Order Number 9020745

The effects of caffeine and noise on visual two-pulse discrimination and vigilance

Bienstock, Barbara, Ph.D.

City University of New York, 1990

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

/1

THE EFFECTS OF CAFFEINE AND NOISE ON VISUAL TWO-PULSE DISCRIMINATION AND VIGILANCE

by

BARBARA BIENSTOCK

A dissertation submitted to the Graduate Faculty in Psychology in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York.

1990

This manuscript has been read and accepted for the Graduate Faculty in Psychology in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

1/25/90
Date

Chair of Examining Committee

January 26, 1990
Date

Executive Officer

Dr. Gad Hakerem

Dr. Philip H. Ramsey

Dr. Stanley Novak

Dr. Jacques Rutschmann

Supervisory Committee

The City University of New York

Abstract

THE EFFECTS OF CAFFEINE AND NOISE ON VISUAL TWO-PULSE DISCRIMINATION AND VIGILANCE

by

Barbara Bienstock

Advisor: Dr. Gerard Bruder

The present study investigated the effects of caffeine and noise on arousal and activation. In double-blind fashion, 12 male college students (18-26 years old) were tested on the effects of three doses of caffeine (0.25 mg/Kg; 2.0 mg/Kg; and 4.0 mg/kg) with and without a white noise background. Three visual tasks were employed: a threshold task, to set sensation level for the main tasks and to check for changes in sensitivity across session and treatment conditions; a two-pulse discrimination task, to test phasic arousal; and a two-pulse vigilance task, to test tonic activation. The light stimulus was presented foveally. The 70 dB SPL(A) white noise was administered via loudspeaker.

Data were analyzed in terms of signal detection theory, which allowed both sensitivity (d'e) and response bias (log beta

and arcsin false alarm) measures to be obtained. Rating scale Receiver Operating Characteristics (ROC) curves were plotted for the vigilance data. A 3 x 2 x 2 Analysis of Variance (ANOVA) of the vigilance data yielded, for d'_e, a significant decrement in performance over time; a significant interaction for noise and time, and a significant interaction for noise, dose, and time. The noise background reduced the vigilance performance decrement over time and caffeine moderated that effect. There were no significant log beta effects. However, the arcsin false alarms increased significantly with dose, indicating that subjects' decision criteria became laxer with increasing dose. Contrary to expectation, A 3 x 2 ANOVA of the threshold, as well as of the two-pulse discrimination data, yielded no significant effects of noise or caffeine. In conclusion, there was evidence for effects of noise and caffeine on tonic activation, as measured by a vigilance task, but not phasic arousal, as measured by two-pulse discrimination.

ACKNOWLEDGMENTS

At the outset of this study, two people, whose names do not appear on my dissertation committee, were originally on it. These were Dr. Mitchell L. Kietzman and Dr. Samuel Sutton. Dr. Sutton (my "co-sponsor"), who was the head of the Psychophysiology (now Biopsychology) Department of the New York State Psychiatric Institute, was very much involved with the design and implementation of this study, which ended with his unfortunate and untimely death. It was a privilege to work under him and he is very much missed. Dr. Kietzman, who was my sponsor and mentor from the beginning of this study and just prior to its defense, was required, for medical reasons, to retire early from his academic position. I thank Dr. Kietzman for his guidance throughout my graduate career and for being a good friend and making the rough going only a small hurdle to pass. It is thus with gratitude and sorrow that I thank Dr. Kietzman and Dr. Sutton for their tremendous support and their time.

Dr. Gerard Bruder graciously consented to take over Dr. Kietzman's place as my sponsor. I thank him for freely giving his time and attention and for his extremely useful comments on my dissertation. He "pitched in" when things looked bleak. I

also would like to thank Drs. Philip Ramsey and Gad Hakerem for serving on my committee and for their valuable contributions. I am grateful to Drs. Stanley Novak and Jacques Rutschmann for serving on my committee and for their very helpful comments.

This study was conducted at the Vision Laboratory of the Psychophysiology Department of the New York State Psychiatric Institute. I would like to thank the entire staff of the department for being there for me when I needed it, professionally and for moral support. I, especially, would like to thank my friend and colleague, Dr. Joseph E. Herskovic, for his participation in this study by keeping track of the stimulus conditions and thus safeguarding the integrity of this double-blind study.

Many thanks to Ray Simon for building the optical equipment and to Ben Djoleto for his maintenance of the system.

Also, many thanks to John Leong of the Computer Department at Queens College for his assistance with data preparation for statistical analysis and to Dr. David Friedman, of the New York State Psychiatric Institute, for his assistance with the epsilon correction factor.

To my subjects: Please take a bow! You deserve it.

I would like to thank my late grandmother and my parents for instilling in me a love of education early in life.

Last, but not least, I would like to thank my husband

Manny (Menachem Zachodin) for his work on the Tables and for his tremendous assistance in the final preparation of this manuscript, but without whom this study could have been finished much earlier -- it was interrupted by his giving me the greatest gift: my daughter.

This thesis is dedicated to my four-year daughter, Maya, in the hope that some day it will inspire her.

TABLE OF CONTENTS

Section		Page
I.	Introduction	1
II.	Method	20
III.	Results	54
IV.	Discussion	84
V.	Appendices	103
VI.	References	114

LIST OF TABLES

Table — Page

1. Cell Means and Standard Deviation of the d'es for the Twelve Subjects Under All Vigilance Conditions..... 55

2. ANOVA (3 x 2 x 2) for d'e..... 57

3. (a). Mean d'e and Standard Deviation for Period (Combined across dose and background)..... 59

(b). Mean d'e and Standard Deviation for Dose (Combined across period and background).. 59

(c). Mean d'e and Standard Deviation for Background (Combined across dose and period)..... 59

4. Comparison of Background and Period (Mean d'es (Combined across dose)..... 60

5. Log Beta Cell Means and Standard Deviation for Dose, Background and Period..... 69

6. ANOVA (3 x 2 x 2) for Log Beta..... 70

7. Arcsin False-Alarm Cell Means and Standard Deviation for Dose, Background, and Period.... 72

8. ANOVA (3 x 2 x 2) for Arcsin of the False-Alarms..... 74

9. Two-Pulse Discrimination Thresholds (ms)..... 76

10. ANOVA (3 x 2 x) for Discrimination Thresholds..... 77

11. Threshold Energy (mL * msec) Obtained at Beginning and End of Each Session (Dose)..... 78

12. Spearman-Rho Rank Correlations of d'es and Two-Pulse Discrimination Thresholds..... 82

LIST OF FIGURES

Figure		Page
1.	A diagram of the one-channel optical system..	23
2.	Rating ROC curves for one subject. Condition: Low dose, no-noise background.....	49
3.	Mean d'es (Combined across dose) for no-noise and noise background conditions.....	63
4.	Mean d'es for the three doses under each of the no-noise and noise background conditions.	65
5.	Differences in d'es between periods for the three doses of caffeine for both background conditions.....	68

THE EFFECTS OF CAFFEINE AND NOISE ON MEASURES OF TWO-PULSE DISCRIMINATION AND VIGILANCE

The overall aim of the present study was to examine the effects of caffeine and noise on arousal and activation. The terms "arousal" and "activation" have been used and misused in the literature in many ways. Some researchers have used the two terms interchangeably, while others have conceptualized them as separate underlying processes or levels of functioning. In addition, there is also a lack of consistency in the definitions of these processes. In the present study, arousal and activation were conceptualized along the lines of Pribram and McGuiness (1975) with arousal being related to a phasic level of functioning and activation referring to underlying changes in performance with time. Two different tasks commonly employed to test these two processes were used: a two-pulse discrimination procedure to test for changes in phasic arousal, and a vigilance task to test for changes in tonic

activation. Thus, in the present study, these two processes are being defined operationally, by the tasks traditionally used to measure them.

Caffeine

Caffeine is one of the most widely used psychoactive substances in the world. Millions of people experience its effects on a daily basis and have done so for many years. Although caffeine has been in use for more than a thousand years, it was not until 1980, with the publication of the third edition of the American Psychiatric Association's Diagnostic and Statistical Manual of Mental Disorders (DSM-III), that psychiatry officially recognized the possibility of disorders stemming from the use of caffeine; they included caffeine disorders in the psychiatric nomenclature of disorders. The lack of clinical recognition until now was not because the effects of caffeine were unknown. During prohibition, some prohibitionists demanded that not only alcohol, but also coffee and tea be banned as well. Also, at the turn of the century, coffee raids by authorities occurred periodically throughout the world (Greden, 1980).

The lack of recognition of the clinical significance of caffeine is perhaps due to the fact that it is

such a common substance and is found in most households. Even children are exposed to caffeine early in life since it is an ingredient of many soft drinks and candy bars. The statistics are startling; for example, it has been estimated that in 1977, infants, age 6 to 11 months, consumed a mean of 4.2 mg of caffeine per day. Considering only those infants that actually consumed caffeine, the average intake of these infants was 77 mg per day. Also the intake of caffeine by children increases with age (Graham, 1978).

Although caffeine is not commonly referred to as a drug, it has potent psychological effects and may even induce distinct syndromes. One of the syndromes, caffeineism (caffeine intoxication), is mentioned in the psychiatric nomenclature (DSM-III), although it is not classified as a mental disorder. Caffeine withdrawal is not mentioned. This decision may be because there is no maladaptive behavior associated with the use of caffeine, and because it is so socially acceptable.

The prevalence of caffeine in our society is obvious. It is common fare in the morning. Many people do not start their day without a morning cup of coffee. When dinner is eaten, coffee and cake follow. In most American restaurants, when coffee is served, it is served in unlimited quantities. The media, in its advertising, portrays drinking coffee as an enjoyable, thoroughly relaxing experience, or one that helps a

person make it through the day. Either approach associates coffee with pleasant social experiences. Of course, there are also coffee houses and tea houses and there are also coffee breaks.

Unfortunately, the effects of caffeine, both positive and negative, have been greatly exaggerated due to inadequate documentation, although some documentation has been available. The long standing popularity of caffeine and caffeine-containing substances originated with the ancient belief that this substance had a stimulant action that elevated mood, decreased fatigue, and increased capacity to work.

The most common source of caffeine is beverages (Graham, 1978). Persons ingesting these beverages report experiencing a reduction in drowsiness and fatigue, and a more rapid and clearer flow of thinking. Under experimental conditions, caffeine has had similar effects in that it produces faster reaction times, significant increases in muscular work in humans, greater capacity for sustained intellectual effort, and a better association of ideas (Ritchie, 1975). For example, with caffeine, typists type faster and with fewer errors. However, the effects of caffeine on newly acquired fine motor skills that require muscular coordination and accurate timing may be adverse (Goldstein, 1965). The above mentioned effects can be produced by as little as 85 to 250 mg of caffeine (Weiss & Laties, 1962),

the amount present in one to three cups of coffee.

A very large percentage of the population consumes caffeine. Many people consume more than 250 mg per day, and from 20 to 30% consume over 500 mg per day, which far exceeds the psychologically active dose which is about 250 mg. Thus, a large percentage of the population seems to be at risk for the adverse effects of caffeine.

Caffeine ingestion can become a medical problem since it can produce substantial effects on several organ systems. Caffeine is a central nervous system stimulant that also acts as a diuretic, stimulates cardiac muscle, and relaxes smooth muscle (Rall, 1980). As dosage increases above 250 mg, it may produce nervousness, restlessness, tremors, insomnia, excitement, gastrointestinal irritations, and diarrhea. At even higher doses, restlessness and insomnia may progress to delirium, and sensory disturbances from caffeine such as flashes of light and ringing in the ears are not uncommon. Nausea and vomiting, cardiac arrhythmia, and definite tachycardia also occur, as may convulsions and seizures. Fatal poisoning by an overdose of caffeine has occurred, but it is rare (Dimaio & Garret, 1974).

Given the fact that there are several possible side effects of caffeine and that there is some seemingly contradictory evidence, as well as the possibility of there

being individual responses to the drug, it is surprising that more research on caffeine has not been done. The lack of systematic research over the past several decades represents a scientific vacuum.

Most of the research that has been done on caffeine has been on its physiological effects, while research on its effects on performance has been relatively scant. Since caffeine is known to accelerate certain physiological functions, it is logical to assume that caffeine would affect motor performance in a similar manner. Smith, Tong, and Leigh (1977) found that caffeine significantly improved decision time as well as motor time in a choice-reaction time task. However, hand steadiness was significantly impaired. Caffeine also shortened auditory reaction time over a 15 minute period, from one and-three-quarter to two hours after ingestion at doses of 75 and 150 mg. (Clubley, Henson, Peck, & Riddington, 1977).

It is hypothesized that the effects of caffeine on performance are achieved via its actions on the central nervous system and its actions through sympathetic excitation. There exists evidence that high levels of caffeine can be related directly to high levels of anxiety (Greden, 1974). Since caffeine is a stimulant, it is expected to increase one's level of excitation.

Arousal and Activation

Two basic processes have been postulated in the literature to describe the effects of increased levels of excitation (the term, here, being used conceptually to encompass both arousal and activation and, not to be confused with neuronal excitation): arousal and activation (Pribram & McGuiness, 1975; McGuiness & Pribram, 1980). Arousal has been considered to be related to a phasic level of functioning, or a response to a stimulus at a given moment in time, and is thought to be reflected in changes in amygdala activity. Lindsley (1960) has suggested arousal to be associated with changes in cortical activity.

The concept of activation is equally important to that of arousal. Very often performance is not only a function of impinging stimuli at a given moment, but also is a function of ongoing activity occurring throughout or ongoing activity changing with time. Tonic activation, then, refers to underlying changes in performance with time, or as Pribram and McGuiness (1975) propose, a tonic readiness to respond. Pribram and McGuiness relate activation to changes in the basal ganglia system. Lindsley suggests activation to reflect activity of the reticular activating system (RAS). The relationship between

arousal and activation has yet to be fully determined.

Behaviorally, an increase in the level of physiological activity or excitation, which would encompass both phasic arousal and tonic activation, does not necessarily increase or improve one's performance. One of the most common and often propounded theories of arousal, which derives from the original work of Yerkes and Dodson (1908), states that there is an inverted U-shaped function between performance and the level of excitation. When excitation is low, at the beginning of the continuum, performance is not expected to be optimal. However, as excitation increases, for whatever reason, performance improves and eventually reaches an optimal peak. But, as excitation continues to increase (beyond the optimum) a resulting state of "overexcitation" can produce a decrement in performance. Thus, low or poor performance could be attributed to either a low level of excitation or to a high level of overexcitation. This theory also implies that given levels of excitation are necessary for optimal performance in various tasks. The concepts of arousal and activation are important because they have been used to explain certain forms of mental illness. For example, schizophrenia has been related to a state of overarousal (Lapidus & Schmolling, 1975).

The concepts of arousal and activation often been used interchangeably in the literature with and without intent.

That is, researchers do not necessarily use both terms to mean one idea (although many do), but specify two concepts, with some researchers using the term "arousal" for one concept and others using the term "activation" for the same concept. Also, some researchers (Routenberg, 1968) propose two "arousal" systems and then refer to the second arousal system as being a result of "activation" of the first arousal system (a poor choice of words). Another example is Sharpless and Jasper's (1956) proposal of two types of "activation," tonic and phasic. To add more confusion, the concept of attention has also been used interchangeably with both arousal and activation. Many studies discuss a vigilance decrement (see below) in terms of arousal theory and, often, different researchers interpret arousal theory differently. Stroh (1971), in his very informative and important book on vigilance, when discussing vigilance theory, states (or accuses authors) very briefly that some researchers are "couched in semantics" i.e., buy using such terms as Attention Theory and Observing Responses. But well they should be couched in semantics. Stroh, himself, provides evidence that semantics, in this area of research, must be dealt with, since he too is guilty of adding to the confusion (and thus taking something away from his thorough and well written book). For example, in trying to support his statement that those who condemn arousal theory often misinterpret it, he cites Bevan et.

al. and begins with with the following quote "following the activation hypothesis." He (Stroh) then goes on to say "This is not what one should expect from the activation hypothesis. The mere fact that the monotony of the vigilance situation is being disrupted by introducing a new factor (sensory restriction) will serve as an arousing factor." He does, however, redeem himself somewhat several pages later, by stating that "It can be seen that what was once a clear-cut and unitary concept, is steadily becoming more and more complex, under stress of experimentation." Nonetheless, trying to formulate new theories and unifying existing ones with experimental findings becomes a very tedious, confusing, and difficult task.

In the present study, there is agreement with Stroh that a unitary concept is no longer sufficient and that there may be two systems or two processes operating: a phasic process (in this paper referred to as arousal) and a tonic process (in this paper referred to as activation). That is, one can look at responses on a trial by trial basis and see how the subject responds to a specific trial or series of trials (phasic process) and one can also look at how reactions or responses to these stimuli change over prolonged periods of time (activation process), or responses to the whole environment. What may be very useful is to define the concepts operationally, that is, by the tasks used to measure them.

One task that has frequently been used as a measure of attention/activation is the vigilance task--or sustained attention task. In a typical vigilance task an observer must detect an infrequent or subtle stimulus change occurring in a display over time; for example, detecting a clock hand moving two steps at once three or four times per minute. Thus, vigilance, then, can be defined as the ability to sustain attention. N. H. Mackworth's (1950) and J. F. Mackworth's (1968) classical studies of vigilance have shown that the typical finding of such tasks is a decrement in performance with the passage of time.

Data from vigilance tasks often permit a signal detection analysis, thereby providing a measure of performance sensitivity (d'), and measures of response bias (beta and arcsin false alarm), which are assumed to be independent of each other. Signal detection theory assumes that discriminations of stimuli are performed against a constant background of neural noise and that all decisions are made against such a noise background with various degrees of uncertainty. Stimuli, therefore, make up a signal-plus-noise distribution. Thus decisions are based on both sensory and non-sensory factors, and these are separated by statistical analysis.

Two other assumptions underlie signal detection theory: 1) the noise and the signal-plus-noise distributions are

normal (Gaussian), and 2) these distributions are of unit (equal) variance (Green & Swets, 1974). However, Jerison (1977) pointed out that these two assumptions are unlikely to be fulfilled in most vigilance studies. Warm and Jerison (1984), though, indicate that these two conditions make computation of signal detection theory data much easier and that is often the reason why they are assumed. If one of these two assumptions is not met, signal detection theory allows for alternative measures to use for sensitivity and response criteria. Thus, for example, if the unit variance assumption is not met, d'e -- another parametric measure of sensitivity -- can be used instead of d'. As far as response criterion is concerned, log beta can be used instead of beta or the arcsin of the false alarms. Thus, by the application of signal detection theory to vigilance studies, activation may be discussed in terms of sensory and nonsensory factors.

Noise

It has been argued (Lindsley, 1960) that behavior could be viewed as a continuum of arousal ranging from sleep to agitated excitation. Some investigators have equated the level of arousal in the central nervous system to the concept of drives (Hebb, 1955) and have pointed out that changes in the environment such as temperature, light, etc., can be expected to

increase the level of arousal since neural pathways carrying information from these stimuli pass through the reticular activating system on their way to the cortex. An additional stimulus that has been documented to be an arousing one is noise (Broadbent, 1971; Frith, 1967; Hebb, 1955). For example it has been shown that white noise improves resolution of flickering lights (Harper, 1979), which have been used as indicators of levels of arousal.

However, Fisher (1983) points out that noise may be argued to be a de-arousing stimulus since it can mask some aspects of other auditory stimulation in the environment. Poulton (1977; 1979) proposes that better performance (positive effect) under noise is associated with its effects on arousal while poorer performance (negative effect) is due to noise's masking essential auditory information as well as inner speech, which is required for rehearsal in memory. Broadbent (1978) argues that there are cases where deterioration occurs but auditory cues are unlikely to be involved and that there are cases where a shift in decision criteria, a non-sensory factor, are evident, thus not supporting Poulton's position. Broadbent supposes that all aspects of performance that change under noise are due to its changing of arousal level, thus, for example, complex tasks should show deterioration because arousal is too high. Poulton (1979) compromises by stating that positive

effects of noise are due to arousal shift, but negative effects are due to masking; the benefits of arousal change may overcome the negative effects of masking.

The concept of varying levels of excitation can be applied to the action of drugs on behavior as well as to environmental stimuli, such as noise, since many drugs are known to affect RAS activity. Thus, depending on what effect a drug has on the RAS, that effect can be altered by environmental stimuli. That is, there could well be an interaction between drugs and environmental stimuli that results in effects that are different than when either is presented alone.

Purpose of Study

The present study investigated the effects of caffeine and white noise on arousal and activation by administering to the subjects three doses of caffeine (the low dose was essentially a placebo dose), each presented on one of three separate days. Performance was assessed with and without a white noise background used as an additional load-imposing stimulus.

One rationale for this study was to investigate phasic arousal and tonic activation as different processes by employing two tasks: two-pulse discrimination as a measure of

phasic arousal and vigilance as a measure of tonic activation. Phasic arousal measured the ability of the subjects to respond to immediate changes, similar to orienting responses. Tonic activation was tested by a procedure for which the demands were different, i. e., were not as immediate but were associated with relatively long-term changes over time.

Another rationale for this study is to determine the effects of caffeine on arousal and activation. With the advent of new technologies and their extremely rapid growth (for example computer hardware, software and displays), there is greater reliance on monitoring tasks. In addition, there is an increase in the number of motor vehicles on the roads and longer traffic delays. Both of the above situations are examples of situations which can put people in hypnotic-like trances, or in situations where novel stimuli are increasingly absent. These situations are similar to what happens in laboratory vigilance tasks. If caffeine is indeed an alerting stimulus, then performance over the duration of the vigilance task should be better than without the use of caffeine. Furthermore, is there also an improvement of performance with caffeine on threshold and two-pulse discrimination tasks that do not look at changes over time? That is, this study was to find out if caffeine affects tasks that require maintained attention differently than it does tasks that do not.

Tasks

Subjects were tested on three tasks: an absolute visual threshold, a two-pulse discrimination threshold, and a measure of vigilance. The absolute visual threshold was obtained so that a luminance level could be established for the stimuli used in the subsequent tasks and so that comparisons could be made across sessions and treatments. The visual discrimination task was a two-pulse discrimination threshold task, which can be viewed as an indicator of phasic arousal. Measures of resolution have been traditionally used as indicators of arousal (Lindsley, 1960; Venables, 1963). The measure of resolution traditionally used was the two-flash threshold, which is similar to the two-pulse discrimination threshold. In both the two-flash threshold task and in the discrimination task the stimuli used are two flashes of light separated by various dark intervals. In the former task, the subject must say whether he perceived one or two flashes of light and in the latter task, must choose which of three stimuli is different. The different stimulus in the discrimination task is the one with a longer dark interval separating the two pulses. The other stimuli have the same minimal dark interval. It is possible that brightness may be a cue that some observers use in the discrimination task, but

resolution can also play a role. In the classical temporal integration function (Kietzman & Gillam, 1972), there exists an area of partial integration. The stimuli used in the discrimination task in the present study fall within that area, where only some of the energy of the two-pulse stimulus is being integrated, but in which case some resolution can also occur. The two-flash threshold, presumably, is in the region where partial integration ends and no integration begins. Thus, two-pulse discrimination can be viewed as a precursor to the two-flash threshold. This view is supported by findings of two-pulse discrimination threshold values, traditionally, being of shorter durations than the two-flash threshold values. In both tasks brightness and resolution may be used as cues in the observer's response decision. In addition, Tong (1970), in a discussion of the two-flash threshold states that the threshold is "compounded by the subject's neural sensitivity and discrimination criterion he adopts in what is essentially a series of decisions as to whether the signal can be described as 'one' or 'two' flashes."

The vigilance task, which is a continuous monitoring task, can be used as an indicator of tonic activation, since results are measured and compared over relatively long periods of time. Thus, the effects of both caffeine and noise can be assessed on two distinct types of

excitation -- arousal and activation. In a typical vigilance task, an observer must detect an infrequent or subtle stimulus change occurring in a display over time; for example, detecting a clock hand moving two steps at once three or four times per minute). Vigilance then, can be defined as the ability to sustain attention.

Data from the present vigilance task permitted a signal detection analysis, thereby providing a measure of performance sensitivity (d'e -- a parametric signal detection measure analogous to d') and measures of response criterion (beta and arcsin false alarm). Signal detection theory assumes that discriminations are performed against a constant background of neural noise and that all decisions are made against such a noise background with various degrees of uncertainty. Thus, decisions are based on both sensory and non-sensory factors, and these are separated by the statistical analysis upon which signal detection theory is based.

In addition to analyzing data within a given task, attempts were made to compare performance across tasks. That is, rank correlations were performed in order to determine whether arousal and activation are parallel processes, independent processes or ordered on one continuum of excitation. Findings are discussed with reference to existing theories of arousal and activation.

Hypotheses

In the vigilance task, it was expected, as in all vigilance tasks, that performance would decrease with time. It was also expected that caffeine, as a stimulant, would improve performance. Also, in the vigilance task, performance will not deteriorate as much over time for the higher dosages as for the low (essentially placebo) dose. Since in vigilance tasks performance deteriorates with time, presumably due to the reduction of activation, caffeine, with its stimulating properties, should reduce or eliminate the vigilance decrement. It was also expected that white noise will have effects that are similar to caffeine, that is, will improve performance, or reduce the vigilance decrement. However, the combination or interaction, of these two variables may improve performance, or may actually produce excessive excitation resulting in a decreased level of performance.

Similar results were expected in the two-pulse discrimination task. It was expected that noise would improve (shorten) the two-pulse discrimination threshold. Caffeine, too, was expected to improve the discrimination threshold for both noise and no-noise background conditions.

Method

Subjects

Subjects for the experiment were 12 male college students between the ages of 18 and 26 years, with a mean age of 20.3 years (SD=2.1). Subjects were recruited from college campuses through bulletin board advertisements and word of mouth. All subjects were prescreened with respect to caffeine consumption by the use of a Substance Intake Packet I (SIPI) caffeine questionnaire (see Appendix I). Subjects chosen for participation in the experiment were all low caffeine users, i.e., they consumed less than 150 mg of caffeine per day. It was decided to test subjects who were low caffeine users in order to avoid possible confounding effects of caffeine tolerance (Colton, Gosselin, & Smith, 1968).

Eleven out of the 12 subjects had normal 20/20 vision while one of the subjects had 20/20 vision with corrective lenses. All subjects were naive with respect to the purpose of the experiment. They were, however, told that the effects of caffeine on visual performance were being investigated.

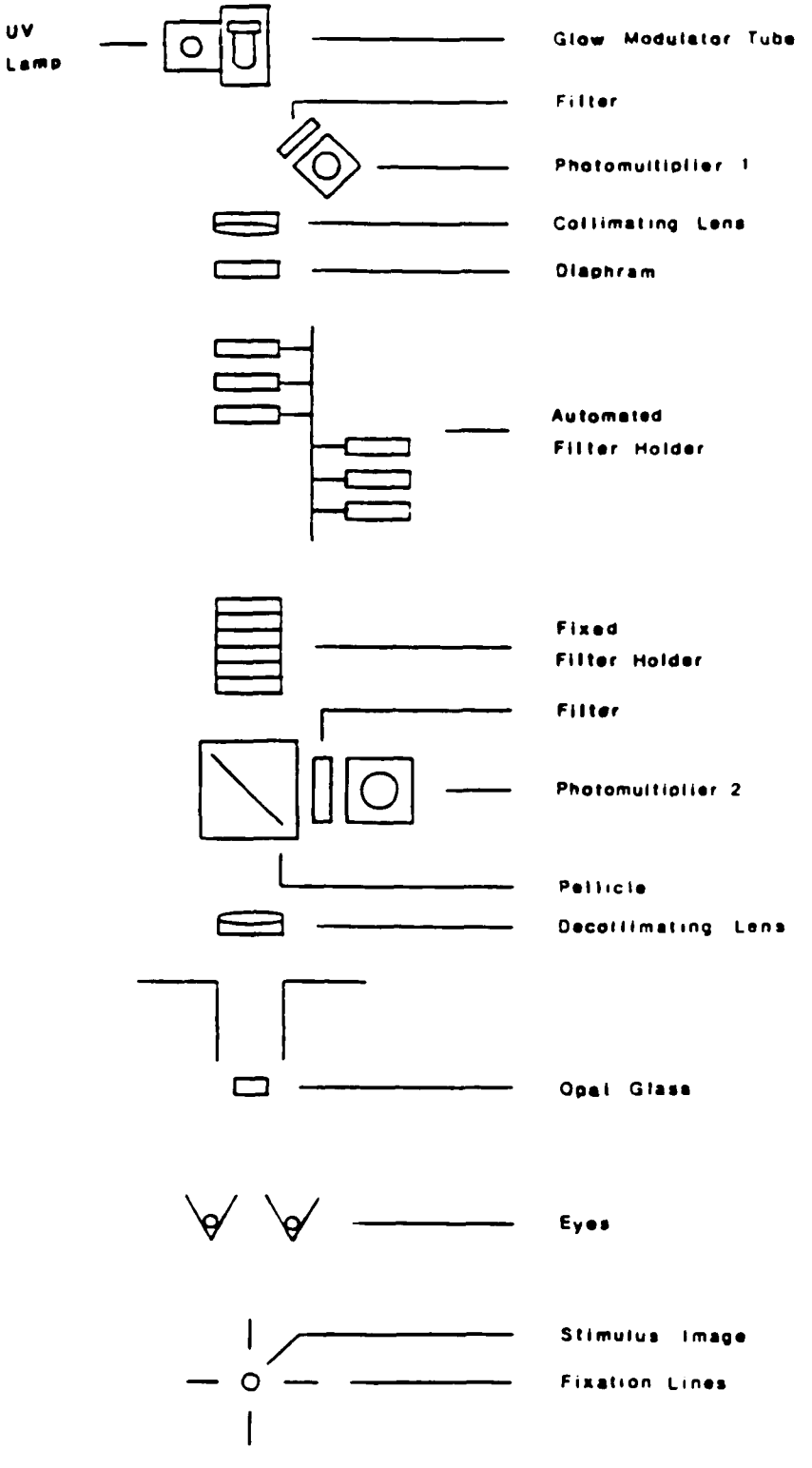
All subjects were informed that they were to be given caffeine in capsule form, in three separate doses, each administered on a different testing day, with the highest dose being equivalent to about two to two-and-one-half cups of coffee (depending, of course, on the weight of the subject, since dosage was based on weight, i.e., the heavier the subject, the more caffeine he received). Informed consent forms were signed by all subjects. They were compensated for their services by being paid $5.00 per hour for the time they participated.

Apparatus

The circular white-appearing light stimulus was produced by a glow-modulator gas discharge tube (Sylvania R1131C) whose rise and decay times were approximately 20 microseconds and 10 microseconds, respectively (see Kietzman & Gillam, 1972). A constant current of 23 mA was supplied to the glow modulator tube. An argon ultraviolet lamp (General Electric AR-4) irradiated the glow modulator and provided short and stable ionization times of less than 20 microseconds (see Matin, 1964). The resultant pulses, therefore, were rectangular.

The light from the glow modulator passed through a single-channel optical viewing system (see Figure 1.) consisting of a collimating lens, filters, and a decollimating lens. Light

Figure 1. A diagram of the one-channel optical system.


UV
Lamp
Glow Modulator Tube
Filter
Photomultiplier 1
Collimating Lens
Diaphram
Automated
Filter Holder
Fixed
Filter Holder
Filter
Photomultiplier 2
Pellicle
Decollimating Lens
Opal Glass
Eyes
Stimulus Image
Fixation Lines

emitted from the glow modulator tube was immediately passed through a six-filter intensity programmer, which consisted of a box containing Tiffen metallic filters that enabled the experimenter to manipulate the intensity, by programming the appropriate filter combination values, on a trial by trial basis. The light then passed through another filter box (the common filter box), which contained gelatin neutral density filters (Kodak Wratten No. 6) inserted manually. This common filter box held the calibrating filters and any filters needed to reduce the light intensity for a given task. For a complete schematic of the system see Berenhaus (1981).

The intensity programmer, which was driven by solenoids, was controlled by switches manipulated by the experimenter. It allowed 64 possible combinations of intensity over a range of 1.4 log units. This enabled the experimenter to use steps of 0.1 log units for the absolute threshold procedure. The common filter box enabled the experimenter to insert filters in order to reduce the light intensity emitted by the glow modulator so that intensities near threshold can be obtained for use in the absolute threshold procedure. These filters, used in combination with the intensity programmer, enabled the experimenter to raise the threshold intensity by 1.0 log units for the discrimination tasks. The common filters, as well as the 64 possible combinations of the intensity programmer, were

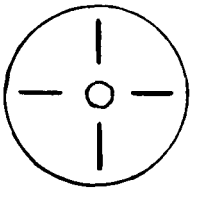
calibrated in the laboratory (see Appendix II).

Two photomultiplier tubes (RCA 1P21), whose inputs were filtered (Kodak Wratten No. 106) to approximate the C. I. E. photopic-luminance-efficiency curves, monitored relative luminance before and after filtering. Photomultiplier 1 received its input directly from the glow modulator tube while photomultiplier 2 received reflected light from a pellicle beamsplitter (National Photocolor Corp.) Situated on the other side of the filters from the glow modulator tube. An oscilloscope (Tektronix 532) display of the photomultiplier outputs enabled the experimenter to monitor intensity and duration characteristics of the stimuli on a trial by trial basis.

After passing through the filters and pellicle, the light was decollimated to transilluminate an opal glass target, which was viewed by the subject from the other side. The stimulus was a white-appearing, circular target subtending a visual angle of 21 minutes. By displacing the opal glass beyond the focal length of the decollimating lens, the stimulus appeared homogeneous, with luminance differences around the circumference and center, as assessed with a Pritchard photometer (Photo Research Corp.), of less than 10 per cent.

The subject was seated in a light-tight booth and foveally viewed the stimulus binocularly from a distance of 48.5

cm. Stable head position was maintained with the aid of an adjustable chin and forehead rest and the stimulus was on the same horizontal plane as the subject's eyes. Four dim red fixation lights (LEDs) surrounded the target in an incomplete cross pattern.

red fixation lights  stimulus target

Daily calibration of the light output of the glow-modulator tube, at the start of each testing day, was done to ensure that the light output was the same on each day. This calibration was achieved by the use of the Pritchard photometer, which was modified to convert luminous flux, in the plane of the stimulus target, into voltage output on an oscilloscope display. Before each testing day, a 0.41 log unit neutral density filter and low density watch glasses were placed in the common filter box until the photometer output was lowered to 0.8 volts, yielding a luminance of approximately 1.5 mL. The calibration procedure was repeated at the start of each testing day. Changes were found to be negligible over the course of the testing, i.e., not more than 0.04 log units.

The presentation of all stimuli, including their duration as well as their sequences, and the presentation of the warning click and interval tones, were controlled by a nine-channel multivibrator timer (Logical Instruments Co.) with an indeterminancy of 1 part in 10,000 and by a Digital Time Generator, Model TG-760 (Analog and Digital Instruments Co.) with an accuracy of plus or minus 0.005 ms. These timers' stimulus channels delivered logic pulses to a gating device that activated and deactivated the glow modulator tube for the preset durations and intervals.

A Grasson Stadler Noise Generator (Model 901B) was used to generate the white noise background. It has a 70-10,000 Hz range with total harmonic distortion of less than 1% at mid-range. The output of the noise generator was amplified by a Bogen power amplifier presented via a loudspeaker in the subject booth at 70 dB SPL(A) as measured by a Quest Electronics Sound Level Meter.

The amplitude of the 500 ms warning click and 500 ms interval tones in the three-interval forced-choice procedure was controlled by a model DC-IMC attenuator (Hewlett-Packard) connected to a model LA 750 speaker (Lafayette Instruments) in the subject booth. The frequency of the tones was controlled by a model 200CD wide range oscillator (Hewlett-Packard) set at 275 Hz.

The duration of the feedback tones, provided in the threshold and in the two-pulse discrimination threshold tasks, lasted only as long as the subject pressed the correct response button. The frequency of the feedback tone was 1000 Hz. All auditory stimuli were well above the hearing threshold and well below a painful intensity level. The click and tones were audible under the noise background conditions in the experiment.

Procedure

Each subject was tested on three separate days, with a minimum of two days but no more than one week separating each testing session. Subjects who regularly ate breakfast were instructed to have a very light breakfast, and all subjects were instructed not to drink tea, or coffee or cocoa, or have any substance containing caffeine on the day of testing. All testing was done in the morning so as not to be confounded by diurnal variations (Humphreys, Revelle, Simon, & Gilliland, 1980).

Data Collection. The experiment was conducted in a double-blind fashion. For each subject, the experimenter calculated the amount of caffeine to be contained in each of three capsules corresponding to each of the three doses used in the experiment. The three doses were 0.25 mg/Kg; 2.0 mg/Kg; and 4.0 mg/Kg. The capsules were prepared by the pharmacy at the New

York Psychiatric Institute. All capsules weighed 500 mg. The total amount of caffeine present in each capsule was always less than 500 mg, the remainder of the capsule was filled with dextrose so that they all weighed and looked the same. Since the experiment was a double-blind study, a third person determined the order of dosage presentation for each subject. The capsules were put in unmarked envelopes labelled Day 1; Day 2; and Day 3, and then given to the experimenter. The third person kept track of the order of presentation and made sure that order was counterbalanced across subjects, as instructed by the experimenter. With 12 subjects, there were two complete counterbalanced runs; that is, across subjects, each of the possible combinations of order of doses was presented twice. The mean low dose was 18.1 mg (SD=1.8), the mean medium dose was 144.8 mg (SD=14.5), and the mean high dose was 289.6 mg (SD=28.8). The strongest dose was equivalent to about three cups of strongly-brewed coffee.

As soon as the subject arrived for testing, he was given the envelope designated for that day's testing and was also supplied with a glass of water with which to swallow the capsule and the experimenter watched him do so. Twenty-five minutes later, the subject was placed in a light-tight testing booth and left there for another five minutes to dark adapt. During those twenty-five minutes prior to being placed in the

booth, the experimenter explained to the subject, or reviewed with him (depending on the day of testing) the tasks to be performed. Thus testing began 30 minutes after the ingestion of caffeine, the time that reports indicate caffeine begins to take effect (Robertson, Frolich, Carr, Watson, Hollifield, Shand, & Oates, 1978).

The subject performed three different tasks in a fixed, prearranged order; first task, to determine his absolute visual detection threshold; second task, to determine his two-pulse discrimination threshold; and the third task, to provide a measure of vigilance. For all three tasks, the subject was required to place his head in a chin and forehead rest and to remove it only for brief scheduled rest periods.

The stimulus consisted of two 1-ms pulses of light equal in energy and separated by 5-ms of darkness, thus measuring 7-ms in total duration. Although the stimulus consisted of two separate pulses of light, the separation between them was so brief as to be fully integrated and perceived as a single flash of light (Kietzman, & Gillam, (1972).

In each of the threshold tasks--the absolute and the two-pulse discrimination thresholds--a trial was preceded by an auditory warning sound (click) lasting 500-ms in duration. In the absolute visual threshold task, the click was followed, one

second later, by three tones presented in succession, signaling three temporal intervals. Each tone was 500-ms in duration and its offset was separated by 500-ms from the onset of the following one. The tones, served as indicators for the subject as to when to look for the stimulus to be detected. The light stimulus to be detected was presented during one of these tone intervals and was presented 100-ms after the onset of the tone. The other two tones did not include a light stimulus. The temporal interval during which the light stimulus was presented was varied from trial to trial. The subject indicated which interval contained the stimulus by pressing the appropriate button corresponding to that interval, located on a board on which the subject's right hand rested.

The presentation of stimuli in the threshold task involved a modified version of the adaptive blocked up-down, three-interval, forced-choice procedure (BUDTIF; Bruder, Spring, Yozawitz, & Sutton, 1980; Mannuzza, Spring, Gottlieb, & Kietzman, 1980), a method that arrives at the 67% level of detection (50% when corrected for chance). In this modification, only a descending staircase was used, and in the first few trials intensity was not blocked but was changed on a trial by trial basis. Since these stimuli were easily detected (because intensity was high since the subject threshold level was unknown and so that there would be room to descend in the staircase),

they were reduced by relatively large steps (0.2 log units). Stimulus intensity was continually reduced from trial to trial until the subject finally made an error in detection, i.e., he incorrectly identified the temporal interval which contained the stimulus. At that point, the procedure was switched to a blocked form of three trials per block, where, as before, each trial consisted of the three temporal intervals. Stimulus intensity was then changed only after responses to a complete block of three trials and remained constant for each of the three trials making up the block. The decision as to the direction of the intensity change for the following block depended on the outcome on all three trials. Any change in intensity could occur only after responses to the three trials in the previous block in the following way: If all three trials were correctly detected, then the intensity for the next block was reduced by 0.2 log unit steps. If two out of three trials were correctly detected, the intensity that was presented in the next block remained the same. If one or none of the three trials was correctly detected, then the intensity was increased by 0.2 log units). This blocked process continued until an intensity value was repeated in three blocks, i. e., there were three blocks with the same intensity, though not necessarily consecutively. At that point the blocked presentation was continued but was changed to 0.1 log unit steps and remained so until there were three blocks repeated at the

same intensity, again, not necessarily consecutively. That last repeated value was then used as the subject's threshold value. Feedback was provided to the subject after every trial by presenting a beep if he was correct. Feedback was provided in order to obtain a more accurate indicator of the subject's threshold in order to avoid obtaining an artificially high threshold level. This procedure provided a detection threshold estimate in 5 to 10 minutes and was presented with one or no rest periods.

Instructions to the subject for the absolute threshold task were given out loud as follows:

> A click will signal the onset of each. Trials are presented in fairly quick succession. About one second after the click you will hear three tones presented in succession. During one of these tones, and only during one of them, a light comes on. Some lights are bright and some lights so dim you may not even see them. You are to indicate, by pressing the appropriate button, which interval you thought contained the stimulus. You are to wait until all three intervals have been presented before you make your decision, but you are to respond as soon as the third interval was presented. If you are correct, you will hear a

beep, if you are wrong you will not hear a beep. Some of the stimuli will be easily detected and others will be difficult to detect, sometimes, so difficult, that you may not see them. Regardless of the case, you are to respond. If you were not sure as to which interval contained the light, or thought you did not see any light, you are to give it your best guess and respond nonetheless by pressing the button corresponding to the interval you guessed. You will be given a break when the task is over.

The subject was then given a few practice trials to determine whether he understood the task, before actual testing commenced.

For the remaining two tasks--two-pulse discrimination and two-pulse vigilance--the experimenter increased the threshold luminous energy by 1.0 log unit so that detection of the stimuli was not a problem. The mean luminous energy of the stimuli, across the group, for both the discrimination and the vigilance procedures was 2.68 mL ms (SD=1.14) for the low dose; 2.09 (SD=0.73) for the medium dose; and 1.97 (SD=1.17) for the high dose. These minimal differences were insignificant.

In the two-pulse discrimination task, as in the absolute threshold task, the onset of each trial was also signalled by a click. However, in this task the tone signaling the three temporal intervals were omitted since the stimuli were all above the absolute luminance threshold and, therefore, readily visible. Each click was then followed after one second by three light stimuli, two identical ones ("standard" stimuli) and a third, different from the other two ("different" stimulus). Each of the two identical stimuli consisted of two 1-ms stimuli separated by a 5-ms dark interval, as in the absolute threshold procedure. The third stimulus also consisted of two 1-ms stimuli, but they were separated in time by a variable dark interval, based on the subject's previous responses.

As in the absolute threshold task, the stimuli were also presented in a modified version of the BUDTIF procedure. Again, a descending staircase was used. In the initial trials, the interpulse interval of the different stimulus was large enough (usually 135-ms was the starting point) for the subject to discriminate it as being different from the other two standard stimuli. The subject's task was to detect the different stimulus and to indicate its position by pressing the appropriate button. The temporal position of the different stimulus was randomly varied from trial to trial. That

is, the subject had to indicate whether the different stimulus was the first, the second, or the third of the three light stimuli presented. Feedback, via a beep, was provided after each trial as to the correctness of the response. Feedback was provided so that an accurate estimate of the two-pulse discrimination threshold be obtained. This was necessary so that the subject will be able to use all response categories in the next task, the vigilance task. If the two-pulse discrimination threshold value is artificially high the subject will have no difficulty in identifying the "different stimulus" (see below) in the vigilance task and thus always be sure of his decision and, therefore, no ROC curves will be obtained.

The interpulse interval of the different stimulus was reduced by 10 msec until the subject could no longer correctly discriminate above a chance level between the three stimuli. From then on, that particular interpulse interval value was then presented in the blocked form. That is, the different stimulus had the same interpulse interval for the three trials compromising a block. Subsequently, the interpulse intervals were changed in 5-msec steps; the decision to increase or decrease the interpulse interval was based on the same stepping rules as were used in the absolute threshold task.

Unlike the absolute threshold task, the discrimination task was presented under both a noise background

condition (about 70 dB above 0.0002 dynes/cm), and a no-noise condition. The presentation of noise or no-noise conditions were randomized from block to block. In the noise background condition, each trial was accompanied by the white noise background from the onset of the click until the subject responded. The noise was not presented during the intertrial intervals in the given block, but was presented during each trial of a given noise block. The interpulse intervals within each of the noise or no-noise background conditions were kept separate. That is, although backgrounds were randomized, the interpulse interval in each subsequent block was based on the subject's responses in the previous block under the same background condition. This was done to reduce the possibility of habituation to the noise. However, the initial trial-by-trial stepping continued until the subject was no longer correct, and was separately done for each background condition. That is, trial-by-trial testing was completed for one background condition before it was done for the other background condition. The randomized order of the noise and no-noise blocks, which followed, was determined by a third person, for each subject so that each dose combination was presented with both sequences of noise and no-noise backgrounds. That noise/no-noise sequence was kept constant for each of the three testing days, with dose changing from day to day. For example, if in a given dose

combination, testing began with a noise block before no-noise on each of the three days, the next time that same combination was presented, testing began with the no-noise block on each of the three days. Testing continued until one interpulse interval value was repeated in three blocks for each of the two backgrounds. Those values were then used as the two-pulse discrimination threshold values for the respective backgrounds. The subject was given a break every ten blocks. The two-pulse discrimination threshold task took about 20 to 30 minutes to complete.

Instructions to the subject for the two-pulse discrimination procedure were read out loud as follows:

> As in the previous task, a click will signal the onset of a trial. However, this time you will see three bright lights presented in succession, beginning one second after the click and separated from each other by one second. Two of these lights will be the same "standard lights in every trial throughout the procedure. Also in each trial, there will be a third stimulus that is different from the other two, and may also be different from trial to trial. You are to indicate, by pressing the appropriate button,

which of the three light stimuli was the different one. You are to wait until all three lights have been presented and then respond immediately. If you are correct, you will hear a beep. Some of the time it will be easy to discriminate the different pulse from the other two, while at other times it may be difficult, and still at other times you may not even notice a difference. Nonetheless, you are to respond on every given trial. That is, on those trials that you are not sure, or when all the stimuli appeared the same, you are still to respond by giving your best guess. Some of the trials will be accompanied by a loud noise background, like the roar of a crowd in a stadium, while other trials will not. Do not worry about the noise, just go on and do the task as instructed. There will be rest periods, at which time you will be allowed to sit back and relax and at the end of the rest, you will again position your head in the chin rest and will continue. You will be given some practice trials.

For the third task, a vigilance procedure, the

stimuli were presented in 12 blocks of 60 trials each for a total of 720 trials per session, which lasted for a total of one hour. Blocks under no-noise and noise backgrounds were counterbalanced in an ABBA design throughout the session, beginning with the same background as was presented in the first block of the discrimination procedure. This counterbalancing provided for analysis of the vigilance data by two temporal divisions, or periods (lasting 30 minutes each), for which performance could be compared and thus period serving as an additional independent variable. The subject did not have a rest period after 30 minutes of testing, but rather had a very brief rest after each block, just enough time to move his head around so that back and neck discomfort from staying on the chin rest for a long time would not interfere with performance. Longer rest periods would have made this something other than a vigilance task, which, by definition, tests attention or performance over continuous and prolonged periods of time.

In the vigilance task, a "click" signaled the onset of a block of 60 stimuli, one stimulus per trial, which were presented one every five seconds. Eighty per cent of these stimuli (or 48 stimuli) were the same standard stimulus in both the noise and no-noise background conditions, while 20% (12 stimuli) were different from the standard but not from each other, within each of the respective background conditions. This

procedure followed the usual vigilance paradigm where it is required to detect a rare event from other events. As before, the standard stimulus was the two 1-msec pulses of light separated by a 5-msec dark interval (the 1-5-1 msec stimulus). The standard stimulus served as the non-signal stimulus in the present task. The different, rare stimulus, which served as the signal stimulus, was not the same for the noise and no-noise conditions. The two pulse discrimination threshold values obtained for the noise and no-noise background conditions were used for the corresponding vigilance conditions. In the vigilance task, the noise remained on throughout an entire block of 60 trials. It was important to use the discrimination threshold stimuli as the signal stimuli, because signal detection theory was to be used in the analysis of the data. Had the rare signal stimulus been too easily discriminated from the standard stimulus, there would have been no uncertainty in the subject's performance and false alarm responses would not have been obtained.

The subject's task was to verbally identify each stimulus presented as being either a standard stimulus or a signal stimulus. He then had to rate the confidence of his decision with one of three ratings: "one" meaning very sure; "two" meaning somewhat sure; or "three" meaning a pure guess. This procedure yielded six categories of responding for either

the different or standard stimulus as shown below.

Response	Rating		
Different	one	two	three
Standard	one	two	three

Feedback was not provided since it would have tended to keep the subject more alert by allowing him to keep track of his responses, and thus defeating the purpose of the vigilance task. The subject was given several practice trials.

Prior to each of the first noise and no-noise blocks, the subject was presented with five standard stimuli followed by one different stimulus, for the respective block. The subject was told that these six stimuli indicated what was meant by standard and signal stimuli.

Instructions to the subject for the vigilance task were as follows:

> A click, this time, will signal the onset of a block of stimuli and not the onset of a trial. This task is different from the other two in that you will not choose from among three intervals or temporal positions, but you are to respond after each and every stimulus. A block will consist of many stimuli presented in succession, several

second apart. Most of the stimuli will be the same standard stimuli, but a small percentage will be different. As in the discrimination task, background noise will be on during some of the blocks, but not during others, but when it is on it will be for an entire block. You are not to worry about the noise but are to go on with the task as instructed. You are to indicate, verbally, after each stimulus whether you thought it was the "standard" or the "different" stimulus as soon as the stimulus is presented. You are also to rate your confidence in what you had just said as follows: say "one" if you are very sure in the decision you had just made; say "two" if you were somewhat sure; and say "three" if you had made a pure guess. Make sure that you use all the categories of responding and ratings, since, as you know, we are never sure 100% of the time, nor are we guessing 100% of the time. At the end of a block of trials you will be given a chance to move your head about for a few seconds.

After the vigilance task and a rest period, a final absolute threshold was again obtained, usually this was

done about two hours after the initial threshold had been taken. The two absolute thresholds--the initial and final--enabled a comparison of the effects of caffeine on absolute threshold within each dose as well as across doses.

Data Analysis. Both absolute visual luminance threshold (first task) and the two-pulse discrimination threshold (second task) were obtained by a three-alternative forced-choice procedure, which is a criterion-free measure based on Campbell's (1963) and Corwin, Kintz, & Beaty's (1979) procedures. The subject is forced to make a decision and cannot choose to refrain from responding. He must choose one of x alternatives given to him, none of which include a response of "I do not know," or "I am not sure." Thus the subject is not given the opportunity to adopt a criterion of willingness to respond, i. e., there is no response bias measure to be obtained, and the resulting threshold is a truer reflection of the subject's sensitivity than that obtained by other psychophysical methods. However, this procedure does not control a subject's motivation.

In the third task, the vigilance task, response bias was measured (by beta and arcsin false alarm measures-see last three pages of this section). In addition to measuring response bias, a measure of sensitivity (d'e) was also obtained. This was achieved by the use of a six-point rating scale and

fitting the resultant data into Receiver Operating Characteristics (ROC) curves. All the ROC curves were generated according to a description by McNicol (1972, chapter 5).

Rating ROC curves are plots of cumulative responses to signal stimuli plotted against cumulative responses to noise stimuli in Z-scores. In the present experiment, these were the two-pulse discrimination threshold value and the 1-5-1 msec stimulus, respectively. The six response categories ranging from a response of "different, very sure," to a response of "standard, very sure," yielded an ROC curve with a maximum of five points. Response categories left blank were combined with adjacent categories, which yielded ROC curves of less than five points (the minimum of obtained points was three). An ROC curve was obtained under each of the conditions of dose, noise and period, for each subject, resulting in 12 ROC curves per subject.

The shape of the ROC curves was determined by the use of a mutual regression equation (Grice, 1966) that bypasses the inadequacies of the least-squares method of curve fitting, which is the usual method for curve-fitting signal detection data (McNicol, 1972). The mutual regression equation, which, through the use of derived data points, produces the best fitting straight-line fit, is given as follows:

$$Y = sy/sx \cdot X + My - (sy/sx \cdot Mx) \qquad (1)$$

where:

Y = the Z-score of the responses to the signal trials (the two-pulse discrimination trials)

sy = the standard deviation of the signal trials

sx = the standard deviation of the standard (1-5-1 msec) trials

sy/sx = the slope of the ROC curve

X = the Z-score of a standard trial

My = the mean Z-scores of the responses to the signal trials

Mx = the mean of the standard trials

The individual data points fitted by the above formula proved to be fairly well represented by a straight line. However, many of the slopes of the ROC curves deviated from unity (obvious by inspection), violating a basic assumption of signal detection theory, that the shapes of the underlying frequency distributions of the signal and nonsignal conditions are equal in variance.

The most preferred parametric measure of sensitivity in signal detection theory is d', which is used when

the data show that the basic assumptions that underlie the theory have been met. The present data generated straight-line ROC curves, which indicate that the underlying distributions were normal in shape. However, the slopes of the straight lines were both greater and less than 1.0, indicating that the equal-variance assumption was not valid. Therefore, an alternative parametric measure of sensitivity, d'e, was used for data analysis.

The d'e measure is defined as twice the value of the Z-score to the signal or noise stimuli, ignoring sign, at the point where the ROC curve intersects the negative diagonal (see Figure 2). The d'e measure is not as affected by changes in the slope of the ROC curve as is the d' measure. The d'e is calculated from the mutual regression equation as follows (Grice, 1966; McNicol, 1972):

$$d'e = -2b/(1+a) \qquad (2)$$

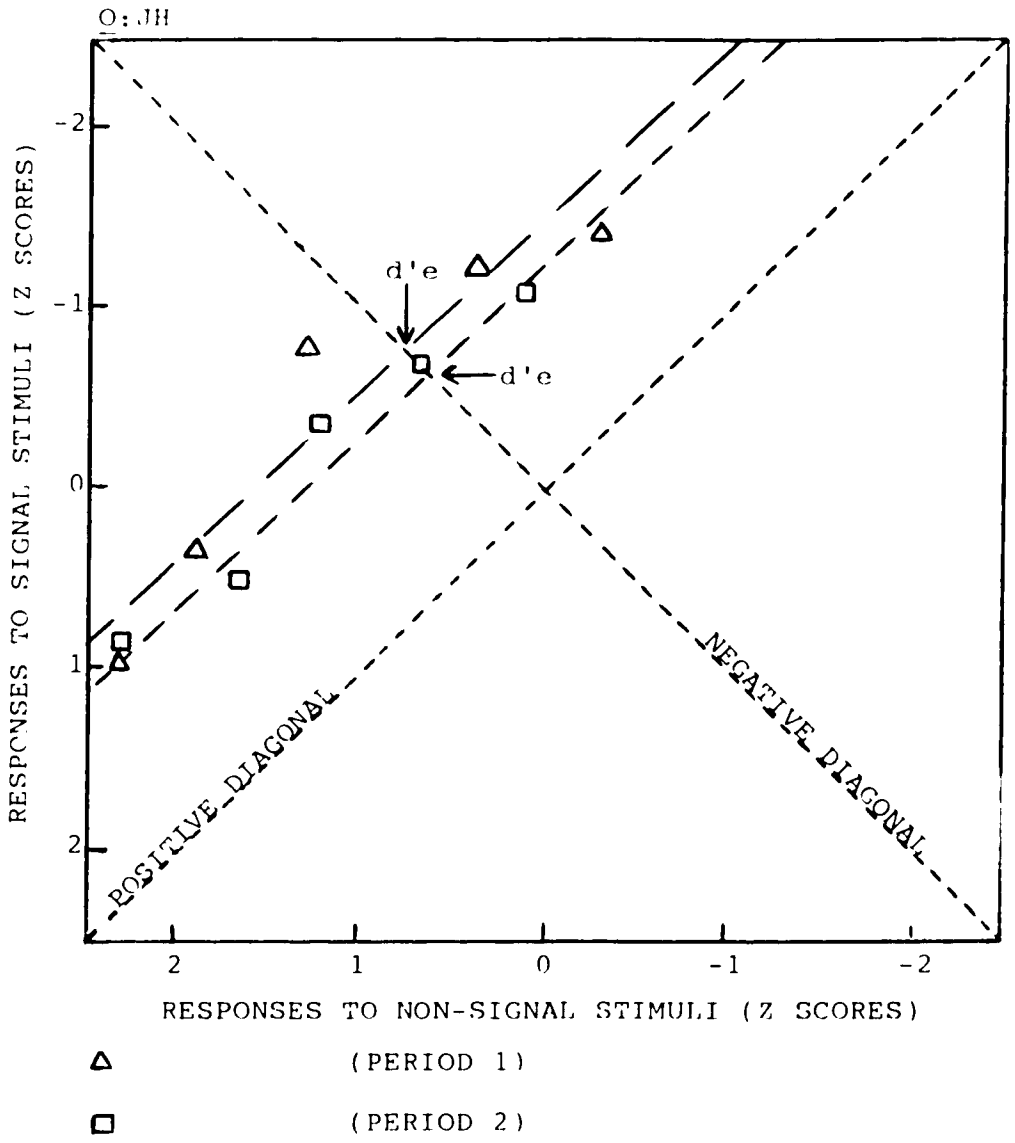
where:

b = the Y-intercept of the ROC curve [My - (sy/sx · Mx), from equation 1]

a = the slope of the ROC curve (sy/sx, from equation 1)

Figure 2 is illustrative data of one subject's ROC curves under

Figure 2. Rating ROC curves for one subject. Condition: Low dose, no-noise background.


O: JH
RESPONSES TO SIGNAL STIMULI (Z SCORES)
-2
-1
0
1
2
d'e
d'e
POSITIVE DIAGONAL
NEGATIVE DIAGONAL
2
1
0
-1
-2
RESPONSES TO NON-SIGNAL STIMULI (Z SCORES)
(PERIOD 1)
(PERIOD 2)

the low dose of caffeine with a no-noise background condition.

In the vigilance task, two measures of response bias were derived: log beta and arcsin transform of the false alarms. Log beta takes into account both hits and false alarms, whereas the arcsin transform takes into account only the false alarms. A hit is achieved when the subjects says that a signal stimulus was presented when indeed it was a signal stimulus. A false alarm is when the subject says that a signal stimulus was presented when, in fact, it was a nonsignal stimulus. There has been some question as to the validity of log beta, although it is the traditional response bias measure in Signal Detection Theory. Mackworth and Taylor (1963) point out that serious violations of the statistical decision theory can and do occur in vigilance studies that may preclude the use of the beta statistic. Swets and Kristofferson (1970) point out that often, in vigilance, ROC curves may be asymmetrical and it will, therefore, be necessary to use less biased variants of criterion. Therefore, it was felt that it would be very useful to look at the false alarms without the hits, that is, to look at errors of "commission," where the subject is more likely to report a signal as present when he is unsure of its presence.

Beta is the ratio of height of the underlying signal and non-signal distributions at a given criterion point. However, because the distribution of beta scores is often skewed

(McNicol, 1972), the beta scores were converted to log beta scores to reduce possible skewness. Log beta, is a parametric measure of response bias. Beta was calculated at each of the points on each ROC curve and, as with d'e, adjacent empty-celled categories (when the subject did not use some of the rating categories) were combined. The equation for beta is as follows:

$$\text{Beta} = Ys/Yn \qquad (3)$$

where:

Ys = the ordinate of the signal stimulus distribution for the given response category

Yn = the ordinate of the nonsignal distribution for the given response category

The logarithm for the first data point on the ROC curve is known as log beta 1, the second point is known as log beta 2, etc.

In the event that a subject is unbiased in his responses, he obtains what is known as optimal beta (the point at which he did not show any bias), which can be calculated using a priori probabilities of signal and non-signal stimuli. Optimal beta is calculated as follows:

$$\text{Optimum beta} = P(N)/P(S) \qquad (4)$$

where:

P(N) = per cent non signal trials

P(S) = per cent signal trials

In the present experiment P(N) was 80% and P(S) was 20%. Thus, 80/20 = 4.0. The common log of 4.0 is 0.6021 [McNicol, (1972) points out that common logs are easier to work with and better for graphical representation than are natural logs], which is optimal beta. Log beta values that are less than optimal beta are indicative of a lax criterion, while log beta values that are greater than optimal beta are indicative of a strict criterion.

The arcsin false alarm measure of response bias was obtained by calculating the cumulative proportions of the subject's responding "different," (that is, signal) to the non-signal stimulus and converting that proportion score into an arcsin transform score (Cohen, 1988; Kraemer & Thieman, 1987) in order to rule out skewness.

The subject was informed that the "different" stimuli comprise a relatively small percentage" of the stimuli, while the remainder of the trials were "standard" stimuli. In order to avoid an artificial build-up of hits and false alarms (many responses of "different" at the end of the session) the subject was never told the actual percentages of the different

types of trials.

Results

The following description of results of the three tasks are ordered according to the extent that significant findings were obtained. Specifically, vigilance, two-pulse discrimination, and absolute threshold data are presented, in that order. Finally, some comparisons across tasks are presented.

Vigilance Data

A separate 3 x 2 x 2 analysis of variance (ANOVA) with repeated measures (BMDP2V) comparing dose, noise and period, was performed on each of the three dependent variables of the vigilance task--d'e (the sensitivity measure), and log beta and arcsin false alarm rate (the two response bias measures).

The d'e measure. Table 1 lists the cell means and standard deviations of the d'e for the twelve separate condition combinations. Inspection of this table facilitates the

Table 1

Cell Means and Standard Deviations of d'e for the 12 Subjects Under All Vigilance Conditions

Dose	Period	Background: No-Noise d'e (SD)	Background: Noise d'e (SD)	Difference
Low	1	1.23 (.64)	1.00 (.64)	
	2	0.75 (.45)	0.99 (.70)	
	Difference	0.48	0.01	0.47
Medium	1	0.93 (.74)	0.86 (.43)	
	2	0.96 (.71)	0.84 (.38)	
	Difference	- 0.03	0.02	0.05
High	1	0.91 (.60)	0.77 (.51)	
	2	0.77 (.52)	0.67 (.54)	
	Difference	0.14	0.10	0.04

interpretation of various groupings of data that are presented later in this section.

One of the assumptions of a repeated measures analysis of variance is that of equal variability of difference scores (Kirk, 1982, pp. 254). Therefore, all data with significant F-ratios were subjected to an epsilon transformation, which corrects for the possibility of such violations by correcting the degrees of freedom (Jennings & Wood, 1976) . If the assumption fails, then the epsilon is less than 1.0, and the degrees of freedom are appropriately corrected. If the assumption is met, the epsilon equals 1.0 and the degrees of freedom remain the same and the F-ratio is looked up under the original degrees of freedom and no correction is necessary. Table 2 contains the F-ratios and the degrees of freedom (epsilon corrected degrees of freedom where results are significant) obtained from the ANOVA. As can be seen from the table, three significant effects were obtained: a significant effect for period; a significant interaction for noise and period; and a significant three-way interaction among dose, noise, and period.

Results from the ANOVA lead to several conclusions about the relationship between caffeine, noise, and period, in the vigilance task with respect to sensitivity. As expected, a significant F-ratio for period [$F(1, 11) = 5.50$, $p < .04$)] was

Table 2

ANOVA (3 x 2 x 2) for d'e

Source	Deg of Freedom	F-ratio	Probability
Main Effect			
Dose	(2,22)	0.70	0.48
Noise	(1,11)	0.84	0.38
Period	(1,11)	5.50	0.04**
Interactions			
Dose-Noise	(2,22)	0.40	0.68
Dose-Period	(2,22)	2.50	0.10
Noise-Period	(1,11)	9.08	0.01**
Dose-Noise-Period	(2,18)*	3.94	0.03**

* Corrected degrees of freedom are presented where epsilon corrections were applied

** Significant, $p < .05$

obtained and indicates that performance was better in the first period (first half hour) in the vigilance task than in the second period (second half hour), as indicated by a higher d'e for the first period in five out of the six comparisons in Table 1. Table 3 is a composite of Table 1 and in it are presented the mean d'e for period 1 and period 2. These data are consistent with the literature on vigilance, in which performance declines over time.

Contrary to expectations, dose alone (caffeine condition) did not affect performance, as the main effect of dose was not significant. In the present experiment, it may be concluded that caffeine alone did not significantly affect sensitivity as a function of vigilance, since the d'es for the medium and high doses of caffeine did not differ significantly from the low dose, which was essentially a placebo (Table 3b).

Like the caffeine condition, there was no statistically significant main effect of noise background conditions (Table 3c).

The results of the ANOVA for vigilance indicated that in addition to the significant main effect of period, there were two significant interactions. Data from Table 1 were regrouped to facilitate interpretation of these interactions.

Table 4 compares performance in each of the two periods under no-noise and noise background conditions, i.e., it

Table 3

(a) Mean d'e and Standard Deviations for Period (Combined across dose and background)

	Period 1	Period 2
$\bar{x}$	0.95	0.83
SD	0.15	0.12

(b) Mean d'e and Standard Deviations for Dose (Combined across period and background)

	Low Dose	Medium Dose	High Dose
$\bar{x}$	0.99	0.90	0.78
SD	0.20	0.05	0.10

(c) Mean d'e and Standard Deviations for Background (Combined across dose and period)

	No-Noise	Noise
$\bar{x}$	0.93	0.86
SD	0.17	0.13

Table 4

Comparison of Background and Period (Mean d'e) (Combined across dose)

	Background		
Period	No-Noise	Noise	Difference
1	1.02	0.88	0.14
2	0.83	0.83	0.00
Difference	0.19	0.05	

portrays the significant F-ratio for the noise-period interaction [F (1, 11) = 9.08, p < .01]. The difference in d'e between period 1 and period 2, by inspection, was greater under the no-noise background (0.19) than under the noise background condition (0.05). (See also Figure 3.) Furthermore, the d'e differences between no-noise and noise conditions, by inspection, was greater in period 1 (0.14) than in period 2 (0.00). Inspection of Table 4 indicates that the introduction of noise reduced the period effect, that is, noise moderated the vigilance decrement effect.

What then accounts for the significant three-way interaction among dose, noise, and period [F (2, 18) = 3.94, p , 0.03]? Table 1 and Figure 4 allow a comparison of the difference in d'e performance between period 1 and period 2 under noise and no-noise background conditions for each of the three doses. A comparison, or contrast among treatment means (Winer, 1971, p171-175) was done to determine the basis for the triple interaction. The first comparison was done to determine if there is a two-way interaction between period and noise between the medium and high doses of caffeine. There was no significant interaction between the two doses [F (1, 18) = 0.20, p, 0.05]. Therefore, for the next comparison the medium and high doses of caffeine were averaged (built in as part of the statistical procedure in comparison among treatment means) and

Figure 3. Mean d'es (combined across dose) for no-noise and noise background conditions.

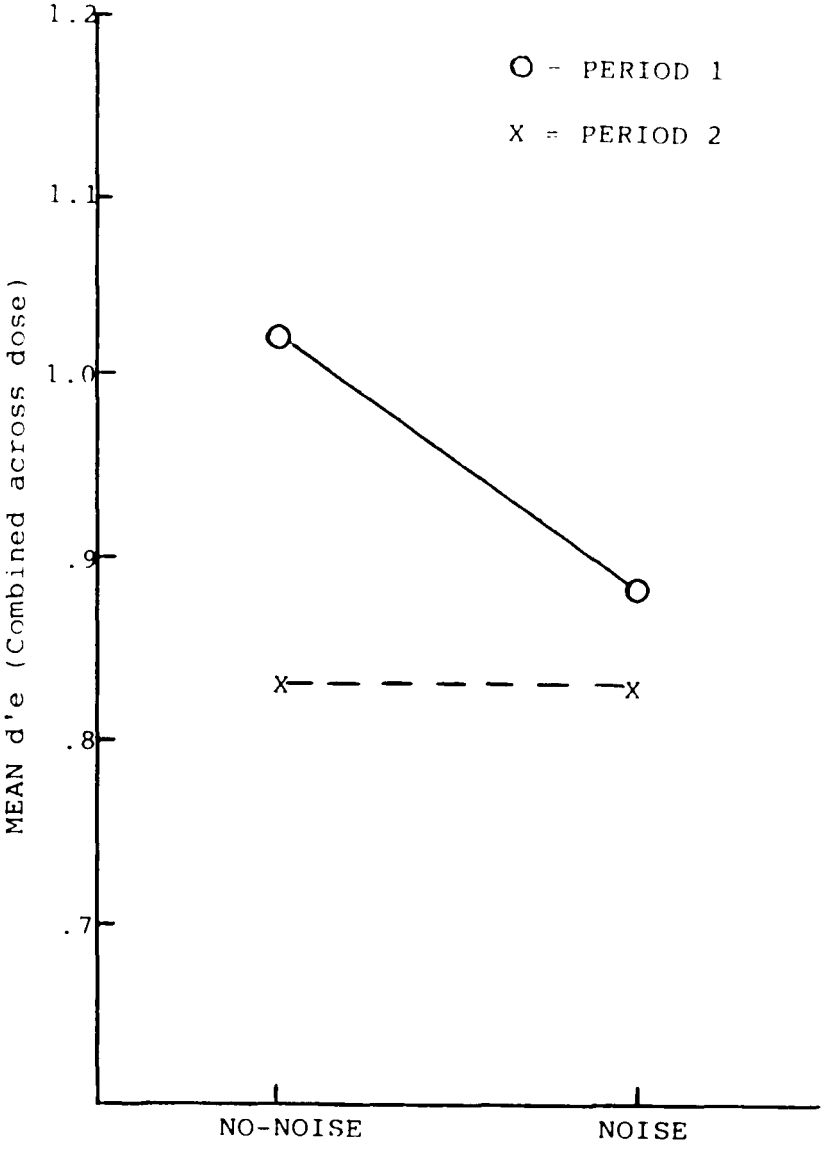
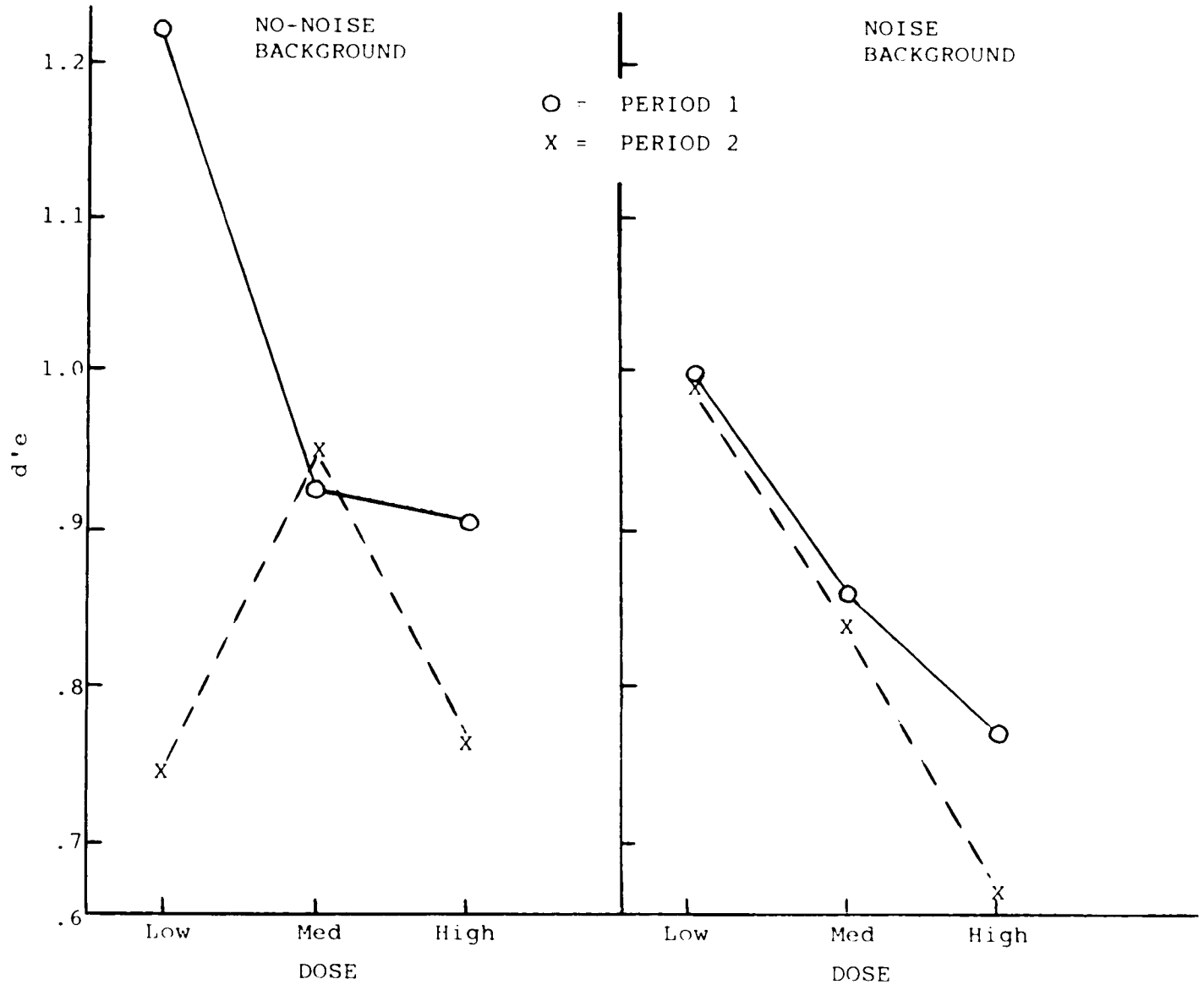

1.2
1.1
1.0
.9
.8
.7
MEAN d'e (Combined across dose)
O - PERIOD 1
X - PERIOD 2
X
X
NO-NOISE
NOISE

Figure 4. Mean d'es for the three doses under each of the no-noise and noise background conditions.

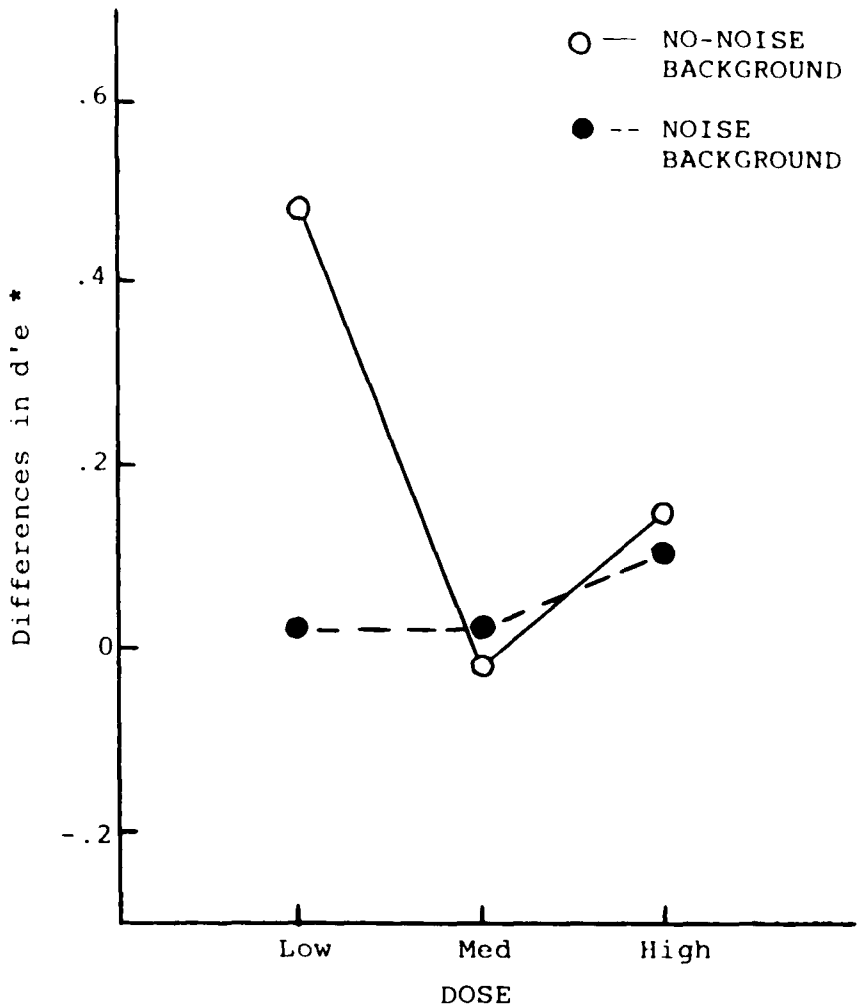

NO-NOISE
BACKGROUND
NOISE
BACKGROUND
O = PERIOD 1
X = PERIOD 2
d'e
1.2
1.1
1.0
.9
.8
.7
.6
Low
Med
High
DOSE
Low
Med
High
DOSE

that average was then compared with the low dose of caffeine. When the low dose was compared with the medium and high doses, a significant difference was obtained [F (1, 18) = 7.68, p, 0.05]. Thus, the period effect and its dependence on noise was primarily present in the low dose and was less evident in the other two doses of caffeine (see Figure 5). The above data, by inspection, indicate that the significant triple interaction is due to differences in period by noise interactions under the doses of caffeine. The two-way interaction between period and noise is more evident under the low dose of caffeine than under the medium dose or the high dose (see difference scores in Table 1).

Response bias: log beta. Table 5 lists the cell means and the standard deviations of log beta for the 12 separate conditions. The data presented in this table represent comparisons at log beta 3. Herskovic (1985) indicated that comparisons at log beta 3 may be the most sensitive since that point is the middle point on the ROC curve and represents cumulative frequencies of all responses of "signal," or, in the present experiment, all responses of "different."

A three-way analysis of variance for repeated measures (see Table 6) yielded no significant F-ratios for any of the independent variables in this experiment, nor any significant interactions. The log beta values obtained ranged

Figure 5. Differences in d'es between periods for the three doses of caffeine for both background conditions.



* Period 1 d'e minus Period 2 d'e

Table 5

Log Beta Cell Means and Standard Deviation for Dose, Background and Period

Dose	Background	Period	Mean Log Beta	SD
Low	No-Noise	1	0.0137	0.6288
		2	0.2266	0.5041
	Noise	1	0.1690	0.5381
		2	0.0877	0.8665
Medium	No-Noise	1	0.1335	0.4998
		2	0.0288	0.7534
	Noise	1	0.1463	0.3470
		2	0.1394	0.4031
High	No-Noise	1	0.0249	0.5157
		2	0.0820	0.4706
	Noise	1	0.1718	0.5236
		2	0.1012	0.4133

Table 6

ANOVA (3 x 2 x 2) for Log Beta

Source	Deg of Freedom	F-ratio	Probability
Main Effect			
Dose	(2,22)	0.30	0.97
Noise	(1,11)	0.82	0.38
Period	(1,11)	0.00	0.98
Interactions			
Dose-Noise	(2,22)	0.23	0.80
Dose-Period	(2,22)	0.53	0.60
Noise-Period	(1,11)	1.13	0.31
Dose-Noise-Period	(2,22)	1.37	0.28

from 0.0137 (for the low dose, no-noise background condition, period 1) to 0.2266 (for the low dose, noise background, period 2) with an overall average of 0.1104 (see Table 5). It is interesting to note, that the largest difference in cell means, although not significant, was in the low dose, no-noise, period 1 vs. period 2 condition, where the biggest difference in d'e was also obtained, although that effect was significant.

Optimum log beta for the present experiment was 0.6021, with higher values indicating relative strictness and lower values relative laxness. Each of the obtained log beta values was separately compared with optimum log beta through the use of 12 individual single-sample t-tests. All of the obtained log beta values, with the exception of one (low dose, noise, period 2) were significantly different ($p < 0.05$) from optimum log beta. Thus, the subjects' log beta response bias measure was lax for all conditions regardless of dose, background condition, or period.

Response bias: arcsin false alarms. In the present study, false alarms were calculated from the subjects' responses of "different" to the standard (non-signal, 1-5-1 ms) stimulus. These proportions (percentages) were converted into arcsin false alarm scores. Table 7 presents the cell means and standard deviations for the 12 conditions.

Results from the 3 x 2 x 2 ANOVA are presented in

Table 7

Arcsin False-Alarm Cell Means and Standard Deviation for Dose Background, and Period

Dose	Background	Period	Arcsin*	SD
Low	No-Noise	1	0.98	0.28
		2	0.99	0.36
	Noise	1	0.96	0.31
		2	0.98	0.36
Medium	No-Noise	1	1.00	0.30
		2	1.00	0.35
	Noise	1	1.06	0.27
		2	1.00	0.32
High	No-Noise	1	1.11	0.25
		2	1.12	0.25
	Noise	1	1.07	0.26
		2	1.21	0.30

* In radians

Table 8. As can be seen, there was only one significant effect,which is the significant main effect for dose. The arcsin transform of the false alarms increased significantly with dose (low dose = 0.98; medium dose = 1.02; high dose = 1.13) indicating that as dose increased subjects became laxer. An attempt was made to determine where the effect of dose was happening by the use of the Newman-Keuls test. There was a significant difference between the high and medium doses of caffeine [t(range) = 5.42, $p < 0.05$] and between the high and low doses of caffeine [t(range) = 7.28, $p < 0.05$]; since the arcsin transform increased with dose it is only natural that the two farthest ranked points be significant if two closer ranked point are significant. There was no significant difference between the medium and low doses of caffeine [t(range) = 1.85, $p > 0.05$]. Therefore, it can be concluded, from this response bias measure that subjects became laxer as dose of caffeine increased from medium to high, but not from the placebo, or low dose, to the medium dose. It is interesting to note, here, that there was no significant main effect for dose for the sensitivity measure, d'e. It is also interesting to note that one response bias measure (log beta) did not yield significant results while the arcsin false alarm measure did. The lack of agreement between the two response bias measures raises the question of why should one response bias measure give such different results than

Table 8

ANOVA (3 x 2 x 2) for Arcsin of the False-Alarms

Source	Deg of Freedom	F-ratio	Probability
Main Effect			
Dose	(2,19)*	4.36	0 .03**
Noise	(1,11)	2.15	0.17
Period	(1,11)	0.26	0.62
Interactions			
Dose-Noise	(2,22)	1.88	0.18
Dose-Period	(2,22)	1.74	0.20
Noise-Period	(1,11)	1.21	0.29
Dose-Noise-Period	(2,22)	2.02	0.16

* Corrected degrees of freedom are presented where epsilon correction was applied

** Significant, $p < .05$

another measure.

Discrimination Data

Table 9 lists the two pulse discrimination thresholds for each of the 12 subjects under each of the three doses for each of the two background conditions (no-noise and noise). The average two-pulse discrimination thresholds for all the three dose levels of caffeine were higher (less sensitive) under the no-noise background condition than under the noise background condition. However, a 3 x 2 ANOVA indicated that there were no significant main effects nor was there a significant interaction (Table 10). Thus, neither caffeine, nor noise had a significant effect on the two-pulse discrimination threshold performance.

Threshold Data

Table 11 lists the group mean thresholds and standard deviations in terms of luminous energy (mL ms) for the three doses of caffeine measured at the beginning and at the end of each session. A 2 x 3 ANOVA (order x dose) comparing the order of the thresholds (the first one obtained at the beginning of each session vs the second threshold obtained at the end of

Table 9

Two-Pulse Discrimination Thresholds (ms)

	Dose					
	Low		Medium		High	
	Background		Background		Background	
Subject #	No-Noise	Noise	No-Noise	Noise	No-Noise	Noise
1	60	20	55	55	50	35
2	115	130	150	100	105	85
3	70	60	55	35	65	55
4	30	25	40	40	20	35
5	30	30	30	30	25	25
6	45	50	45	35	50	45
7	50	50	35	50	45	50
8	40	40	50	65	45	45
9	65	55	30	30	25	30
10	60	60	105	90	60	55
11	35	35	40	40	50	45
12	55	40	45	55	75	55
Mean:	54.60	49.60	56.70	52.10	51.20	46.70
SD:	23.30	28.60	35.40	22.90	23.70	15.70

Table 10

ANOVA (3 x 2) for Discrimination Thresholds

Source	Deg of Freedom	F-ratio	Probability
Main Effect			
Dose	(2,22)	0.65	p > .05 Not Sig.
Noise	(1,11)	3.71	p > .05 Not Sig.
Interaction			
Dose-Noise	(2,22)	0.00	p > .05 Not Sig.

Table 11

Threshold Energy (mL * msec) Obtained at Beginning and End of Each Session (Dose)

	Low Dose		Medium Dose		High Dose	
Subject #	Beginning	End	Beginning	End	Beginning	End
1	.2872	.2300	.0744	.0492	.1866	.2514
2	.0405	.0963	.0718	.0372	.0577	.0718
3	.2872	.1478	.2300	.1478	.1478	.2300
4	.3811	.2300	.2872	.2872	.1478	.1866
5	.2300	.1866	.1866	.1866	.0963	.1866
6	.1478	.2872	.2300	.1866	.1478	.2872
7	.3811	.2300	.2300	.2872	.3811	.2872
8	.2300	.3811	.2872	.3811	.3811	.2872
9	.2740	.1785	.1821	.2415	.1821	.2415
10	.1182	.1478	.1866	.1182	.1182	.0963
11	.3811	.1866	.2300	.2872	.3811	.2872
12	.3811	.2300	.2300	.3811	.1866	.2300
$\bar{X}$	.2621	.2110	.2022	.2159	.2012	.2202
SD:	.1130	.0736	.0691	.1152	.1148	.0734

each session) for the three doses of caffeine did not yield significant differences for dose [$F(2, 22) = 0.80$, $p > 0.05$], for order [$F(1, 11) = 0.22$, $p > 0.05$), or for the order by dose interaction [$F(2, 22) = 2.56$, $p > 0.05$]. Thus, it can be concluded that caffeine did not affect the luminous energy needed for threshold, i.e., there were no significant threshold differences within a given dose of caffeine comparing initial and final thresholds, or between the three caffeine dosages.

Discrimination and Vigilance

Several analyses were done in an attempt to compare performance between the vigilance and discrimination measures. A direct comparison between the two measures cannot be made for two reasons: (1), the two-pulse discrimination threshold measure was obtained via an adaptive procedure (using stimuli in one sequence that were based on the outcome of responses to stimuli in a previous sequence), while the vigilance data were obtained using pre-set stimuli; (2) the requirements of each task were such that they can only be done in the sequence presented, since each task was somewhat dependent on data obtained in the previous task.

All but one of the d'e values obtained in the vigilance task were less than in the discrimination task. In the

discrimination task, testing was stopped once a 67% performance level was reached and, in a three-alternative forced-choice, this is equivalent to a d' of 1.13 (Hacker & Ratliff, 1979). The d' measure provides the closest possible approximation to d'e and enables a comparison of the sensitivity between the two tasks. Sensitivity in the vigilance task, as measured by d'e, ranged from 0.67 to 1.23 (see Table 1). Multiple t-tests, individually comparing each of the 12 d'e values obtained in the vigilance task with the constant d' of the discrimination task have shown that 5 out of 12 tests were significantly different ($p < 0.05$). Since when doing several t-tests, a certain number of them may come out to be significant by chance, it was decided to "weed" out that possibility by making the significance criterion more stringent and comparing results at the $p < 0.01$ to determine which comparisons remain significant and, therefore, less likely to be due to chance. At $p < 0.01$ none of the t-tests are significant. That is, the d'e obtained in the vigilance task were not significantly different from the d' in the discrimination task, where a more stringent significance level was used.

It may be argued that since there were statistical differences in the vigilance data but not in the discrimination data, no further statistical analyses need be done. However, in the present study it was decided to do one additional analysis

(Spearman-rho rank correlation) to determine whether subjects themselves were consistent in their sensitivity responses from task to task. Thus, every subject was compared to himself with respect to rank.

For the vigilance task, the d'e values were averaged for the two periods because the discrimination task was not divided into periods. Separate rank correlational analyses were done for each of the three doses in the following way: for the discrimination data the two-pulse discrimination threshold value for the noise background was subtracted from that of the no-noise background and these differences were ranked. For the vigilance data, the d'es for the noise background were subtracted from those of the no-noise background, and these differences, too, were ranked. The data were analyzed this way in order to permit an evaluation of whether individual subjects had maintained the same differences in performance between noise and no-noise background conditions from the discrimination task to the vigilance task. These differences constituted the data for the rank correlations.

Table 12 presents the results of the rank correlations. As can be seen from the table, under the low dose there was a significant positive correlation between the discrimination and vigilance data. This positive correlation indicates that those subjects who performed better in the

Table 12

Spearman-Rho Rank Correlations of d'e and Two-Pulse Discrimination Thresholds

Dose	Spearman-Rho r	Probability	
Low	0.73	$p < .02$	Sig.
Medium	0.41	$p > .05$	Not Sig.
High	-0.04	$p > .05$	Not Sig.

no-noise background than in the noise background in the discrimination task, also performed better in the no-noise background in the vigilance task. From this we can conclude that, at least under the low dose of caffeine, subjects were consistent in their performance with regard to the effects of noise versus no-noise background conditions. The rank correlations between discrimination and vigilance performance under the medium and under the high doses of caffeine were not significant.

DISCUSSION

Vigilance

The d'e. In the vigilance task, the finding that performance significantly deteriorated with time (period 1 showing a higher d'e than period 2) is consistent with the literature (see review by Mackworth, 1968; Parasuraman & Davies, 1977). When attention has to be sustained for prolonged periods, there generally is a decrement in performance over time. This is not surprising since it is difficult to maintain heightened attention for long periods of time, especially when doing boring, monotonous tasks.

In the present study, the decrement in performance over time was a decrement in sensitivity and not in response bias. Specifically, the d'e for period 1 was significantly higher than for period 2 in the vigilance task while neither of the response bias measures (log beta and the arcsin false alarms) proved to be significantly different over periods. The decline in sensitivity over time can be considered to be a result of a drop in the level of activation. That is, during the first half hour of the vigilance testing the subjects were more

alert and, therefore, more accurate in their identification of the standard stimuli (1-5-1 ms) and the different stimuli (the two-pulse discrimination threshold value stimulus) than they were in the second half hour. However, their decision criterion remained the same, regardless of the deterioration of performance with time. The assumption that the subjects were more alert in the first part of the vigilance task was evidenced by the subjects' reporting that they were finding it difficult to stay awake in the second half of the vigilance task, regardless of the dosage of caffeine that was administered. The fact that the subjects had difficulty in this respect lends further credence to the notion that tasks such as this one that consist of repeated frequent and rare stimuli and require subjects to respond to each stimulus are valid measures of vigilance. There is also evidence demonstrating that procedures that require subjects to respond as instructed to every stimulus, signal or not, yield vigilance data (d' or beta) that differ very slightly or not at all from those requiring responses to signal stimuli only (Guralnick & Harvey, 1970; Parasuraman & Davies, 1975).

The drop of performance sensitivity over time supports both Lindsley's (1960) and Pribram and McGuiness' (1975) contention of a change in the level of activation, or tonic readiness to respond, over time. However, if one

interprets Pribram and McGuiness literally, then a tonic readiness to respond can be viewed as a criterion factor, i.e., as a change in the subject willingness to respond. The present data indicate that the deterioration of performance over time is one in sensitivity, not in criterion.

Guralnick and Harvey (1970) compared response requirements and performance in a visual vigilance task by using three different signal detection procedures. Subjects were divided into three groups as follows: one group was given the standard vigilance task (respond to signal only); the second group was given a binary procedure (respond to every trial as to it being a signal or non-signal); and the third group was given a rating procedure. According to signal detection theory, these methods should yield identical results since sensitivity and response bias can be separated statistically. These investigators found that the d' measure remained constant over time (80 minutes of testing) and was not different for the three different procedures. However, beta increased with time and there was a statistically significant difference between the standard and binary procedures group.

There are several factors that may account for the discrepancy between the present experiment's vigilance results and those of Guralnick and Harvey's. The task required by all subjects by Guralnick and Harvey was essentially an absolute

threshold task. That is, the stimuli were at threshold level and the subjects were required to make decisions as to the presence or absence of a stimulus. The present study employed supra-threshold stimuli and subjects made decisions on the basis of differences between stimuli presented individually, one stimulus per trial. That is, the subjects in this experiment had to "remember" what was meant by the different or signal stimulus, as well as what was meant by the standard stimulus. This latter type of decision, a decision where the standard and different stimuli are not presented simultaneously, or where a reference stimulus is not available, may be more demanding on an individual. In the present experiment, this task was even more demanding because these differences between the two types of stimuli were at threshold level, even though the energy level of the stimuli was suprathreshold.

One other factor to consider, when comparing the present experimental results to those of Guralnick and Harvey is that the present study obtained the well documented vigilance effect of decreasing sensitivity over time (even with a rating procedure) while Guralnick and Harvey did not. However, this discrepancy is not singular. Broadbent and Gregory (1963) also reported that sensitivity was invariant as time increased in the vigilance task. Thus, several researchers have postulated that performance in a vigilance task may be task specific

(Parasuraman, 1976; Swets & Kristofferson, 1970). In tasks that require continuous observation and responding, a decrement in sensitivity over time is more likely to occur than in tasks where stimuli are presented infrequently or where a response is made only to the signal stimuli. With continuous observations there is a drop in perceptual efficiency with time and thus the decrement in performance sensitivity.

Some discrepancies across vigilance studies have been attributed to the event rate of stimulus presentation. Several studies have reported that the higher the event rate, the greater the decrement in sensitivity (Loeb & Binford, 1968; Parasuraman, 1979). Some caution must be taken in interpreting these findings. The definition of what is meant by a high event rate is not consistent among researchers. A high event rate can mean in some studies simply to be a more frequent presentation of stimuli, signal or non signal, while in other studies it may mean an increase in the a priori probabilities of signal presentation. These discrepancies have been found to affect signal detection measures differently, e. g., signal probability seems to affect criterion but not sensitivity (Broadbent & Gregory, 1965; Parasuraman & Davies, 1976).

In the present study, there was a significant interaction effect for period and noise. The difference between period 1 and period 2 was greater for no-noise than for noise

conditions. This indicates that noise reduced or eliminated the vigilance decrement in sensitivity. In that respect, noise acted not so much as an arousing stimulus, but more as one that maintains arousal over time.

The present study also found a three-way interaction for noise-dose-period. This triple interaction indicates that the two-way interaction of noise and period is either greatest or restricted to the low (placebo) dose of caffeine. Thus, the greatest difference (decrement) in performance was between period 1 and period 2 in the low dose under the no-noise background condition. The presence of caffeine and noise interacts to reduce the vigilance decrement in performance over time. This may be taken as evidence that both caffeine and noise, although not necessarily improving or decreasing performance, do help to maintain the initial performance level over prolonged stimulation.

A factor of interest, although not testable in the present study, was the subjects' report on all three days of testing, that they were finding it difficult to stay awake as time went on. Yet the triple interaction for sensitivity among dose, noise, and period seems to indicate that noise and caffeine erased or reduced the period effect, or time decrement. This may be an indication that subjective reports of arousal/activation are not consistent with objective

measurements, which is not inconsistent with the literature.

Response bias. Neither of the response bias measures, log beta and the arcsin false alarms, yielded a change in response bias over time. The typical finding in vigilance experiments using signal detection analyses, is that of an increment in strictness or caution, usually referred to as a criterion increment, over time (Broadbent & Gregory, 1963; Guralnick & Harvey, 1970). Guralnick and Harvey (1970), in their comparison of a vigilance task under three different response requirements, found that beta significantly increased with time. However, the results from the rating procedure differed significantly from the other two procedures, which did not differ from each other. The rating procedure produced the lowest beta (a much higher false alarm rate than the other two procedures), that is, individuals tested with that procedure were laxer.

The finding that the rating procedure is more likely to produce laxer criteria than other traditional methods (Guralnick & Harvey, 1970) may explain why, in the present experiment, subjects adopted a laxer criterion than was optimal. The present result, though, of criterion remaining invariant over time, does not agree with the majority of those in the literature. However, Davies & Parasuraman (1982) found instances of lax criterion associated with less change over time than

instances of strict criterion. Thus, the present data may indicate that the laxer the criterion adopted to begin with, or the laxer the criterion adopted overall, the less likely it is to change with time.

Based on the present experiment and previous research results, the obtained data suggest that not all psychophysical procedures and/or signal detection studies yield comparable results and suggest that particularly the rating procedure may be different from other signal detection procedures.

Although there was no significant vigilance effect for criterion regardless of the measure, there was one difference in findings between the two response measures. While log beta yielded no significant effects at all, the arcsin false alarms yielded a significant main effect for dose. This discrepancy raises an important question. Why should one response bias measure give different results than another measure? It is possible that the difference in the response bias measures may be due to the relationship between response bias and sensitivity. That is, the problem may lie in the assumption of independence between response bias and sensitivity. This leads to the suggestion that future research might profitably analyze both the log beta and the arcsin false alarms measures, and then choose the more stable and independent measure of the

two as the response bias indicator. The present data suggest that one must analyze the data with respect to the assumed independence of response bias and sensitivity measures and not simply assume them to be independent.

There are many different types of response bias measures (Dusoir, 1975), yet it is unclear as to what each measure is tapping. Researchers have tended to treat the different measures as if they were interchangeable, but results of the present study indicate that response bias measures must be investigated more closely, and particularly, that the relationship between them must be empirically established.

Caffeine

The d'e. Caffeine did not significantly affect performance in the vigilance task, over the range of caffeine doses used in this study. Since caffeine is a stimulant, one might have expected the d'e measure of sensitivity to change for the different doses. Because one of the reported effects of caffeine is to increases alertness, one would expect the d'e to be higher as dose increased. However, since no d'e effect for dose was obtained it can be concluded that caffeine did not affect the sensitivity of the subjects in the vigilance task.

Given the literature, it was surprising that

caffeine did not have an effect on sensitivity in the present study. This may be due to variability. In Table 3b, one can see a trend for d'e to decrease with dose. However, the variability (as much as 20% in the low dose) may have masked possible differences between doses. The standard deviations are large when compared to differences between means. This may be due to the dosages selected for testing. The differences between the dosages administered in the present experiment may not have been large enough to obtain a statistically significant sensitivity difference. The dosages in the present experiment were not randomly selected. One reason for their selection was an administrative one. If higher dosages had been used, constant physical monitoring would have been required; also it would have been necessary to have a physician available for intervention, in the event that side effects occurred, which would have been likely with higher dosages. Another reason for the selection of the doses used in the present experiment was the desire to minimize side effects. Yet another reason was the subject population. It was decided to use non- or low-caffeine users so that caffeine tolerance would not have been a confounding factor. Because the behavioral and physiological effects are partially dependent on the tolerance level of caffeine (Childs, 1978) it follows that chronic users need greater doses to achieve or maintain the behavioral effects of caffeine. Since

the subjects were predominantly "abstainers," there was also more of a chance that caffeine would cause side effects at lower doses than for subjects who were not "abstainers."

Another possible reason to account for the lack of significant differences in d'e is the number of subjects tested. The possibility exists that had more subjects been tested, the variability may have been smaller and the trend of decreasing d'es with dose would have been significant. However, the significant changes in d'e over time, and the significant three-way interaction between dose, period, and noise indicate that the present number of subjects may have been sufficient, but that large variability may be inherent for prolonged tasks.

The decreasing trend in d'e with caffeine increments suggests, that for some subjects, caffeine may interfere with, instead of facilitate performance. This possibility may be related to an inverted-U shaped activation curve, i. e., the Yerkes-Dodson Law. If we assume that the subjects tested in the present experiment were at the mid- or upper-range of the excitation/activation continuum and caffeine acted as a stimulant to further increase their activation, this could have resulted in their performance deteriorating with increasing dosage. The inverted-U shaped hypothesis may also explain why these subjects were low caffeine users. These subjects may be avoiding caffeine because they may also be at a

higher level of arousal/activation to begin with, and thus, any additional stimulant can arouse/activate them even more.

The above suggestions are not in conflict with caffeine acting as a stimulant, only in the predicted direction of that effect. That is, the stimulant action of caffeine is not necessarily to increases sensitivity, as originally expected, but to affect sensitivity indirectly through its actions on arousal/activation mechanisms. Caffeine may move a subject further along on an activation continuum, and depending on where the subject's baseline is, performance will shift accordingly. That is, if a subject's baseline level is essentially at a low activation level, performance improves (sensitivity increases). If activation is at some medium level, performance may not change or may decrease somewhat and, if baseline is at a higher level of activation, performance may decrease, as was the trend in the present data.

Noise

As can be seen from the d'e data (Table 2 and Figure 3) there was a significant two-way interaction between noise and period. Other than this interaction (discussed in the vigilance section), there are no noise effects. An almost significant trend can be seen in the results of the ANOVA for

the discrimination data in Tables 9 and 10. The F-ratio for noise, approaches significance (see Table 10). Inspection of Table 9 reveals that performance under noise yielded lower two-pulse discrimination thresholds than under no-noise. Although not significant this may be indicative of a trend, suggesting that noise improved the two-pulse discrimination thresholds. It may be possible that had a louder noise level been used, significant results may have been obtained. This trend, if we may consider it so, is consistent with the view of noise being an arousing stimulus, since the thresholds were better than without noise. However, one must bear in mind that this is only considered a trend and that significant differences were not obtained.

The lack of any clear noise effects, other than the interactions discussed in the vigilance section may not be surprising. Although theories of how noise affects performance abound (Berlyne & Lewis, 1963; Broadbent, 1978; Fisher, 1983; Hamilton, Hockey, & Rejman, 1977; Poulton, 1979) these theories fail to give a cohesive picture of noise's effects on performance. At times, task complexity is offered as a factor in explaining the influence of noise as is noise's interference (masking) with inner speech necessary for some tasks. The masking effect does not necessarily need to be viewed as a separate factor or something other than changing one's arousal

level. These "negative" effects of masking may also serve to change arousal level. This view may be consistent with the Yerkes-Dodson Law. That is, noise may simply be an arousing stimulus in the absence of other auditory stimulation, external or inner, and where the opportunity for masking exists and occurs, that effect may also shift arousal to the point of increase beyond optimum, where performance begins to decrease. That is, in processing where a one stimulus masks or is masked by another stimulus, that event in itself may be seen as a form of overload and, therefore, a shift on the arousal/activation continuum. Naatanen (1973) suggested that although noise may have an irritating and annoying effect, one must not necessarily view that effect as a necessary experience connected with high activation, but that it may be specific to a stimulus situation which may cause one to lose concentration. Easterbrook (1959) suggested that attention may be restricted during arousal conditions. This position may lead to a hypothesis concerning individual differences or preferences. That is, if attention is restricted, it may be focused on some aspects and not others and that focus may at times be dictated by the nature of the task, the complexity of the task, or, it may simply be individually directed by the subject such that subjects may use their own cues for selective attention. If such were the case, subjects could effectively choose to ignore certain cues or stimuli,

while channeling their attention to more demanding cues. This proposed view is consistent with both arousal and masking effects of noise as well as with the concept of adaptation, where subjects may habituate to the noise. This view also suggests that the effects of noise may very well depend on interaction of stimulus dimensions of noise (such as duration and loudness) and experimental design (for example, ongoing vs intermittent noise) and task complexity as well as how important is the noise in the subject's response decision.

Two-Pulse Discrimination

In the two-pulse discrimination data, neither caffeine nor noise had a significant effect on the two-pulse discrimination threshold. It is not surprising that caffeine did not have an effect on the threshold, since it is more likely that caffeine should change the level of tonic activation rather than phasic arousal. It is surprising, however, that noise, an alerting stimulus, presented over trials only, did not have an effect. As an alerting stimulus, it should have changed phasic arousal in this discrimination task. For example, Harper (1979) found that white noise of about 70 dB maximally and reliably (across subjects) improved sensitivity to critical flicker fusion, that is, improved resolution, a measure which may used

as an indicator of arousal. Other values also improving CFF were 100 dB and 40 dB, while values between 40 and 70 and between 70 and 100 dB did not, creating a sawtooth function. Although Harper used a noise level that was used in the present experiment, his study measured sensitivity (d'e) of critical flicker fusion as a function of white noise. The present study kept sensitivity constant (i. e., measured two-pulse thresholds) and did not manipulate white noise levels. Harper's task, although a signal detection one, was not a vigilance task and thus does not provide information on tonic activation, but does for phasic arousal.

As mentioned earlier, the present study did not find noise effects for arousal or activation. A possible explanation for the lack of an effect for activation may be noise's reportedly transient effects, which disappear with prolonged stimulation (Broadbent, 1971).

Threshold

Since caffeine is a stimulant and is known to dilate the pupil, it might be expected to affect the visual threshold, either in a general way through increasing the subjects alertness, or specifically via pupillary effects. However, the present data indicate that caffeine did not affect

the energy needed for threshold. The absence of expected differences from pupillary dilation across doses may be explained by the small stimulus target size (21' visual angle).

Discrimination and Vigilance

The findings of significant differences in the vigilance task but not in the discrimination task indicates that these do not tap parallel processes. These finding support Parasuraman's (1976) view that results of tasks testing activation or arousal are very task "type" specific, especially with respect to monitoring tasks. Thus, the the two tasks, when compared, in the present study need not produce the same results. However, the significant rank correlation coefficient for the low dose, but not for the medium and high doses of caffeine, suggests that subjects displayed some consistency with respect to their individual responses to noise across the two-pulse discrimination and vigilance tasks, but that consistency across tasks was not present with caffeine. This may indicate that caffeine and noise may affect individuals differently in different tasks.

Summary

The present study supported other research in finding a decrement in performance over time in a vigilance (prolonged monitoring) task. It did not find significant effects for either caffeine or noise alone in altering performance. However, caffeine and noise interacted to maintain an initial level of performance and, thus erasing or reducing the vigilance decrement over time. Caffeine did affect response criterion as measured by the arcsin of the false alarms. That is, the subjects became laxer with increasing dose. This result however, was not found with the other criterion measure, log beta. The discrepancy between the two response measures raised questions about the assumptions of signal detection theory.

Recommendations for Future Research

It is suggested that for future research, both more levels and higher levels of the independent variables be used, as well as more subjects. It is also suggested that subjects be equated on several measures of arousal and/or activation, which have been demonstrated to be correlated with the measures to be used in the proposed study. Thus, screening

criteria could be used to equate subjects for their responses to caffeine on some task other than the experimental task. It is also extremely advisable that if arousal or activation are to be examined via performance, that concomitant physiological measures be obtained simultaneously. It is also recommended, that while carrying out studies that purport to test arousal or activation, however they are defined, that subjective responses be obtained in order to correlate them with performance and physiological activity. Future research could also benefit by a closer examination of signal detection theory. That is, why do log beta and the arcsin false-alarms yield different results? Are all the assumptions of signal detection theory met; if not, why not?

APPENDIX I

FILE: IPI .OPS A *** PSYCHIATRIC INSTITUTE ***

I.APPLEBY 10/91 (X56291
PAGE 1 OF 1

SUBSTANCE INTAKE PACKET I (S.I.P.I) ****

(ID. NO.) ____
(STUDY) ____
(CARD NO.)____
(D.O.B.) ____
(1/2) ____
(CP) ____
(RG.I) ____

THIS IS A QUESTIONNAIRE REGARDING YOUR INTAKE OF CAFFEINE-CONTAINING SUBSTANCES. PLEASE ANSWER ALL QUESTIONS BELOW THE DOUBLE DOTTED LINE BELOW. ALL INFORMATION IS CONFIDENTIAL AND ANONYMOUS. IT USUALLY TAKES ABOUT 15 MINUTES TO COMPLETE THE SIXTY QUESTIONS, BUT YOU MAY TAKE AS MUCH TIME AS YOU NEED. IF ANYTHING IS UNCLEAR, PLEASE ASK. IF A QUESTION DOES NOT APPLY TO YOU, PLEASE FILL IN THE SPACE WITH A DASH.

THANK YOU FOR YOUR COOPERATION.

PLEASE FILL IN THE INFORMATION REQUESTED BELOW.

1. TODAY'S DATE__/__/__
2. DATE OF BIRTH __/__/__
3. SEX (CIRCLE) MALE/FEMALE
4. HEIGHT _____FT_____INCHES
5. WEIGHT _____LBS

WE ARE INTERESTED IN YOUR CURRENT CONSUMPTION OF CAFFEINE-CONTAINING SUBSTANCES. PLEASE FILL IN THE CHECKLIST BELOW USING YOUR INTAKE OF THE PAST 4 WEEKS AS A GUIDE:

USE THE SPACE (DAY, WEEK OR MONTH) WHICH WOULD BEST INDICATE YOUR FREQUENCY OF CONSUMPTION. IF YOUR ANSWER IS "NONE" PUT A "0" IN THE SPACE PROVIDED. IF THE QUESTION DOES NOT APPLY TO YOU, PUT A DASH (-) IN SPACE PROVIDED.

A. BEVERAGES

6. HOW MANY CUPS OF COFFEE DO YOU DRINK PER DAY? ______
OR PER WEEK? ______
OR PER MONTH? ______

7. PLEASE SPECIFY HOW MUCH OF EACH TYPE OF COFFEE IN THE SPACE BELOW. CONSIDER PERKED, DRIP, INSTANT, EXPRESSO, CAPPUCHINO, AND DECAFFINATED TYPES OF COFFEE. INCLUDE BOTH HOT AND ICED FORMS. IF YOU HAVE MORE THAN ONE TYPE, PLEASE LIST THEM ALL.

TYPE	NUMBER OF CUPS PER DAY OR WEEK OR MONTH	BRAND

PILOT SITE CODE A *** PSYCHIATRIC INSTITUTE ***

7. HOW MANY CUPS OF TEA DO YOU DRINK PER DAY? _____
OR PER WEEK _____
OR PER MONTH? _____

8. PLEASE SPECIFY WHICH BRAND BELOW. INCLUDE HOT AND ICED (INCLUDING CANNED) FORMS. INDICATE IF REGULAR TEA, SPECIAL HERBAL TEA, OR DECAFFEINATED TEA, IF KNOWN.

KIND	BRAND	NUMBER OF CUPS PER DAY OR WEEK OR MONTH	SIZE OF CUP/CAN SMALL	REGULAR	LARGE

10. HOW MANY GLASSES OR BOTTLES OF COLA DO YOU DRINK PER DAY? ____
OR PER WEEK? ____
OR PER MONTH? ____

11. PLEASE SPECIFY BOTH BRAND AND USUAL SIZE OF PORTION BELOW. FOR EXAMPLE, SOME OF THE BRAND NAMES ARE COCA-COLA R, PEPSI-COLA R, TAB R, ROYAL CROWN R, SHASTA R, AND R.C. R. PLEASE SPECIFY IF IT IS THE DIET VERSION.

BRAND	NUMBER OF CUPS PER DAY	WEEK	MONTH	SIZE OF GLASS/CAN/BOTTLE SMALL	REGULAR	LARGE

13. DO YOU DRINK ANY OF THE FOLLOWING SOFT DRINKS: ORANGE CRUSH R, MOUNTAIN DEW R, MELLOW YELLOW R, OR DR. PEPPER R?

13. IF SO, PLEASE SPECIFY WHICH BRAND, WHETHER OR NOT IT IS THE REGULAR OR DIET VERSION, AND SIZE OF USUAL PORTION.

BRAND	REG./DIET	NUMBER OF CUPS PER DAY OR WEEK OR MONTH	SIZE OF GLASS/CAN/BOTTLE SMALL 6 OZ OR LESS	REGULAR 8-16 OZ	LARGE 32 OZ OR MORE

14. HOW MANY CUPS OF CHOCOLATE MILK OR HOT CHOCOLATE DO YOU DRINK
PER DAY?___
OR PER WEEK ___
OR PER MONTH___
OR PER YEAR ___

DRAFT

[illegible] PSYCHIATRIC INSTITUTE [illegible]

[illegible]

15. HOW MANY CHOCOLATE BARS DO YOU EAT PER DAY? ____
OR PER WEEK? ____
OR PER MONTH? ____
OR PER YEAR? ____

16. HOW MUCH IN THE WAY OF CHOCOLATE FLAVORED OR COATED FOODS (IN TERMS OF NUMBER OF SERVINGS OF CAKE, ICE CREAM, ETC.) DO YOU CONSUME PER DAY? ____
OR PER WEEK? ____
OR PER MONTH? ____
OR PER YEAR? ____

17. PLEASE DESCRIBE BELOW.

ITEM	AMOUNT	FREQUENCY

C. PHARMACEUTICALS

18. DO YOU TAKE ANY PRESCRIPTION DRUGS THAT CONTAIN CAFFEINE? PLEASE CHECK THE LIST BELOW. CIRCLE AND FILL IN APPROPRIATE ANSWERS. IF YES, DESCRIBE FREQUENCY; DISTINGUISH BETWEEN PAST AND CURRENT USE.

CAFERGOT NO MIGRALAM R, MIGRAL, ESGIC, APECTOL, SOMA COMPOUND, CARNES R, DARVON COMPOUND N, DOLOR R, EMPIRIN COMPOUND R, FIORINAL-C COMPOUND N, A.P.C. R, PC-105 TABLET R, TRIGESIC R, FIORINAL...
OTHER_____
NONE.

19. IF YES, WHICH ONE(S)?_____
HOW OFTEN?_____

20. DO YOU USE ANY OVER-THE-COUNTER PHARMACEUTICALS THAT CONTAIN CAFFEINE? YES/NO____

21. IF SO, PLEASE SPECIFY NAME AND FREQUENCY_____

[illegible] PSYCHIATRIC INSTITUTE [illegible]

[illegible] DO YOU OR HAVE YOU EVER USED ANY OF THE FOLLOWING STIMULANT, ALERTNESS, OR STAY-AWAKE TYPE PREPARATIONS?

NO-DOZ R, ENERGETS R CANDIES, CAFFEDRINE R, DOUBLE-E ALERTNESS R, PROLAMINE R, QUICK-PEP TABLETS R, TIRED R, VIVARIN R, VIVA T.O. R, WAKOZ R...
OTHER_____
NONE_____

[illegible] IF YES, PLEASE SPECIFY NAME AND FREQUENCY_____

[illegible] DO YOU OR HAVE YOU EVER USED ANY DIET AIDS THAT CONTAIN CAFFEINE? PLEASE CHECK THE LIST BELOW, AND CIRCLE AND FILL IN APPROPRIATE ANSWERS. IF YES, PLEASE SPECIFY HOW OFTEN (FREQUENCY). DISTINGUISH YOUR PAST (EVER) FROM YOUR CURRENT USE (PAST 4 WEEKS).

DEXATRIM R, DIETAC R, PROLAMINE R, EIC SLIM T R, DEX-A-DIET R, HUNGREX PLUS R, SLIM ONE R...
OTHER _____
NONE_____

[illegible] IF YES, PLEASE DESCRIBE KIND, FREQUENCY, AND CURRENT USE.
_____
_____

[illegible] DO YOU OR HAVE YOU EVER USED ANY ANALGESICS, COLD, OR ALLERGY OR DIURETIC MEDICATIONS THAT CONTAIN CAFFEINE? PLEASE CHECK THE LIST BELOW, AND CIRCLE AND FILL IN, DESCRIBE PAST AND CURRENT USE.

ANACIN R, BROMA-SELZER R, COPE R, EXCEDRIN R, FLORINAL R, STANBACK TABLETS R, STANBACK POWDER R, TRIGESIC R, AQUA-BAN R, VANQUISH R, CRISTAN R, TRIAMINICIN R, CORYBAN-D R, EMPIRIN R, PHRENS R, MIDOL R, EASY-MENS R, GOODY'S HEADACHE POWDERS, PERMATHENE H2OFF R, HERBAL DIURETIC TABLETS R, CENEGESIC R, DURACYCLE-FORTE R, ELPHENEX R, MISTA-COMPOUND R, MIDRAN R, NEO-SYNEPHRINE R, SINTABILS R, SUPER-DECON R,
OTHER_____
NONE_____

[illegible] IF YES, SPECIFY YOUR CURRENT USE_____
_____

HISTORY

[illegible] WOULD YOU CONSIDER THE TOTAL CAFFEINE INTAKE YOU JUST DESCRIBED AS YOUR USUAL INTAKE? YES/NO_____

[illegible] IF NO, DESCRIBE.
_____
_____

DRAFT

[illegible] PSYCHIATRIC INSTITUTE ###

31. [illegible] HAS THIS PATTERN OF CAFFEINE INTAKE BEEN CHARACTERISTIC OF YOU?

NUMBER OF WEEKS---
OR MONTHS---
OR YEARS ---

32. HAVE YOU HAD ANY ADVERSE EFFECTS ASSOCIATED WITH TAKING ANY CAFFEINE CONTAINING SUBSTANCES? ANY UNPLEASANT REACTIONS OR ILL EFFECTS WHICH YOU WOULD SAY WERE CAUSED BY CAFFEINE?

YES/NO_____

33. IF YES, PLEASE DESCRIBE.

---

34. DID YOU EVER ATTRIBUTE ANY OF THE FOLLOWING SYMPTOMS TO THE INTAKE OF CAFFEINE? PLEASE CHECK.

INCREASED ANXIETY
INCREASED IRRITABILITY
INCREASED MUSCLE TENSION
GASTROINTESTINAL DISTURBANCES

35. DO YOU NOW FEEL CAFFEINE CAUSES ANY OF THE FOLLOWING SYMPTOMS? PLEASE CHECK APPROPRIATE COLUMN BELOW:

SYMPTOM	NO	SLIGHT	MODERATE	SEVERE
HEADACHE PAIN?	---	---	---	---
FEAR OF DYING	---	---	---	---
IMPENDING DOOM	---	---	---	---
PALPITATIONS	---	---	---	---
DIFFICULTY IN BREATHING OR RAPID BREATHING	---	---	---	---
FREQUENT URGENCY TO URINATE	---	---	---	---
FREQUENT URGENCY TO DEFECATE	---	---	---	---
DIZZINESS	---	---	---	---
CONFUSION	---	---	---	---
FEELING THAT THINGS ARE PEOPLE ARE UNREAL	---	---	---	---
DIFFICULTY IN CONCENTRATION	---	---	---	---
SWEATING	---	---	---	---
DIFFICULTY IN SPEAKING	---	---	---	---
DIFFICULTY IN DOING A JOB	---	---	---	---
TWITCHING (LIMBS, OR GRIMACING)	---	---	---	---
VOMITING OR NAUSEA	---	---	---	---
OTHER (DESCRIBE)_____	---	---	---	---
----------	---	---	---	---

[illegible] A ••• PSYCHIATRIC INSTITUTE •••

37. DID YOU EVER MARKEDLY CHANGE YOUR CAFFEINE INTAKE BY INCREASING IT FOR ANY REASON (FOR EXAMPLE, EXAMS, DRIVING, WORK, ATHLETICS)? YES/NO_____

38. IF YES, PLEASE DESCRIBE HOW MUCH AND UNDER WHAT CIRCUMSTANCES.

39. DID IT ACHIEVE THE DESIRED EFFECT? YES/NO_____

40. DID IT HAVE ANY OTHER EFFECTS? YES/NO_____

41. IF YES, PLEASE SPECIFY.

---

42. WHAT IS THE MOST CAFFEINE (THE HIGHEST DOSE) YOU HAVE HAD IN A SHORT PERIOD OF TIME? DESCRIBE BELOW

---

43. DID YOU EVER TRY TO REDUCE YOUR INTAKE OR STOP INGESTING ANY CAFFEINE COMPLETELY? YES/NO_____

44. IF YES, PLEASE ANSWER THE QUESTIONS BELOW:
HOW OFTEN? NUMBER OF TIMES_____
FOR HOW LONG? DESCRIBE._____
WHY? DESCRIBE._____
---

45. DID YOU HAVE ANY DIFFICULTIES ASSOCIATED WITH ANY OF THESE TIMES? YES/NO_____

46. IF YES, PLEASE SPECIFY.

---

47. DURING THESE TIMES, DID YOU EXPERIENCE ANY CHANGES, IN TERMS OF AN INCREASE OR DECREASE IN ANY OF THE FOLLOWING: PLEASE CHECK APPROPRIATE COLUMN.

EFFECT: INCREASE	ITEM	DECREASE
_____	HEADACHES	_____
_____	ANXIETY	_____
_____	IRRITABILITY	_____
_____	MUSCLE TENSION	_____
_____	DEPRESSION	_____
_____	MUSCLE WEAKNESS	_____
_____	ABILITY TO WORK EFFECTIVELY	_____
_____	ABILITY TO RELAX	_____
_____	ABILITY TO THINK CLEARLY	_____
_____	ABILITY TO CONCENTRATE	_____
_____	OTHER, PLEASE SPECIFY_____	_____

48. DID YOU EVER TAKE CAFFEINE CONTAINING SUBSTANCES TO COUNTER-ACT THE EFFECTS OR SIDE EFFECTS OF OTHER DRUGS? YES/NO_____

[illegible] PSYCHIATRIC INSTITUTE [illegible]

[illegible] IF YES, PLEASE SPECIFY.

50. IF YOU EVER TAKE A PILL TO COUNTER-ACT THE EFFECTS OF ANY OF THE FOLLOWING: IF YES, PLEASE CHECK BELOW.

ALCOHOL		_____
VALIUM		_____
LITHIUM		_____
BARBITURATES (SPECIFY)		_____
IMIPRAMINE (TOFRANIL)		_____
MAO INHIBITORS		_____

51. HAS CAFFEINE INTAKE, IN ORDER TO COUNTERACT ANY OF THE FOLLOWING EFFECTS, IF YES, PLEASE CHECK OR SPECIFY.

DEPRESSION	_____
INSOMNIA	_____
MENTAL FATIGUE	_____
PHYSICAL FATIGUE	_____
MUSCLE WEAKNESS	_____
OTHER _____	

52. DID IT COUNTERACT THE EFFECTS YOU WANTED IT TO? YES/NO/PARTIALLY_____

PLEASE SPECIFY:_____

53. DID YOU REGULARLY DRINK ANY OF THE FOLLOWING AS A CHILD? PLEASE CHECK THE ITEMS THAT APPLY.

	COFFEE	____
55.	TEA	____
56.	SOFT DRINKS WITH CAFFEINE	____

57. AS BEST YOU CAN RECALL, AT WHAT AGE DID YOU BEGIN CONSUMING REGULARLY THE FOLLOWING CAFFEINE CONTAINING BEVERAGES:

COFFEE	_____
TEA	_____
SOFT DRINKS WITH CAFFEINE	_____

[illegible] PSYCHIATRIC INSTITUTE [illegible]

[illegible] HOW MUCH YOU DESCRIBE THE REASONS FOR YOUR CURRENT CAFFEINE CONSUMPTION? PLEASE CHECK THOSE THAT APPLY.

MOOD RELATED _____
MENTAL ALERTNESS _____
COMBAT PHYSICAL FATIGUE _____
NEITHER _____
OTHER PRIMARILY PHYSICAL. PLEASE SPECIFY_____
OTHER PRIMARILY MENTAL. PLEASE SPECIFY_____

[illegible]. WHICH OF THE FOLLOWING ARE REASONS YOU CONSUME CAFFEINE-CONTAINING SUBSTANCES? PLEASE CHECK THOSE THAT APPLY. PLEASE INDICATE WITH A DOUBLE CHECK ( ) THOSE WHICH ARE YOUR MAJOR REASONS.

IT IS INEXPENSIVE _____
IT IS READILY AVAILABLE AND CONVENIENT _____
ITS TASTE _____
IT IS LOW-CALORIE _____
IT SATISFIES THE DESIRE FOR HOT OR COLD DRINKS _____
IT COUNTER-ACTS OTHER DRUG EFFECTS _____
OTHER. PLEASE SPECIFY._____

[illegible]. HAVE YOU HAD ANY DIFFICULTY IN ANSWERING ANY OF THE QUESTIONS? IF YES, PLEASE DESCRIBE BELOW.

_____
_____
_____
_____
_____
[illegible] SIP [illegible]

APPENDIX II

Filter calibration. Calibration of the filter combinations provided by the intensity programmer and fixed filter holder were determined within the optical system using photomultipliers 1 and 2 (see Apparatus section). Each photomultiplier was connected to its own resistor box (Shallcross) such that photomultiplier voltage gain was proportional to resistance. The resistance boxes provided inputs to an oscilloscope high-gain differential preamplifier (Tektronix, Type D).

The glow mudulator tube was pulsed on for 10 msec every second. With no filters between the two photomultipliers, the resistance of photomultiplier 1 was set at 100 kilohms and the resistance of photomultiplier 2 was adjusted to "null" the voltage difference between the outputs of the photomultiplier resistor boxes. After filters were introduced between the photomultipliers, the resistance of photomultiplier 1 was reduced to again "null" the voltage difference. This procedure was repeated three times for each filter combination.

The ratio of final resistance of photomultiplier 1 to 100 kilohms (multiplied by 100) indicates the per cent transmission for that filter combination. Per cent transmission was corrected for the input impedance of the oscilloscope

preamplifier using the following formula (R equals resistance in ohms):

per cent transmission = 1.0125/(1/R x 105) + 0.125)

density = log (100/per cent transmission)

References

American Psychiatric Association. (1980). Diagnostic and statistical manual of mental disorders (3rd ed.). American Psychiatric Association Committee on Nomenclature and Statistics, Washington, D. C. Author.

Berenhaus, I. J. (1981) One-pulse and two-pulse measures of visual temporal integration. Unpublished doctoral dissertation, City University of New York.

Berlyne, D. E., & Lewis, J. L. (1963). Effects of heightened arousal on human exploratory behavior. Canadian Journal of Psychology, 17, 398-411.

Broadbent, D. E. (1971). Decision and stress. New York: Academic Press.

Broadbent, D. E. (1978). The current state of noise research: Reply to Poulton. Psychological Bulletin, 95, 1052-1067.

Broadbent, D. E., & Gregory, M. (1963). Vigilance considered as a statistical decision. British Journal of Psychology, 54, 309-322.

Broadbent, D. E., & Gregory, M. (1965). Effects of noise and signal rate upon vigilance organised by means of decision theory. Human Factors, 7, 155-162.

Bruder, G., Spring, B. Yozawitz, A, & Sutton, S. (1980). Auditory sensitivity in psychiatric patients and non-patients: Monotic click detection. Psychological Medicine, 10, 133-138.

Campbell, R. A. (1963). Detection of a noise signal of varying duration. Journal of the Acoustical Society of America, 35, 1732-1737.

Childs, J. M. (1978). Caffeine consumption and target scanning performance. Human Factors, 20, 91-96.

Clubley, M., Henson, T., Peck, A. W., & Riddington, C. (1977). Effects of caffeine and cyclizine alone and in

combination on human performance and subjective ratings. Proceedings of the British Psychological Society, July, 652.

Cohen J. (1988). Statistical power analysis for the behavioral sciences (2nd Ed.). Hillsdale, N. J.: Lawrence Erlbaum Associates.

Colton, T., Gosselin, R. E., & Smith, R. P. (1968). The tolerance of coffee drinkers to caffeine. Clinical Pharmacology and Therapeutics, 9, 31-39.

Corwin, T. R., Kintz, R., & Beaty, W. J. (1979). Computer-aided estimation of psychophysical thresholds by Wetherill tracking. Behavior Research Methods & Instrumentation, 11, 526-528.

Davies, D. R., & Parasuraman, R. (1982). The psychology of vigilance. London: Academic Press.

Dimaio, V. J. M., & Garriot, J. C. (1974). Lethal caffeine poisoning in a child. Forensic Science, 3, 275.

Dusoir, A. E. (1975). Treatments of bias in detection and recognition models: A review. Perception & Psychophysics, 17, 167-178.

Easterbrook, J. A. (1959). The effects of emotion on cue utilization and the organization of behavior. Psychological Review, 66, 183-201.

Fisher, S. (1983). Memory and search in loud noise. Canadian Journal of Psychology, 37, 439-449.

Frith, C. D. (1967). The interaction of noise and personality with critical flicker fusion performance. British Journal of Psychology, 58, 127-131.

Goldstein, A., Warren, R., & Kaiser, S. (1965). Psychotropic effects of caffeine in man. I. Individual differences in sensitivity to caffeine induced wakefulness. Journal of Pharmacology and Therapeutics, 149, 156-159.

Graham, D. M. (1978). Caffeine--its identity, dietary sources, intake and biological effects. Nutritional Review, 36, 97-102.

Greden, J. F. (1974). Anxiety or caffeinism: A diagnostic dilemma. American Journal of Psychiatry, 131, 1089-1092.

Greden, J. F. (1980). Caffeine and tobacco dependence. In H. I. Kaplan, A. M. Freedman, & B. J. Sadock (Eds.), Comprehensive textbook of psychiatry (3rd ed.). Baltimore, Md: Williams & Wilkins, pp. 1645-1653.

Green. D. M., & Swets, J. A. (1974). Signal detection theory and psychophysics. New York: Wiley.

Grice, G. R. (1966). Dependence of empirical laws upon the source of experimental variation. Psychological Bulletin, 66, 488-498.

Guralnick, M. J., & Harvey, K. G. (1970). Response requirements and performance in a vigilance task. Psychonomic Science, 20, 215-217.

Hacker, M. J., & Ratliff, R. (1979). A revised table of d' for M-alternative forced choice. Psychonomic Society, 26, 168-170.

Hamilton, P., Hockey, G. R. J., & Rejman,. M. (1977). The place of the concept of activation in human information processing theory: An integrative approach. In S. Dornic (Ed.), Attention and performance, VI. Hillsdale, New Jersey: Erlbaum.

Harper, D. W. (1979). Signal detection analysis of white noise intensity on sensitivity to visual flicker. Perceptual and Motor Skills, 48, 791-798.

Hebb, D. O. (1955). Drives and the CNS (conceptual nervous system. Psychological Review, 62, 243-254.

Herskovic, J. E. (1985). Rating scale and response time analysis of visual flicker in depressed patients and normal subjects. Unpublished doctoral dissertation, City University of New York.

Humphreys, M. S., Revelle, W., Simon, L., & Gilliland, K. (1980). Individual differences in diurnal rhythms and multiple activation sites: A reply to M. W. Eysenck and Folklard. Journal of Experimental Psychology: General, 109, 42-48.

Jennings, J. R., & Wood, C. C. (1976). The e-adjustment procedure for repeated-measures analyses of variance. The Society for Psychophysiological Research, 13, 277-278.

Jerison, H. J. (1977). Vigilance: biology, psychology, theory and practice. In R. R. Mackie (Ed.), Vigilance, theory, operational performance and physiological correlates. New York: Plenum.

Kietzman M. L., & Gillam. B. J. (1972). Visual temporal integration and simple reaction time. Perception & Psychophysics, 11, 333-340.

Kirk, R. E. (1982). Experimental design: Procedures for the behavioral sciences (2nd ed.). New York: Brooks Cole

Kraemer, H. L., & Thieman, S. (1987). How many subjects? Newbury Park: Sage.

Lapidus, L. B., & Schmolling, P. (1975). Anxiety, arousal, and schizophrenia: A theoretical integration. Psychological Bulletin, 82, 689-710.

Lindsley, D. B. (1960). Attention, consciousness, sleep, and wakefulness. In J. Field (Ed.), Handbook of physiology (Vol 3). Washington, D. C.: American Psychological Society.

Loeb, M. & Binford, J. R. (1968). Variation in performance on auditory and visual monitoring tasks as a function of signal and stimulus frequencies. Perception & Psychophysics, 4, 361-367.

Mackworth, N. H (1950). Researches on the measurement of human performance (Special report No. 268). Medical Research Council, London: HMSO.

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

Mackworth, J. F., & Taylor, M. M. (1963). The d' measure of signal detectability in vigilance-like situations. Canadian Journal of Psychology, 17, 302-325.

Mannuzza, S., Spring, B. J., Gottlieb, M. D., & Kietzman, M. L. (1980). Visual detection threshold differences between psychiatric patients and normal controls. Bulletin of the Psychonomic Society, 15, 69-72.

Matin L. (1964). Use of a glow modulator tube for visual research. American Journal of Psychology, 77, 650-651.

McGuiness, D., & Pribram, K. (1980). The neuropsychology of attention: Emotional and motivational controls. In M. C. Wittrock (Ed.), The brain and psychology. New York: Academic Press, p. 95-136.

McNicol, D. (1972). A primer of signal detection theory. London: George Allen & Unwin, Ltd.

Naatanen, R. (1973). The inverted-U relationship between activation and performance: A critical review. In S. Kornblum (Ed.), Attention and performance, IV. New York: Academic Press, p 155-174.

Parasuraman, R. (1976). Consistency of individual differences in human vigilance performance: An abilities classification analysis. Journal of Applied Psychology, 61, 486-492.

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

Parasuraman, R., & Davies, D. R. (1975). Response and evoked potential latencies associated with commission errors in visual monitoring. Perception & Psychophysics, 17, 465-468.

Parasuraman, R., & Davies, D. R. (1976). Decision theory analysis of response latencies in vigilance. Journal of Experimental Psychology: Human Perception and Performance, 2, 578-590.

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

Poulton, E. C. (1977). Continuous intense noise masks auditory

feedback and inner speech. Psychological Bulletin, 84, 977-1001.

Poulton, E. C. (1979). Composite model for human performance in continuous noise. Psychological Review, 86, 361-375.

Pribram, K., & McGuiness, D. (1975). Arousal, activation, and effort in the control of attention. Psychological Review, 82, 116-149.

Rall, T. W. (1980). Central nervous system stimulants. The Xanthines. In G. A. Goodman, L. S. Goodman, & A. Gilman (Eds.). The pharmacological basis of therapeutics (6th ed.). New York: McMillan, p. 592-607.

Ritchie, J. M. (1975). Central nervous system stimulants (continued), the xanthines. In L. S. Goodman, & A. Gilman (Eds.), The pharmacological basis of therapeutics. New York: McMillan.

Robertson, D., Frolich, J. C., Carr, R. K., Watson, J. T., Hollifield, J. W., Shand, D. G., & Oates, J. A. (1978). Effects of caffeine on plasma renin activity, catecholamines and blood pressure. New England Journal of Medicine, 298, 181-186.

Routenberg, A. (1968). The two-arousal hypotheses: Reticular formation and limbic system. Psychological Review, 75, 51-80.

Sharpless, S., & Jasper, H. H. (1956). Habituation of the arousal reaction. Brain, 79, 655-680.

Smith, D. L., Tong, J. E., & Leigh, G. (1977). Combined effects of tobacco and caffeine on the components of choice reaction-time, heart rate, and hand steadiness. Perceptual and Motor Skills, 45, 635-639.

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

Swets, J. A., & Kristofferson, A. B. (1970). Attention. Annual Review of Psychology, 21, 339-366.

Tong, J. E. (1970). Psychophysical considerations for arousal and signal discrimination research in experimental psychopathology. Bulletin of the British Psychological Society, 23, 66-78.

Venables, P. H. (1963). The relationship between level of skin potential and fusion of paired light flashes in schizophrenic and normal subjects. Journal of Psychiatric Research, 1, 279-287.

Warm J., & Jerison, H. J. (1984). The psychology of vigilance. In J. Warm (Ed.), Sustained attention in human performance. New York: Wiley & sons, p 15-61.

Weiss, B., & Laties, V. G. (1962). Enhancement of human performance by caffeine and the amphetamines. Pharmacological Review, 14, 1-36.

Winer, B. J. (1971). Statistical principles in experimental design (2nd ed.). New York: McGraw-Hill.

Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit formation. Journal of Comparative and Neurological Psychology, 18, 459-482.