval	
						Lower Bound	Upper Bound
200 mg	1	4	-2.080	1.083	.059	-4.243	-.0829
20 mg	1	4	1.857	1.182	.121	-.503	4.217
0 mg	1	4	1.217	1.129	.285	-1.038	3.472

Time Block	Caffeine	Caffeine	Mean Difference	Standard Error	Significance	95% Confidence Interval	
						Lower Bound	Upper Bound
4	0 mg	20 mg	2.402	2.207	.280	-2.004	6.808
	0 mg	200 mg	-4.897	2.113	.024	-9.115	-.680
	20 mg	200 mg	-7.299	2.164	.001	-11.620	-2.978

Subsequent Tests for False Alarms

						95% Confidence Interval	
Caffeine	Time Block	Time Block	Mean Difference	Standard Error	Significance	Lower Bound	Upper Bound
200 mg	1	4	-3.240	1.885	.090	-7.003	.523
20 mg	1	4	3.952	2.056	.059	.153	8.058
0 mg	1	4	.130	1.965	.947	-3.792	4.053

						95% Confidence Interval	
Time Block	Caffeine	Caffeine	Mean Difference	Standard Error	Significance	Lower Bound	Upper Bound
4	0 mg	20 mg	2.455	2.719	.370	-2.972	7.883
	0 mg	200 mg	-4.743	2.602	.073	-9.939	.453
	20 mg	200 mg	-7.198	2.666	.009	-12.521	-1.875

Subsequent Tests for Response Times

Caffeine	Time Block	Time Block	Mean Difference	Standard Error	Significance	95% Confidence Interval	
						Lower Bound	Upper Bound
200 mg	1	4	-7.320	14.252	.609	-35.774	21.134
20 mg	1	4	46.286	15.550	.004	15.239	77.332
0 mg	1	4	10.957	14.858	.463	-18.709	40.622

Time Block	Caffeine	Caffeine	Mean Difference	Standard Error	Significance	95% Confidence Interval	
						Lower Bound	Upper Bound
4	0 mg	20 mg	48.752	18.936	.012	10.944	86.559
	0 mg	200 mg	-3.071	18.127	.866	-39.263	33.121
	20 mg	200 mg	-51.823	18.571	.007	-88.902	-14.744

Subsequent Tests for Stanford Sleepiness Scale Scores

Caffeine	SSS	SSS	Mean Difference	Standard Error	Significance	95% Confidence Interval	
						Lower Bound	Upper Bound
200 mg	1	2	-1.120	.297	.000	-1.712	.528
20 mg	1	2	-1.524	.324	.000	-2.170	-.878
0 mg	1	2	-1.348	.309	.000	-1.965	.731

SSS	Caffeine	Caffeine	Mean Difference	Standard Error	Significance	95% Confidence Interval	
						Lower Bound	Upper Bound
2	0 mg	20 mg	-.327	.450	.470	-1.225	.571
	0 mg	200 mg	.155	.430	.720	-.705	1.014
	20 mg	200 mg	.482	.441	.278	.399	1.362

Subsequent Tests for NASA-TLX Scores

		95% Confidence Interval				
Caffeine	Caffeine	Mean Difference	Standard Error	Significance	Lower Bound	Upper Bound
0 mg	20 mg	-.4762	2.21439	.830	-4.8974	3.9450
	200 mg	6.1600	2.11977	.005	1.9278	10.3922