INFORMATION TO USERS

This reproduction was made from a copy of a document sent to us for microfilming. While the most advanced technology has been used to photograph and reproduce this document, the quality of the reproduction is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help clarify markings or notations which may appear on this reproduction.

1. The sign or “target” for pages apparently lacking from the document photographed is “Missing Page(s)”. If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure complete continuity.

2. When an image on the film is obliterated with a round black mark, it is an indication of either blurred copy because of movement during exposure, duplicate copy, or copyrighted materials that should not have been filmed. For blurred pages, a good image of the page can be found in the adjacent frame. If copyrighted materials were deleted, a target note will appear listing the pages in the adjacent frame.

3. When a map, drawing or chart, etc., is part of the material being photographed, a definite method of “sectioning” the material has been followed. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.

4. For illustrations that cannot be satisfactorily reproduced by xerographic means, photographic prints can be purchased at additional cost and inserted into your xerographic copy. These prints are available upon request from the Dissertations Customer Services Department.

5. Some pages in any document may have indistinct print. In all cases the best available copy has been filmed.
Hungerford, John Charles

EXPLORATORY STUDY OF RESPONSE TIME, EYE MOVEMENTS, EKG, AND EEG IN A SUSTAINED ATTENTION TASK

The Ohio State University

University Microfilms International

Copyright 1984 by Hungerford, John Charles

All Rights Reserved
EXPLORATORY STUDY OF RESPONSE TIME, EYE MOVEMENTS, EKG, AND EEG IN A SUSTAINED ATTENTION TASK

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

John C. Hungerford, B.S., M.S.

* * * * *

The Ohio State University

1984

Reading Committee:

Thomas H. Rockwell, Ph.D.
George L. Smith, Jr., Ph.D.
Richard J. Jagacinski, Ph.D.
Philip J. Smith, Ph.D.

Approved By

[Signature]

Adviser
Department of Industrial and Systems Engineering
DEDICATION

This study is dedicated to my wife, Ruth Faye Hungerford, and daughter, Julia Michelle Hungerford, for their love and faith in me.
ACKNOWLEDGMENTS

The author would like to express his deep gratitude to Cedric Sze, Dr. Kenneth Funk, Don Broach, Dr. Asa Bishop, and Becky Harbin.

I would especially like to thank Dr. Thomas Rockwell, the chairman of my committee, whose encouragement, support, and advice made this research possible. I also would like to extend my thanks to the reading committee for their help and suggestions: Dr. George L. Smith, Jr., Dr. Richard J. Jagacinaki, and Dr. Philip J. Smith.

I would like to thank the Ohio State Graduate College for providing a research fellowship to support the research. The author would especially like to thank Dr. John N. Snider, Department Head of the Industrial Engineering Department at The University of Tennessee, for his encouragement and support in this endeavor.

The author is indebted to Dr. Kenneth Heathington, Vice President for Research and former Director of The University of Tennessee Transportation Center and Dr. Michael S. Bronzini, Director of the University of Tennessee Transportation Center for making the resources of the Center available for the typing and preparation of this manuscript. Thanks to Mrs. Bette Huber who has provided moral support and unselfishly given of her time during the typing of this dissertation. I would like to especially thank Mrs. Renée Hall who has helped so much in
the preparation of this manuscript. Her expert and professional touch
have added a quality to the dissertation. Thanks also to Mrs. Kathy
Sjodin for her assistance.

Finally, I would like to thank my wife, Ruth who has encouraged
and believed in me throughout this research process. Thanks to my
3½ year old daughter who keeps asking, "Daddy when are you going to
finish your desecration?"
VITA

February 6, 1939 ...................................................... Born - Detroit, Michigan

1967 ................................................................. B.S.M.E., Ohio University
                                                Athens, Ohio

1968 ................................................................. M.Sc., Ohio University
                                                Athens, Ohio

1968-1970 ......................................................... Target Fellowship - Industrial and
                                                Systems Engineering Department,
                                                Ohio State University, Columbus, Ohio

1976-1979 ........................................................... Teaching Associate and Research Associate
                                                in the Industrial and Systems Engineering
                                                Department, Ohio State University,
                                                Columbus, Ohio

1979-1984 ........................................................... Assistant Professor, Industrial Engineering
                                                Department, The University of Tennessee,
                                                Knoxville, Tennessee

PUBLICATIONS

Final Report: Juvenile Delinquency Prevention and Control Plan for
the State of Tennessee, HEW Grant #70-20319/4-0, 1973.

Rockwell, T. H.; Bala, K. N.; Hungerford, J. C., "Evaluation of
Illumination Designs for Accident Reduction at High Nighttime--
Accident Highway Sites."

Rockwell, T. H. and Hungerford, J. C. "Improving Driver Nighttime
Performance on Rural Curves Through Novel Delineation Systems." Pre­
sent ed at the 1981 Human Factors Conference, Los Angeles, California.

Hungerford, J. C. and Heathington, K. W. "Conceptualization and
Subjective Evaluation of Innovative Railroad-Highway Crossing Active
Warning Devices." The University of Tennessee Transportation Center,
September, 1981.

Snider, John N. and John C. Hungerford. "Laboratory Specification
and Testing of Alternative Active Railroad-Highway Grade Crossing
Warning Devices." The University of Tennessee Transportation Center,
March, 1982.
FIELDS OF STUDY

Major Field: Industrial and Systems Engineering.

# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEDICATION</td>
<td>ii</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>iii</td>
</tr>
<tr>
<td>VITA</td>
<td>v</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>x</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>xii</td>
</tr>
<tr>
<td><strong>Chapter</strong></td>
<td></td>
</tr>
<tr>
<td>1. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Statement of the Problem</td>
<td>1</td>
</tr>
<tr>
<td>Purpose of the Study</td>
<td>3</td>
</tr>
<tr>
<td>Importance of Topic</td>
<td>3</td>
</tr>
<tr>
<td>Limitations</td>
<td>5</td>
</tr>
<tr>
<td>Definitions of Terms</td>
<td>6</td>
</tr>
<tr>
<td>Organization of the Study</td>
<td>10</td>
</tr>
<tr>
<td>2. LITERATURE REVIEW AND CONCEPTUAL FOUNDATIONS</td>
<td>11</td>
</tr>
<tr>
<td>Introduction</td>
<td>11</td>
</tr>
<tr>
<td>Organization of the Literature</td>
<td>11</td>
</tr>
<tr>
<td>Material Acquisitions</td>
<td>15</td>
</tr>
<tr>
<td>Sustained Attention Theories-Overview</td>
<td>16</td>
</tr>
<tr>
<td>Detailed Theory Review, Theory Assessment</td>
<td>17</td>
</tr>
<tr>
<td>and Present Research</td>
<td></td>
</tr>
<tr>
<td>Classical Conditioning/Inhibition Theory</td>
<td>22</td>
</tr>
<tr>
<td>Observing Responses-Operant Conditioning</td>
<td>25</td>
</tr>
<tr>
<td>Theory</td>
<td></td>
</tr>
<tr>
<td>Elicited Observing Rate Response Theory</td>
<td>31</td>
</tr>
<tr>
<td>Signal Detection Theory (SDT)</td>
<td>47</td>
</tr>
<tr>
<td>Probability Matching</td>
<td>59</td>
</tr>
<tr>
<td>Arousal/Activation Theory</td>
<td>62</td>
</tr>
<tr>
<td>Information Processing/Channel</td>
<td>74</td>
</tr>
<tr>
<td>Capacity Models</td>
<td>74</td>
</tr>
<tr>
<td>Expectancy Theory</td>
<td>81</td>
</tr>
<tr>
<td>Habituation Theory</td>
<td>93</td>
</tr>
<tr>
<td>Motivation Theory</td>
<td>102</td>
</tr>
<tr>
<td>Davies and Tune Model (1969)</td>
<td>108</td>
</tr>
<tr>
<td>Stroh IAF Model 1971</td>
<td>111</td>
</tr>
<tr>
<td>Status of Vigilance Theory Today</td>
<td>114</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>-----</td>
</tr>
<tr>
<td>Task Influences on Performance</td>
<td>117</td>
</tr>
<tr>
<td>Time Perception and Vigilance Performance</td>
<td>124</td>
</tr>
<tr>
<td>Task Duration and Time on the Watch Effects</td>
<td>127</td>
</tr>
<tr>
<td>Intersignal Interval Effects</td>
<td>128</td>
</tr>
<tr>
<td>Signal Duration Effects</td>
<td>129</td>
</tr>
<tr>
<td>Artificial Signals</td>
<td>130</td>
</tr>
<tr>
<td>Warning Signals</td>
<td>133</td>
</tr>
<tr>
<td>Signal Intensity Effects</td>
<td>134</td>
</tr>
<tr>
<td>Performance Measures in Sustained Attention Research</td>
<td>135</td>
</tr>
<tr>
<td>Literature Summary and Research Questions</td>
<td>148</td>
</tr>
</tbody>
</table>

3. METHODOLOGY ................................................................. 149

Scope ........................................................... 149
Selection of Subjects ........................................... 149
Instrument and Measurement System ....................... 149
Computer-Data Collection and Experimental Supervisory Control | 150
Display Panel and the Presentation of Signals | 155
Eye Activity Video Recording System | 155
Eye Movement Recording System | 157
Cardiac Recording System | 158
Recording Brain Wave Activity | 160
Research Design .................................................. 163
Experimental Procedures ....................................... 166
Analysis Strategy .................................................. 168

4. RESULTS OF RESEARCH ...................................................... 172

Introduction ........................................................... 172
Response Time Analysis Overview | 173
Inferential Analysis of Response Time Data | 191
Intersignal Interval Effects | 198
EEG Analysis Results ............................................. 200
Correlation Analysis of EEG Dependent Measures | 207
Inferential Analysis of EEG Measures | 210
Eye Movement Analysis Results | 219
Correlation Analysis of Eye Movement Performance Measures | 239
Heart Activity and Information Processing Analysis Results | 252
Combined Analysis of Performance Measures | 272
LIST OF TABLES

Table | Page
---|---
1. Matrix of Continuous Performance Tasks | 120
2. Efficiency as a Function of Vigilance Level | 137
3. Brain Signal Frequencies and Associated Behaviors | 146
4. Independent Variables and Variable Levels Used to Obtain Response Time Analysis Results | 175
5. Summary of Response Time Data for Each Subject and Run | 177
6. Response Time Data Summary, Data Averaged Across Subjects, Runs, Times, and Treatments | 187
7. Analysis of Variance for Response Times for Runs 1 and 2 Using an S x A x B Design | 193
8. Summary of Treatment Statistics for Response Time | 197
10. EEG Measures | 206
11. Independent Variables Used in the Analysis of the EEG | 208
12. Correlation Coefficient (r) Matrix of EEG Measures | 209
13. Summary of Predictor Variables That Appear in AID3 Analyses of EEG Measures | 211
14. Summary of One-Way ANOVA for EEG Measures and Independent Variables | 214
15. Subjects x Type of Prior Signal ANOVA for Alpha (α) | 217
16. Predictor Variables Used in Eye Movement Analysis | 223
17. Eye Movement Raw Measures and Performance Measures | 225
<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>18. Eye Movement Location Reliability Analysis-- Television Recording System Versus Biometrics 5GHV/2 Eye Movement Monitor ........................................ 226</td>
</tr>
<tr>
<td>19. Means, Variances, and Number of Observations for Eye Movement Times in Each Eye Movement Dwell or Transition Region of the Display Panel . . . . . . . . . . 233</td>
</tr>
<tr>
<td>20. Significant Correlations of Means and Variances of Eye Movement Performance Measures Aggregated Across Subjects Versus Time . . . . . . . . . . 241</td>
</tr>
<tr>
<td>21. Summary of Aggregation Intervals and Signals in Intervals for All Subjects .......................................................... 243</td>
</tr>
<tr>
<td>22. Analysis of Variance Summary Table for Eye Movement Measures .......................................................... 245</td>
</tr>
<tr>
<td>23. Heart Signal Performance Measures Aggregated Across 30-Second Intervals .......................................................... 260</td>
</tr>
<tr>
<td>24. Means and Variances of Heart Rate Performance Measures H1 and H2 Aggregated Across 5 Minute Periods of a Run .......................................................... 264</td>
</tr>
<tr>
<td>25. Correlation Coefficients and Levels of Significance for the Mean and Variance of Heart Rate Measures With the Mean and Variance of Eye Movement Measures (N = 40).* ........ 275</td>
</tr>
<tr>
<td>26. Summary of Major Findings .......................................................... 282</td>
</tr>
<tr>
<td>27. Summary of Vigilance Theories and Comments .......................................................... 287</td>
</tr>
<tr>
<td>28. Comparisons of Different Research Results Based on ISI .......................................................... 318</td>
</tr>
<tr>
<td>29. Truth Table to Diagnose and Distinguish Among Phenomena Related to Attention .......................................................... 320</td>
</tr>
<tr>
<td>30. Mean Intersignal Interval (ISI) and Limits on Regularity (Standard Deviation of ISI) for ISIs With a Uniform P.D.F.* .......................................................... 346</td>
</tr>
<tr>
<td>31. Summary of Proposed Research Conditions -- Mean Intersignal Interval and Variance of ISI .......................................................... 348</td>
</tr>
<tr>
<td>Figure</td>
</tr>
<tr>
<td>--------</td>
</tr>
<tr>
<td>2.</td>
</tr>
<tr>
<td>3.</td>
</tr>
<tr>
<td>4.</td>
</tr>
<tr>
<td>6.</td>
</tr>
<tr>
<td>7.</td>
</tr>
<tr>
<td>8.</td>
</tr>
<tr>
<td>10.</td>
</tr>
<tr>
<td>11.</td>
</tr>
<tr>
<td>12.</td>
</tr>
<tr>
<td>13.</td>
</tr>
</tbody>
</table>
14. Experimental Signal Presentation Schedule (Seconds Into Run) .................. 164
15. General Research Analysis ............................................................. 169
16. Response Time Analysis Strategy ...................................................... 174
17. AID3 Analysis for Response Time .................................................. 180
18. Stem and Leaf Plots, Box Plots, Normal Probability Plots and Summary Statistics for Response Times ..................................................... 183
19. Block Graph for Response Time Means for Each Level of Each Independent Variable ................................................................. 185
20. Block Graph for Response Time Means for Each Level of Each Independent Variable ................................................................. 186
21. Plot of Response Times by Run, Time, and Treatment (T1 = no Prior or 315 Seconds, T2 = 5 Second Auditory Warning Signal, T4 = 5 Second Visual Artificial Signal, T5 = 60 Second Visual Artificial Signal) ............................................. 190
22. Means and Variances of Response Times Versus Intersignal Intervals .......................................................................................... 199
23. 95 Confidence Interval for Response Time Data ................................ 201
24. Analysis of Strategy for EEG ............................................................ 204
25. AID3 Analysis of Power in Alpha (8-13 cps) Frequency Range if EEG Segment ........................................................................ 212
26. Eye Movement Analysis Strategy ...................................................... 221
27. Eye Behavior Observed on Video Tape for Runs With Missed Signals ......................................................................................... 228
28. Display Segment (1, 2, or 3) Activity Versus the 20 Second Interval Around a Missed Signal for Subject 3-Run V2 (Horizontal Scale Zero Point Occurs 5 Seconds Before the Onset of a Signal) .................................................. 231
29. Mean and Variances of Dwell Time in Display Segment 1 (E4) and Display Segment 3 (E5) ................................................................. 246
<table>
<thead>
<tr>
<th></th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>Mean of Total Time Spent in Display Segment 1-2, 2-1, 1-3, 3-1, 2-3, and 3-2 (E14)</td>
<td>247</td>
</tr>
<tr>
<td>31</td>
<td>Variance of Total Time Spent in Display Segments 1-2, 2-1, 1-3, 3-1, 2-3, and 3-2 (E14)</td>
<td>248</td>
</tr>
<tr>
<td>32</td>
<td>Effects of Signals and No Signals on the Means of Dwell Times in Display Segment 1 (E14)</td>
<td>253</td>
</tr>
<tr>
<td>33</td>
<td>Effects of Signals and No Signals on the Means of Dwell Times in Display Segment 3 (E5)</td>
<td>254</td>
</tr>
<tr>
<td>34</td>
<td>Effects of Signals and No Signals on the Means of Dwell Times in Display Segment 1-2, 2-1, 1-3, 3-1, 2-3, and 3-2 (E14)</td>
<td>255</td>
</tr>
<tr>
<td>35</td>
<td>EKG Analysis Strategy</td>
<td>257</td>
</tr>
<tr>
<td>36</td>
<td>WH1 = Mean of H1 (Mean of R-R Intervals) Aggregated Across 10 Second Intervals and Across Subject-Run Combinations Versus Time</td>
<td>262</td>
</tr>
<tr>
<td>37</td>
<td>MH2 = Mean of H2 (Mean of Variances) Aggregated Across 10 Second Intervals and Across Subject-Run Combinations Versus Time</td>
<td>263</td>
</tr>
<tr>
<td>38</td>
<td>Autocorrelation of Mean Heart Rate--MH1 (H1 is Aggregated Every 30 Seconds)</td>
<td>267</td>
</tr>
<tr>
<td>39</td>
<td>Autocorrelation of the Variance of Heart Rate--MH2 (H2 is Aggregated Every 30 Seconds)</td>
<td>268</td>
</tr>
<tr>
<td>40</td>
<td>Spectral Density Plot of R-R Intervals for Subject 4-Run A2</td>
<td>270</td>
</tr>
<tr>
<td>41</td>
<td>Spectral Density Plot of R-R Intervals for Subject 4-Run A1</td>
<td>271</td>
</tr>
<tr>
<td>42</td>
<td>Spectral Density Plots (S-01, S-02), Coherency Squared Plot (K-01-02) and Phase Spectrum Plot (PH-01-02) in Radians Versus Frequency in Radians for VE9 and VH2</td>
<td>277</td>
</tr>
<tr>
<td>43</td>
<td>Classification of Theories According to a State-Strategy Diminsion</td>
<td>303</td>
</tr>
</tbody>
</table>
44. Summary of Research Results and Theoretical Predictions for Response Time as a Function of Intersignal Interval .................................................... 310
45. Heuristic Model for Present Research ............................. 336
46. Temporal and Spatial Display and Task Variables in Vigilance Research ............................................. 344
1. INTRODUCTION

Statement of the Problem

The role of the human operator in human machine systems is rapidly changing. The role of the human operator is shifting from that of being a controller of relatively simple manually operated systems to that of being a system monitor and supervisor in complex systems. In the design of modern complex systems more and more of the system's functions are being shared between the human and the machine or being allocated entirely to the machine. According to Sheridan and Johannsen (1976) monitoring means the systematic observation by a human operator of multiple sources of information. The purpose of the monitor is to detect and diagnose difficulties in the operation of the system and to supervise the correction of the difficulties. Wiener and Curry (1980) see the movement of systems towards increased automation as a tradeoff. On the positive side, automation results in increased capacity and productivity, reduction of manual workload and fatigue, relief from routine errors, and economical utilization of machines. On the negative side, however, automation results in low alertness of human operators, lower job satisfaction, systems that are fault intolerant, human operator complacency, false alarms, automation induced failures and an increase in mental workload.
Operator complacency and loss of alertness are often observed in monitoring tasks (for example, Mackworth, 1950). This area of research has come to be known as vigilance research. Jerison (1977) equates vigilance research with sustained attention research and argues that sustained attention is fundamentally different from selective attention with respect to both its psychology and biology. Mackie (1977, p. 3) makes the following comment concerning the biology of sustained attention:

 Fifteen years ago there was practically no mention of the physiological correlates of vigilance performance other than a few remarks concerning muscle-action potentials, although in one interesting theoretical discussion Buckner (1963) concluded, "If the arousal hypothesis is going to be useful in explaining vigilance behavior, it is fairly obvious that we need an independent measure of arousal. I suppose a physiological measure is the most likely candidate (p. 124)." There was no reference to the many now commonly measured correlates such as EEG, cortical evoked potentials, heart rate, heart rate variability, GSR, CFF, and biochemical measures such as the catecholamines. . . .

Gale (1977a) cautions against an over-reliance on behavior data to describe vigilance phenomena. Rather, he suggests that vigilance research proceed along three directions: experiential, behavioral, and physiological.

A considerable number of studies examine vigilance performance in terms of some response measure such as response time or EEG or heart rate variability as a function of either display or task or organismic variables (Mackie, 1977). Few studies use a multilevel approach such as is suggested by Gale.
Purpose of the Study

This research is an exploratory investigation of the effects of changes in display and task variables in a sustained attention task on human visual, physiological, and psychomotor behavior. More specifically, the study investigates the effects of time on the watch, artificial and warming signals, and artificial and warning signal intervals on response latencies, eye movements, EEG, and EKG. Additionally, a further purpose of the study is to examine interdependencies between the dependent measures.

Importance of Topic

Many situations involve maintaining alertness in routine, repetitive tasks. Driving an automobile on an interstate highway at night is often monotonous but requires alertness on the part of the driver in order to avoid accidents. Similarly, a high level of sustained attention is required by the pilot of an aircraft. The pilot of the airplane has a wide range of demands place upon his skills. Near an airport, the pilot must simultaneously watch for other nearby aircraft, maintain radio contact with the control tower, and monitor critical panel instruments. Away from crowded airports, a pilot's alertness is likely to decrease as the result of reduced stimulation. The pilot may miss important instrument readings, route checkpoints, or fail to see other close flying aircraft. Another example of a situation requiring continued attention is air traffic control. The air traffic controller has the job of insuring that "blips" on the radar screen do not come together. Such a failure in detection and monitoring would be disastrous. The engineer of a train is often exceedingly busy but there
are other times when the train essentially runs itself and the engineer acts as a system monitor trying to anticipate what is going to happen or what might happen unexpectedly. Because of the slow and complex response of the train to changes in the controls, the train operator must maintain continuous alertness throughout the train run in order to insure the safety of the train. Situations requiring monitoring and sustained attention are not exclusive to the transportation system. Anesthesiologists are required to maintain a high level of alertness during extended surgical operations. The anesthesiologist must monitor the patient's vital signs as well as administer the proper amount of anesthesia at the proper time. Failure to perform properly may be life threatening to the patient. Surgical and medical complications and misadventures resulted in 3,184 deaths in the United States during 1975 (National Safety Council, 1977). Many industrial processes involve constant inspection of instruments that indicate the state of process being observed. Failure to correct system performance discrepancies based on the available information may result in costly losses.

The National Safety Council estimated that approximately 100,000 people were accidently killed and 10,300,000 people had disabling injuries in the United States during 1976. It is estimated that the cost of these accidents was $52.8 billion—a significant part of the gross national product (National Safety Council, 1977). One could argue that most accidents have a component of inattention involved although the primary cause might be attributed to something else. For example, accident statistics for Ohio for 1976 indicate that "driver inattention" was the probable cause in 16.4 percent of the total number
of accidents. However, "failure to yield, ran stop sign, and ran traffic signal" account for another 20.4 percent of the total number of accidents. The latter categories suggest that driver inattention may also be a component of causation in these classifications (Ohio Department of Highway Safety, 1976). It is obvious that inattention as an accident cause is underestimated and that the associated cost is high in lives, injuries, and money. For these reasons, there is considerable motivation to research the area of sustained attention.

The systems described above suggest that it would be useful to predict human performance in a sustained attention task. The motivation for doing this would be to improve operator and system performance. Ideally, one would like to get to the point of being able to construct an operator alertness indicator. Such an indicator would arise from being able to predict changes in operator performance from physiological changes rather than using task demands to predict physiological changes (Carriero, 1977).

Limitations

This study has several limitations which should be considered in any interpretive use of the data.

The generalization of the findings of this study is limited by the fact that all subjects were drawn from Ohio State University students. The subjects were solicited and recruited by means of a flyer that was posted in Baker Systems Engineering Building. The subjects were all volunteers and the research literature concerning problems with volunteer subjects has been documented by Rosenthal and Rosnow (1975).
Another limitation of the study concerns instrumentation. The validity of the research instruments is accepted for the research purposes. In the case of eye activity measurement equipment, validation of measuring equipment was performed simultaneously using other instruments to cross check measurements. However, the validity of heart and brain activity measurement equipment was accepted without cross checks.

Subjects were advised to come to research sessions rested and not under the influence of any drugs. Rigid control was not enforced, i.e., subjects were not isolated and controlled prior to experimentation.

Another limitation of the study is that imposed by the computer in the experiment. Due to computer memory space limitations, brain activity data was collected for only fifteen seconds around a signal presented for detection.

The results from two meter monitoring task used in this research may not generalize to other monitoring tasks. The extent to which the results generalize to other tasks depends upon the similarity of task components.

**Definition of Terms**

**Alerting Signal.** For the purpose of this study alerting signal is defined as a visual or auditory signal that precedes another visual signal by either 10 seconds or 60 seconds.
Alpha Waves. The frequency of the EEG in the range of 8-13 Hz. These frequencies are often associated with a relaxed wakefulness state.

Artificial Signal. A signal which is not discriminable from a real signal that is injected into the monitoring task for the purpose of maintaining performance level. In this research on artificial signal is used as a synonym for visual alerting signal.

Attention. Attention is used synonymously with information processing. Information processing refers to input-perception-output sequences performed by the human in response to stimuli. Perceptions refers to the interpretation of sensory input. There are three basic attributes associated with attention: intensity or effort, distribution, and regularity. Attentional processes are often categorized into divided attention, selective attention, and sustained attention processes.

Beta Waves. The frequency of the EEG in the range of 14-30 Hz. These frequencies are often associated with an excited state.

Delta waves. The frequency of the EEG in the range of 0.5-3.5 Hz. These frequencies are often associated with a sleep state.

Display. Display refers to the display board that the subject monitors approximately 25 inches in front of him. The display consists of two meters that fluctuate around some mean value. For analysis
purposes, the display is divided into three segments: left, center, and right.

**EEG-electroencephalogram.** The electrical activity of the brain is recorded from the surface of the head and is referred to as the EEG. Decomposition of the EEG into frequency components gives the power in different frequency ranges such as alpha, beta, delta, and theta.

**Eye Movement Activity.** For the purposes of this study, eye movement activity is defined as the number of transitions a person makes from one display segment to another in some unit of time.

**Eye Fixation Duration.** The dwell time of a person's eye in a display segment. The time a person's eye enters a display segment and the time it leaves a display segment is recorded. The difference between these two times is the dwell time.

**Heart Rate.** The heart signal is recorded by an amplifier and recorder. The record is called an EKG (or ECG). The EKG shows several distinct components during a cardiac cycle: P, Q, R, S, and T. The R peak is the largest amplitude component of the EKG waveform and is used to determine heart rate. The determination of the time between R peaks allows one to calculate average heart rate.

**Heart Rate Variability (HRV).** Heart rate variability refers to the variation in the R-R intervals that occurs over some unit of time. HRV is calculated for the R-R intervals just as any variance is determined for a set of measurements.
Monitoring Behavior. Refers to situations where a human operator shares key functions with a machine. Monitoring is the systematic observation of the information provided by the system over an extended period of time.

Non-signal. A non-signal refers to the times when both meters on the display board oscillate around a mean value of 50 with a range of 45-55. A non-signal occurs at least 98 percent of the time during an experimental run.

Response Time. Response time is defined as the time it takes a subject to detect and respond (push a hand held button) to a signal on the display board.

Real-time Experiment. For the purposes of this experiment, real time experiment is defined as an experiment where a computer is in the experiment. The computer supervises the experiment and is also the data acquisition system.

Signal. A signal is defined as a five second interval when one meter on the display panel oscillates around a different mean value than the other meter. For example, the left meter may oscillate between 40 and 50 with a mean value of 45 while right meter oscillates between 45 and 55 with a mean value of 50.

Sustained Attention. Sustained attention refers to the regularity and distribution of attention over an extended period of time while a person is in a monitoring task.
Theta. The frequency of the EEG in the range of 4-7 Hz. These frequencies are often associated with drowsiness, displeasure, and frustration.

Vigilance. Vigilance is the basic paradigm used to study sustained attention.

Warning Signal. An auditory signal that precedes a visual signal that is to be detected. Warning signal is used synonymously with auditory alerting signal.

Organization of the Study

In presenting the results of this research, the following organization is used. This study is organized into five chapters. Chapter 1 contains the statement of the problem, purpose of the study, need for the study, limitations, definitions and organization of the study. Chapter 2 reviews the relevant literature. Chapter 3 discusses the methodology used in this research. A discussion is provided that describes the research protocol, equipment, research design, and analysis techniques. Chapter 4 presents the findings of this research in the light of the research questions. Chapter 5 presents the conclusions, interpretation of the findings, and recommendations for further research.
Introduction

The purpose of this chapter is to examine some of the fundamental concepts involved in attention. Various kinds of attention reported in the literature will be discussed. The experimental paradigms that are used to study attention will be outlined. Sustained attention or vigilance as it is sometimes called is one of the basic kinds of attentional processes with which this research is concerned. A further objective of this chapter is to examine where other people have been in their studies of sustained attention and to use their results as a guide and springboard for the present research. A brief discussion will be presented to examine where the notion of sustained attention fits into a general consideration of attention and information processing. This chapter will present some of the current relevant psychological and physiological research that relates to sustained attention.

Literature searches were performed but did not turn up a single study that related response latency, eye movements, EEG, and EKG to sustained attention.

Organization of the Literature

Early reviews of the literature on attention indicated that the literature on attention could be organized along three dimensions:
process input variables, attentional processes, and process output variables. Figure 1 shows this organization.

Process input variables define the attentional process. For example, if the signals are infrequent, the task duration extends 10-20 minutes or longer and the task is a monitoring task, then a sustained attention process exists. Process input variables are classified as environmental, subject, task, and signal. Environmental variables are variables such as temperature, noise, lighting, and time of day. Subject variables are variables such as intelligence, age, sex, and cognitive style. Task variables are variables that characterize the subject’s task such as duration of task, knowledge of results, practice, pacing, rest pauses, multiple displays, and multiple modality monitoring. Signal variables are variables such as signal duration, signal intensity, signal frequency, inter-signal intervals and artificial signals.

Attentional processes can be grouped into three categories: sustained, intensive, and selective. Sustained attention research examines the fluctuations of attention over time. In terms of subject loading, the sustained attention task usually underloads a person. Intensive attentional research examines a person's ability to time-share between multiple inputs. Usually a dual task paradigm is used to study the intensive aspects of attention. Secondary task loading is increased to determine the effect on primary task performance. In terms of subject loading the dual or multiple task paradigm may overload the subject. Processes that involve choosing a response for a
Figure 1. Relationships between Attentional Process, Process Input, and Process Output.
particular stimulus from all of the available stimuli involve selective attention processes. A selective attention task may or may not overload a person. The study of such processes examine the way in which people prioritize and select certain information and shed the remainder. Obviously, the classification of attention into three categories is a research convenience. Information processing tasks do not generally draw purely on selective, intensive, or sustained attention processes.

Process output variables are dependent variables and can be classified as either psychological or physiological. Psychological measures of information processing may be behavioral such as errors and response times or subjective such as attitudes towards a task. Physiological measures associated with information processing may be chemical, electrical, or physical such as body temperature or heart rate.

The present research is concerned with literature that lies in the highlighted portion of Figure 1, that is, literature that is concerned with the way that task and signal input variables affect psychological and physiological performance measures in a sustained attention task. More specifically, the present research investigates the effect of changes in task variables and signal characteristics or response latencies, eye movements, EEGs, and EKGs. The literature review is organized according to the following outline: (1) introduction, (2) organization of the literature review, (3) material acquisition, (4) sustained attention theories, (5) task influences on
performance, (6) time perception and vigilance performance, (7) task
duration and time on watch effects, (8) intersignal interval effects,
(9) signal duration effects, (10) artificial signals, (11) warning
signals, (12) signal intensity effects, (13) performance measures in
sustained attention research, and (14) literature summary and research
questions.

Material Acquisitions

A survey of the literature related to this study was conducted. The
studies reviewed linked sustained attention to response time, eye
movement measures, EEG and heart rate. The search for related litera-
ture included a Mechanized Information Center (MIC) search at the Ohio
State University in the fall of 1977. Identifiers used in the search
were visual perception, visual search, attention, vigilance, and
information processing. A TRISNET search was also conducted at about
the same time at the Ohio Department of Transportation using the
identifiers alertness, vigilance, attention/inattention, and eye
movements. In 1978, a MEDLINE search was conducted at the Ohio State
University medical library to determine if any literature in that data
base was related to the present research. The identifiers used were
attention, evoked potentials, eye movements, EEG, and heart rate. An
additional computer search was conducted of the Psychological Abstracts
at The University of Tennessee library during August 1981. The iden-
tifiers cognitive processes, physiological correlates, EEG, eye move-
ments, heart rate, and heart rate variability were used. Cognitive
processes were linked with the other identifiers.
Although a large number of studies were found and reviewed, no one study was located that included response time, heart rate and heart rate variability, eye movements, and EEG.

**Sustained Attention Theories—Overview**

Loeb and Alluisi summarized the theories of vigilance in 1970 and again in 1977 (Loeb and Alluisi, 1970, 1977). They grouped current theories into the following categories: (1) inhibition theory, (2) observing response theory, (3) expectancy theory, (4) neural habituation and filter theory, (5) statistical decision theory and (6) activation or arousal theory. All of these theories have been used to explain the vigilance decrement (loss of detection capability over time). Inhibition theory ascribes the vigilance decrement to reactive inhibitive; the tendency not to repeat a response that has just been made. Observing response theory is generally associated with Skinner and Holland a protégé of Skinner. Observing response theory suggests that observing responses, e.g., eye, head, or hand movements are reinforced by signal detections and extinguish when not reinforced. Expectancy theory postulates the vigilance decrement occurs because a discrepancy exists between the expected signal rate and the actual signal rate; an expectancy mechanism acts to cognitively mediate a person's criteria for responding. Neural habituation theory and filter theory are physiologically based theories. The vigilance decrement is explained in terms of inhibitory pathways in the nervous system. Novel signals receive extensive processing and repeated signals received less processing. Statistical decision theory assumes
the vigilance decrement is caused by the monitor shifting his decision criterion or criteria because of fatigue or a lack of motivation. This often results in a decline in detections and false alarms. Activation theory is based on a neurological model. Loeb and Alluisi (1977) identify two types of arousal:

...a primary cortical one, which we shall arbitrarily label 'Type C', which is characterized by suppression of theta electroencephalographic rhythms and is closely linked to the vigilance decrement, and a more generalized 'Type G' arousal, characterized by elevation of blood pressure and body temperature, lowering of BSR, absence of alpha and slow wave EEG activity, etc., and which may be related to absolute level of vigilance and other kinds of performance but not to the vigilance decrement. (p. 747).

Leob and Allusi (1977) summarize the status of vigilance theories in the following way: (1) research fails to confirm any theory exclusively, (2) current data casts doubt that current theories can account for all vigilance phenomena, (3) brain activity analysis may provide useful advances in arousal theory, and (4) factors not encompassed by any of the theories are known to affect vigilance (p. 749).

Detailed Theory Review, Theory Assessment and Present Research

A scientific paradigm is the language, theories, concepts, methods conventions, and limits of a science. A paradigm is the way the scientist looks at a part of the world. One could speak of the "vigilance paradigm" but the problem with speaking about a vigilance paradigm is that limited agreement exists concerning research methods, dimensions of the vigilance task and underlying theories.

Theories are useful for simplifying reality. Theories organize and integrate past data, predict future data, and help guide future
research. A number of vigilance theories have been advanced to explain past research results. These theories explain some of the data but not all of the data. Moreover, some of the data is equally well-explained by several theories. What are the requirements for a comprehensive theory? What phenomenon must be explained? A comprehensive theory of vigilance should explain the following:

1. The "vigilance decrement" observed in some studies and the steady or fluctuating performance level observed in other tasks. In general, a comprehensive vigilance theory must explain the overall level of performance as well as trends over time.

2. The initial level of performance.

3. The progressive increase in response latency observed within a session.

4. This effect of changes in display, task, environmental and subject variables. Included here are the effects of motivational variables and payoff structures.

5. Individual differences.

6. Relationships between dependent variables used to measure vigilance states or performance.

7. Physiological changes that occur within a subject in sustained tasks.

8. Differences between laboratory findings and "real world" results—the effect of task operational-relevance on vigilance performance.

What theories have been reported in the literature to explain some of the previous phenomenon, particularly the vigilance decrement? Some of these theories have previously been discussed in the literature review. Frankmann and Adams (1962), Davies and Tune (1969), Davies and Parasuraman (1982), Stroh (1971), Mackworth (1969, 1970), Warm (1977)
Broadbent (1971), and Loeb and Alluisi (1970, 1980) provide comprehensive reviews of vigilance theories. The theories reported in the literature are:

1. Classical conditioning/inhibition theory
2. Observing responses - operant conditioning theory
3. Elicited observing rate response theory
4. Signal detection theory (SDT)
5. Probability matching
6. Arousal/activation theory
7. Information processing/channel capacity models
8. Expectancy theory
9. Habituation theory
10. Motivation theory
11. Davies and Tune model
12. Stroh IAF model

There is not universal agreement about how to classify various research results and theories. For example, Loeb and Alluisi (1980) classify inhibition theory, reinforcement theory, and observing response theory under the heading "learning models." Arousal and habituation theories are considered neurological models. Expectancy, probability matching, and signal detection theories are considered "psychophysical models." Filter theory and channel capacity are grouped under "information processing models." Davies and Tune (1969) group Broadbent's filter model, Elliot's attentional scan model, and Jerison's observing response model under attentional models. They group inhibition theory, reinforcement theory, motivation theory, and arousal theory under the generic category "drive theory." Warm (1977) and Jerison and Pickett (1963) discuss vigilance in terms of information processing stages: (1) storage of background information; sources of uncertainty, learning, inhibition theory, and expectancy theory are considered relevant to
this stage, (2) selection and transduction of information; sensory factors and filter theory are relevant to this stage, (3) orientation movements and decision-making; coupling and observing, observing response theory, the elicited observing rate hypothesis, and signal detection theory describe this stage, and (4) activity of neural attention units; activation theory and habituation theory describe and explain this stage.

The theories are not mutually exclusive. This is evidenced by the difficulty people have in arriving at a consistent classification of the various theories. For example, Mackworth (1950) used a classical conditioning paradigm to explain his results. This classical conditioning paradigm featured inhibition and reinforcement theory as explanations for the vigilance decrement. Later, Schroeder and Holland (1968), and Holland (1963) used an operant conditioning paradigm to explain eye movement rate decrements. Reinforcement was also a part of this model in the sense that detection of a signal reinforced an observing response.

Another aspect of these various theories is that some are very general and lack specificity. Arousal theory is an example of this kind of theory. Arousal was at one time thought to be a unitary process that could be represented by an inverted U-shaped curve. That concept has given way to a multidimensional concept of arousal. Lacey (1967) and Eysenck (1982) conclude there are three types of arousal which are quite different: behavioral, autonomic, and cortical. Measures of each one of these kinds of arousal are frequently poorly related to each other. For example, if one looks only at autonomic
activation, some physiological processes are mediated by the sympathetic nervous system (SNS) while others by the parasympathetic nervous system (PNS). With practice one can learn self-control over some involuntary autonomic processes. At one time this was not thought possible. Moreover, some processes seem to be controlled by both the SNS and PNS.

In addition to some concepts such as arousal not being well-understood and too general, some of the theories of vigilance appear to be ad hoc theories that are too specific and have been devised to explain a narrow group of studies. McCormack's reinforcement, inhibition, motivation theory and Smith's (1966) motivation theory fall in this category. Some theories such as arousal/activation theory emphasize changes in physiological state; other theories emphasize changes in signal detection parameters, response rates or expectancies. The point is that not all theories have something to say about all measures used in this research.

Another situation arises when reviewing vigilance theory. Theories fade from the mainstream because they are too simple to cover the wide range of phenomenon that need to be explained. The unitary concept of activation theory is such a theory. Another example is the classical conditioning theory used by Mackworth (1950) to explain the vigilance decrement. This theory has been dismissed as a viable explanation of the vigilance decrement by Deese (1955). Originally, Mackworth argued the vigilance decrement resulted from a weakening over time during the experimental session of reinforcement (knowledge of results) received during the training trials. Deese (1955) argued that by Mackworth's logic, an increase is signal frequency during the experimental session
would lead to more unreinforced trials and a larger vigilance decrement would occur. This did not occur; in fact just the reverse occurred, performance increased as the signal rate increased. Each theory and its applicability to this research will be briefly discussed in the following paragraphs.

**Classical Conditioning/Inhibition Theory**

The main exponent of inhibition theory was N. H. Mackworth (1950). Mackworth provided one of the earliest systematic studies of vigilance. The Mackworth studies were performed to determine: (1) the optimum length of radar watch, (2) the effects of interruptions (telephone message), (3) the effects of briefing and instructions, (4) the effects of knowledge of results, (5) the effect of stimulants (benzedrine), (6) the effect of individual differences, and (7) the effect of high atmospheric temperatures on monitoring performance. The motivation for the Mackworth studies was provided by World War II. Field reports had been received of overstrain among radar operators on submarine patrol--this overstrain was manifested in a loss of detection efficiency after about a half-hour on the watch.

Mackworth used three laboratory tasks to study monitoring performance: (1) clock tests, (2) synthetic radar tests, and (3) listening tests.

The clock used in the clock test had a black 6 inch pointer that moved around a white face without scale markings. The pointer jerked to a new position once every second; one hundred of these movements made the full circle. Occasionally, the pointer would make a double jump to which the subject was instructed to respond. Double jumps
(signals) occurred at 3/4, 3/4, 1/2, 2, 2, 1, 5, 1, 1, 2, and 3 minutes in that order followed by ten minutes of single jump movements. This completed the first half hour of a two hour session. The second, third, and fourth half hour periods were repeats of the first half hour. Mackworth used a 8 second limited-hold signal; if a subject did not respond to a signal within 8 seconds it was considered a "miss."

The synthetic radar task closely simulated the visual signals of an actual radar screen. Subjects were seated alone in a darkened room looking at a five inch screen one foot away at eye level. A bright line of light scanned the screen at one revolution/second. Subjects were instructed to detect "echos" that were oval in shape about 2 mm long and 1 mm wide. The signal schedule was the same used in the clock tests.

The listening test was an auditory counterpart of the clock test. Subjects monitored a standard 50 phons, 1000 cps sound which came on every 18 seconds and lasted 2 seconds. Occasionally the sound lasted 2.25 seconds. These longer sounds were the signals that the subject had to detect. The signal schedule was the same as used in the clock tests.

Mackworth found a reliable decrease in efficiency after about half an hour for all three tasks. Performance in the clock test could be improved by alternating the clock task with other work every half hour, knowledge of results, sudden telephone messages, and amphetamines, Mackworth found no end spurt at the end of a two-hour watch which he attributed to faulty time estimation. Large subject differences were reported but could not be explained by Mackworth. Briefings to subjects prior to an experiment did not affect results. High temperatures
and the expectation of a telephone call during the clock test resulted in decreased performance.

Mackworth interpreted his results in the current theory of the day, classical conditioning. The vigilance decrement was very similar to the experimental extinction observed by Pavlov and Hull. During the training session whenever a double jump of the pointer occurred, the experimenter said "now." The subject then pressed the response key. The conditioned stimulus was the double jump of the clock hand, the unconditioned stimulus the command "now" and the conditioned response was the pressing of the key. Mackworth argued that during test sessions, extinction of the conditioned response occurred due to a build-up of reactive internal inhibition. The unconditioned stimulus (KOR) was absent during test trials. Complete extinction was never observed and detection stabilized at about 70-75%. Mackworth attributed this partial extinction to the replacement of the reinforcing function of the unconditioned stimulus during training trials by expectancy and self-instructions during test sessions. Mackworth's observation that declines in performance could be prevented by telephones messages during a test session were explained in term of Pavlovian disinhibition. In a strict Pavlovian sense, a disinhibitory stimulus is a neutral one. Mackworth's telephone messages were meaningful and perhaps motivating.

Deese (1955) pointed out that detection rate does not always decline but may show some improvement or fluctuate irregularly making an inhibitory construct untenable. Deese reasoned that classical conditioning would predict larger and more rapid performance decrements
with increases in signal frequency during test sessions. The build-up of reactive inhibition should occur more quickly due to more unreinforced trials. Deese tested this hypothesis and found just the opposite result; the greater the signal frequency the greater the level of detections. Deese studied four subjects over four three hour periods. He used signal rates of 10, 20, 30, and 40 signals per hour randomized over time and display location. Deese found the average percent targets detected increased from about 50% for 10 signals per hour to nearly 90% for 40 signals per hour. Deese rejected classical conditioning theory in favor of expectancy theory.

**Observing Responses-Operant Conditioning Theory**

Holland (1957, 1958, 1963), a student of B. F. Skinner argued for an atheoretic approach to vigilance. The atheoretic approach promptly got labeled "observing response theory." Holland studied the results of Deese (1955) and others and hypothesized that observing rate and hence performance was controlled by the schedule of signal detections; signal detections were viewed as reinforcing. Holland (1958) devised a series of experiments that tested the applicability of operant conditioning theory to vigilance. The operant conditioning paradigm is as follows:

| Environmental cues, expectancy, often unknown. | Operant response R, some behavior elicited by a subject, e.g. orienting response, eye movements, head movements, etc. | Response contingent stimulus ($S^x$) such as detections. |

$S^x$ is a stimulus called a reinforcer if it increases the frequency of operant behavior R.
Holland had subjects (Navy enlisted men) sit in a darkened room. Their task was to observe and report deflections of a pointer on a dial that could only be seen after they had pressed a key which provided a brief flash of light, thereby illuminating the dial for 0.07 seconds. This was called an observing response. The subject had to release the key and depress it again to get another look at the dial. Whenever, he observed a signal which was a pointer deflection, he had to depress another key to report his detection. His only instructions were to make as many detections as possible and to reset the pointer as quickly as possible. The deflections of the pointer were programmed so as to make possible various schedules of detections (reinforcements). Subjects were given feedback about their performance (detections) at the end of a session. Cumulative response records were made of the frequency of observing responses. The number of detections and average time per detection were also recorded. Subjects were not in any way informed of the nature of the signal schedule or that the experimenter was interested in the frequency of observing responses. Holland used signal schedules as follows: (1) fixed interval (¼, 1, 2, 3, and 4 minutes between intervals), (2) fixed ratio (a needle deflection occurred only after a subject made so many responses; he used 36, 60, 84, 108, 150, and 200 responses), (3) variable interval schedules drawn from rectangular distributions (with mean intervals of 15, 30, 60, and 120 seconds; the range was 5 seconds to double the average interval), (4) multiple schedules (mixed schedule consisting of a combination of 40 response fixed-ratio and a 3 minute fixed-interval schedule), (5)
differential low-rate reinforcement schedule, and (6) the Mackworth signal schedule.

Holland concluded that signal detections can control the rate or probability of emission of observing responses. Particular fixed-interval schedules did not result in large differences in the slopes of the cumulative response records. This was also observed with the fixed-ratio schedules. However, various variable-interval schedules did result in obvious differences in observing rates. The observing rate was lower for a 2 minute average signal interval than for higher average signal intervals of 15, 30, or 60 seconds. Observing response rates were ordered for average signal intervals as follows 15 > 30 > 60 > 120 seconds. Holland observed that when no further signals occurred a reduction in observing responses occurred which he concluded was due to extinction. Holland studied observing behavior using a Mackworth signal schedule to try to obtain results similar to Mackworth. Unlike previous work, he used a limited-hold signal (previous work used unlimited-hold signals which insured no misses). Results were very similar to Mackworth's except Holland observed an end spurt which he said was due to subjects knowing the session was 2 hours long.

Schroeder and Holland (1968) performed an experiment that examined operant control of eye movements. Schroeder and Holland hypothesized that eye movements were observing responses and that eye movement rates could be controlled by signal rates. The investigators used limited-hold signals (2.5 second signal duration with a 2.5 second response time limit after a pointer deflection) with signal rates of 0.1, 1, and
10 signals per minute in 40 minute sessions. Sixteen subjects monitored four dials for pointer deflections. A new eye movement was scored only if a subject looked out of a dial area into another area or back into the same area. The Mackworth camera and corneal reflection technique were used to record eye movements. As signal rate decreased, percentage of signals detected decreased. Eye movement rate also decreased with decreased signal rate. As signal rate decreased, response rates became more variable within sessions and across sessions. Individuals with higher eye movement rates detected more signals. The slower the signal rate and the longer the time on the watch, the greater the tendency to fixate a signal without reporting it.

Schroeder and Holland concluded that time on the watch and signal rates similarly affected both eye movement rates and percentage of detections.

Reinforcement theory, according to Deese (1955) and Baker (1963c), would predict that performance was a function of the interval since the last detection of a signal; the longer the ISI the lower the probability of detection. Inhibition is assumed to grow as the length of the inter-signal interval. Baker (1963b) found percent detections and response time were uncorrelated with ISI.

There is a paucity of work that supports Holland's operant conditioning theory of vigilance performance. Moreover, some research such as Guralnick's (1972) is contradictory. Guralnick employed a variation of Holland's technique. He studied 12 subjects that took part in 10, 60 minutes sessions (only 1 session per day). Subjects
compared the heights of two vertical lines projected on a rear projection screen. A signal stimulus was one in which the upper segment right-hand member of the pair of lines was slightly longer than the left-hand member. The subject had an observing response key that briefly illuminated the display. Subjects could press this key as often as they wanted. Each 60 minute run was divided into 20 minute sections. Ten signals were presented in each 20 minute block and signals were available on a limited hold basis (5 seconds). Inter-signal intervals were drawn from a rectangular distribution with the restriction that no ISI be more than 240 seconds or less than 20 seconds. This schedule was repeated every 20 minutes throughout the session. Guralnick, unlike Holland, used a payoff structure although the payoff variable had no effect on performance. Additionally, he offered a bonus of $25.00 for the best performance. Guralnick found that observing rates increased over time whereas detection rates decreased over time. This result is opposite to that predicted by Holland's reinforcement theory.

Mackworth, Kaplan, and Metlay (1964) examined eye movements during a vigilance task and found that percentage of detections was unrelated to observing responses in a two dial monitoring task. Almost as many unreported signals were fixated as unfixated. Frankmann and Adams (1962) criticized Holland's work on the basis that rate of observing responses is not a theory but a technique for studying vigilance. Frankmann and Adams also suggest that the response of button-push in Holland's task introduces an element that is not present in free-scanning tasks. An alternate interpretation of the high rates of
button pressing responses is that subjects want to keep the display illuminated so they can scan when they want to. Rapid high-rate key pressing may introduce work inhibition or fatigue.

Jerison and Pickett (1963) observe that Holland's illumination response switch required more than three pounds to actuate. The switch used to indicate a detection could be actuated by a force of less than one pound. They assert that Holland's results cannot be duplicated unless similar switches are used. Jerison and Pickett suggested Holland's results reflected the cost of observing.

Blair (1958) performed some work that does not support Holland's work and also does not require button pushing responses. Blair had a continuous light source mounted on the head of subjects viewing a display in a darkened room. The light had to be directed at the display to see critical signals. A light-sensitive germanium diode on the display activated a recorder whenever a subject looked at the display, thereby giving a record of frequency and duration of observing responses. Only two of five of Blair's subjects reproduced Holland's findings of increased observing as signal time approached.

Holland's theory is weak in several areas: (1) the paradigm assumes detections are valued by people as reinforcers, (2) the observing response can be defined in nearly anyway the researcher cares, (3) the paradigm assumes that looking is seeing (processing information), (4) the paradigm does not consider the effects of motivation and payoffs, (5) the research generally emphasizes supra-threshold unlimited-hold signals, (6) not all observing responses can be observed, (7) some
observing responses are idiosyncrotic, and (8) frequency of observing responses may be influenced by mode of response, e.g. button-pressing versus eye-movements, responses.

Elicited Observing Rate Response Theory

One of the major criticisms of Holland's work centered around the issue of the inability to operationally define observing responses. A related issue had to do with the finding that eye movements, considered by some to be observing responses, were unrelated to detections (Mackworth, et al., 1964). Very often subjects fixated on the target but signals nevertheless went undetected. A process of looking without seeing, that is the "blank stare" phenomenon was observed. Jerison and Wing (1963) found the illumination response that Holland used as his observing response was not correlated with observing responses and detections. These findings led Jerison to extend Holland's observing response theory to include provision for the quality of observing.

One of the earliest versions of a theoretical model that considered quality of observing was discussed by Jerison and Pickett (1963). The model discussed detection by the observer in terms of six stages: (1) storage of background information, (2) orientation movements, (3) operation of sensory transducers, (4) activity of neural "attention" units, (5) decision-making and (6) responses required by the decisions. The model assumes that observers store information about the signal distribution and signal characteristics as well as the payoffs and risks associated with correct responses and errors.
The operation of sensory transducers is not considered central to the vigilance problem because signals are usually well above threshold. Orientation movements, activity of neural attention units, decision-making, and responses required by decisions are the stages of processing that are central to the vigilance problem. These stages can be studied in a decision analysis framework. There are two stages of decision-making: (1) decisions about whether to observe the display and (2) decisions about detection. Although these two stages are intimately linked, it is the observing response stage that is more important to vigilance performance according to Jerison and Pickett. If the observer makes the decision to observe, the outcome of the observation would be determined by the detection-decision analysis. If the observer's decision is not to observe, then the outcome can be determined entirely by the stored a priori signal probabilities and payoffs associated with detection responses, independently of the signal parameters. The observer is viewed as making sequential decisions on whether or not to observe the display on the basis of the expected value of an observation. The central feature of the model is the emphasis on cost of observing to the observer. There is a utility for observing which reflects the observer's value for attending or not attending to the display. The model treats the observing response as a discrete unitary event that either does or does not occur. Jerison and Pickett point out that one difficulty in using the model has to do with estimation of utilities and expected values. Moreover, the model assumes that utility for observing changes over a watch and this leads to the decrement function. Nonstationary utilities may explain the
vigilance decrement but such an assumption makes it nearly impossible to test and validate the model. This assumption would also make allowances for vigilance increments or situations where performance is steady over time.

Jerison and Pickett use decision matrices to illustrate the nature of the decisions to make observing responses. Two decisions are possible: 0 (observe the display) and Ō (do not observe the display). Two states of nature are possible: S (signal on display during observing period) and Š (no signal on display during observing period). An a priori probability is associated with each state of nature and a utility is associated with each joint occurrence of decision and state-of-nature. The expected value (EV) of each decision determines whether a person observes the display or not. IF EV(0) > EV(Ō), then the person observes the display. According to the authors, it is not likely that the a priori probabilities change over time. The authors favor the notion of the vigilance decrement resulting from changing observer utilities over time.

Jerison and Pickett's early model is not incompatible with other theories. Other theories tend to focus at the behavioral level and are narrow with regard to focus. For example, expectancy theory emphasizes the observers expectations concerning signal frequency as a determinant of detection performance. Inhibition theory is considered in the Jerison-Pickett model. Inhibition acts to increase the cost of observing over time. Changing utilities over time and the alteration between observing and not observing is likened to Broadbent's filter which acts to block the processing of certain information. Broadbent's filter
model suggests decision rules other than maximizing expected value that
determine which signal is actually observed.

Two studies cited earlier, Blair (1958) and Baker (1960b) did not
support Holland's findings that observing responses were related to
detections. Jerison and Pickett suggest these results could arise
because the cost of observing was low with eye movement responses
(those that have engaged in eye movement research would probably not
agree with Jerison and Pickett's observation).

Jerison and Pickett discussed characteristics of signals, spatial
and temporal uncertainty of signals, KOR, individual differences and
work-rest cycles in terms of their model. Signal intensity generally
increases the probability of detection but the intensity of a signal
can be weakened by imperfect observing. Signal duration may vary from
brief flashes to unlimited hold signals. Longer duration signals
improve detection performance. The temporal uncertainty of a signal
can be measured by the mean, variance, and higher moments of inter-
signal interval (ISI). ISI is the reciprocal of signal rate which
estimates a priori signal probability. In general, as the signal rate
increases (ISI decreases—temporal uncertainty decreases), the percent-
age of signals detected increases. Tasks with only spatial uncertainty
are called search tasks. Real world tasks generally involve both
spatial and temporal uncertainty. Laboratory tasks have shown that
tasks with spatial uncertainty (search tasks) abolish the vigilance
decrement (Jerison, 1963). Spatial uncertainty has the overall effect
of lowering efficiency by about 13% and the addition of temporal un-
certainty reduces efficiency another 20% although these effects may be
task-specific (Baker and Harabedian, 1962b). Providing KOR to subjects improves performance; the more information provided the more improvement occurs. Jerison and Pickett assert that the aforementioned variables are not independent but interdependent. For example, increases in signal intensity, frequency, and duration all may interact to provide KOR to observers.

A good theory should consider individual differences. Jerison and Pickett reviewed the research on individual differences and concluded that vigilance performance is highly variable and no satisfactory analysis of the sources of variation had yet been accomplished. Twenty years later their observation remains descriptive. Jerison and Pickett attribute differences in performance found in the laboratory and those found in field research to differences in the utility structure of the decision matrices.

Jerison and Pickett (1964) published a paper on the importance of the elicited observing rate. Jerison and Pickett presented events at a rate of 5/min or 30/min in an 80 minute vigil. Signal rate was a constant 15/hour for both conditions. Observers missed about 10% of the signals at the low event rate and 70% at the high event rate. An event was a pair of movements to the right of a 2 mm x 18 mm light bar. The light bar moved to the right 29 mm, snapped back to the zero position and then repeated the movement. A signal occurred when the second movement deflection was 35 mm. The interval spacing for events was 0.8 seconds and 10.8 seconds for high and low event rates. Signals occurred at successive intervals of 4.4, 6.0, 5.6, 1.9, and 2.1 minutes every 20 minutes. The authors concluded that detectability of a signal was
determined by what was going on when a signal was not present. Arousal theory would predict better performance for the high event rate unless the event rate is overloading subjects. Jerison and Pickett discussed their results in terms of a decision-theoretic model in which the expected value of observing is lower for low event rates.

Jerison, Pickett, and Stenson (1965) continued their research with variable event rates. A major purpose of this study was to integrate variations in observing behavior with the theory of signal detection (TSD). The monitoring task was similar to the one used by Jerison and Pickett (1964). The researchers examined the effects of 30 per minute and 5 per minute stimulus rates in an 80 minute vigil. A vigilance control condition was also included with stimulus rates of 60 per minute. Detection results similar to those reported by Jerison and Pickett (1964) were found, i.e. lower event rates produced better detection performance with no significant vigilance decrement whereas high event rates produced a decrement and significantly poorer performance. Jerison, Pickett, and Stenson used TSD measures to analyze performance, namely $d'$ and $\beta$. The authors caution that a traditional TSD interpretation should be avoided since the parameters $d'$ and $\beta$ were unrealistically high. The authors used the data from subjects in the high signal probability/low stimulus density group who detected all 20 signals and made no false alarms during the 80-minute vigil to estimate TSD parameters. The researchers treated the detection of all 20 signals as being at least as great as 19.5 out of 20 correct reports. Similarly, the emission of fewer than 0.5 false alarms was assumed when no false alarms occurred. This enabled the researchers to circumvent a
common problem in signal detection theory-based research, that is, the problem of no false alarms being emitted. Jerison, Pickett, and Stenson made the claim that $P_d = (19.5/20) = 0.975$ and $P_a = (0.5/380$ non-signal events) = 0.0013 was a lower bound on the probability of a hit and an upper bound on the probability of a alarm. One problem with defining limit values for $P_a$ and $P_d$ in this way is that $P_a$ and $P_d$ vary with the total number of signals and nonsignal events in an experiment. They calculated $d' = 4.968$ and $\beta = 13.507$ based on the above limiting probabilities. Actual results were then compared to these bounding values. The authors found that actual $d'$ values were lower than the minimum bounding value of $d'$ (median $d' = 3.042$, range was from 1.606 to 4.521). Actual $\beta$ values were considerably higher in some instances than 13.507 (median $\beta = 43.412$, range was from 0.639 to 500.198). The method of determining bounding values is arbitrary. If one chose a value of 17.5 correct reports as being the bounding value, then $P_d = (17.5/20) = 0.875$. Suppose a value of 19.995 correct reports were assumed, then $P_d = (19.995/20) = 0.99975$. Bounding values for $d'$ and $\beta$ for the best observers will vary considerably depending upon one's initial assumptions. Moreover, the actual $d'$ and $\beta$ values reported by the authors were based on the assumption that $N$ and $S+N$ distributions were Gaussian with equal variance. Acceptance of these assumptions requires a "leap to faith"--there was no test of the validity of these assumptions in the research. The authors argued that such high parameter values could not be interpreted as simple indices of detectability and criterion. Instead the authors suggested that TSD parameters could be used to define the quality of the observing response. A heuristic model was
developed based on three observing states: (1) alert observing (perfect attentiveness with true $d'$ and $\beta$), (2) blurred observing (process that increases the variance of the $N$ and $S+N$ distribution, $d' = 1.0$ and true $\beta$), and (3) distracted observing (large values of $\beta$ which results in no responses to any stimuli). Blurred observing was considered an observing state with inappropriate accommodation or fixation, eyes tearing, etc. Distracted observing is a state where there are neither detections nor false alarms. "True states" for $\beta$ were determined on the basis of the equation: $\beta_o = \beta_t^{0.4}$ where $\beta_o$ is the obtained criterion used by observers and $\beta_t$ is the theoretical optimum criterion. $\beta_t = P(N)/P(SN) = (380$ nonsignal events$)/(20$ signal events$)$ assuming that all utilities for responses are equal. True $d'$ was defined as the minimum $d' = 5.0$ value.

The vigilance decrements observed for the high event rates were explained in terms of changes in observing states during a session. Changes in observing states were indicated by changes in $\beta$ and to a lesser extent $d'$ over a session. Performance during a session was described in terms of the heuristic model, e.g. a person may use some mixture of alert observing, blurred observing, and distracted observing.

The elicited observing rate and decision process model was developed as the result of research that indicated that background event rate had a greater influence on performance than signal rate. This suggested that events elicited observing responses. Jerison and Pickett (1963) suggested that performance declined because of higher background event rates which had a greater attentional cost associated
with them. An observer observed or did not observe based on his utility for observing. The vigilance decrement often observed was considered due to changing utilities for observing. Later work expanded upon attentional states: observing could be alert, blurred, or distracted; signal detection theory parameters could be used to indicate what observing state was dominant during any part of a particular session. Increases in criterion ($\beta$) were considered as being related to declines in vigilance. Changes in $d'$ were found to be small and not considered to be related to the vigilance decrement. The difficulty with the elicited observing rate response model is that it is more speculative than factual. The model is suggestive of research that needs to be performed; however, in some instances the state of the art of existing methodologies may not permit validation of the model. For example, utility theory has some difficulty dealing with temporal and contextural effects. Utility theory usually assumes that a person's utility for some entity is a universal constant. Jerison's et al. model assumes a fairly rapidly changing value for observing. Another difficulty with Jerison, Pickett, and Stenson's (1965) model has to do with signal detection model assumptions, namely Gaussian N and S+N distributions with equal variances; although for non-optimal observing, Jerison, et al., do assume increases in the variance of the S+N distributions. Davies and Parasuraman (1982) discuss this issue and claim that no study has been performed to date that investigates increased variability due to distracted observing.

If signal variability increases with time, the ROC plotted on double probability paper will become increasingly more skewed in
relation to a reference unit slope over time. If this relationship can be shown, it would support Jerison's hypothesis and model.

At least two other papers by Jerison have been reported in the literature that elucidate upon the elicited observing rate response model (Jerison, 1967, 1970). Jerison (1967) discussed the role of activation theory and long-term performance in relation to some of his earlier reported results. Jerison criticized activation theory for being too broad and nonspecific but at the same time suggested that some vigilance data cannot be explained without it. According to Jerison, activation theory may have a role in analyzing and explaining the cost of observing. Jerison presented the results of further experiments that manipulated background event rate. He used two background event rates, 360 stimuli per hour and 1800 stimuli per hour. Additionally, he looked at two different signal probabilities, 0.0083 and 0.042. The experimental conditions were Group A (3/360), Group B (15/1800), Group C (15/360), and Group D (75/100). Jerison found results similar to earlier findings, performance was much better for low event rates. For the low event rate condition there was no signal frequency effect. Performance for the high event rate was considerably lower and there was a signal frequency effect that increased average performance from about 30% (low signal frequency condition) to about 55%. Another interesting result was that response latencies for Groups A and C was 1.8 seconds. Response latencies for Groups B and D was about 1.0 second. In other words good detection was accompanied by long latency periods whereas poor detection was accomplished by short latency periods. Jerison suggests that detection frequency is an index
of attending whereas response latency may have little to do with attentiveness—latency effects may follow a different set of rules. Jerison's data could be explained by overactivation; the paced high event rate overloads the observing-response-control mechanisms. The decrement observed with high event rates can be explained in terms of inhibition generated by the observing response. Jerison (1967, p. 373 concluded that "... activation theory in several forms may be necessary to understand different phases of vigilance performance." Jerison posed three questions for activation theory: (1) "What is the mechanism of inhibition that affects performance at high event rates?", (2) can arousal be conditioned over time as was suggested by his animal vigilance studies?", and (3) "What are the motor aspects of the emission of detection-indicating responses?"

Jerison (1970) extended his inquiry into the role of activation in vigilance. He developed a paradigm and discussed the role of physiology in relation to vigilance. The paradigm identified the various possible observer states in a vigilance task. The paradigm identified 20 different stimulus-response histories. An event is either a signal or nonsignal. Observing states may be alert (0), blurred (0') or distracted (0). Following the observing state is a subjective decision that the stimulus is either a signal or nonsignal (s or 0) a response (R_s or R_0) follows the decision that classifies an event as s or 0. For example, the chain E - S - 0 - s - R_s represents one branch of alert observing. Another branch of alert observing is E - 0 - 0 - R_s. Division of observing states into three discrete observing states, 0, 0', and 0 may be artificial; a continuum of attentiveness may be
more appropriate. The paradigm has no place for long-term effects, observer expectancies, or motivation. The value of the paradigm is that it permits a researcher to determine what dependent variable measurements should be taken and at what stage. Jerison suggests that three EEG measures, the contingent negative variation (CNV), a 100-200 msec latency component (N1 - P2) of the averaged evoked responses (AER), and a 350-600 msec AER latency component may be associated with vigilance. Jerison says the latter two measures may be relevant to the paradigm, the early component may be associated with the observing stage and the late component with the decision stage. Jerison does not mention it but the CNV may also be relevant to the vigilance paradigm. The CNV wave is thought to be associated with expectancy. Expectancy could act to evoke the observing response.

Several studies have been performed which cite Jerison's model as an explanation of findings. In general, these studies have assumed a two stage vigilance model. First stage decisions of the observer involve whether or not to observe. The second stage involves decisions about whether or not an event is a signal. Second stage decisions are described by TSD parameters and changes in these parameters reflect the utility for and quality of observing. Hatfield and Soderquist (1970) studied the effects of coupling in a vigilance task. The term "coupling" was introduced by Elliot (1960). Coupling is not a well-defined term. Elliot (1960, p. 360) defines coupling as follows: "By 'coupling' is meant an arrangement of the task so as to ensure that the signal put out by the experimenter's apparatus gets into the appropriate sensory input of the vigilant observer." Auditory tasks are
generally considered more tightly coupled than visual tasks. Subjects in the auditory tasks listened to a series of 0.5 sec, 70 dB pulses of white noise occurring every 2.5 sec. Aperiodically, 1.6 dB increments in intensity were added to the pulses. Subjects in the visual tasks monitored a light source for changes in intensity. There were two versions of this task, a loosely coupled visual task and a tightly coupled visual task. The loosely coupled task had the subject monitoring nonsignal pulses of 237 ft-c and signal pulses of 343 ft-c from a source 4 ft. away at eye level. The tightly coupled task had each subject monitoring pulses of 7 ft-c and signal pulses of 12 ft-c. Under similar conditions as for the other tasks except their eye lids were taped closed with transparent tape to increase coupling. The physical intensities of each stimulus were equated on the basis of $d'$ = 2.50. Experimental sessions were 90 minutes with an average signal rate of 1 per minute with the ISI drawn from a rectangular distribution with a range of 10 to 120 seconds. There was an added restriction that 15 signals had to occur in each 15 minute block of the session. No KOR was given. The authors used a $2 \times 2 \times 6$ factorial design with repeated measures on sense mode and time on the task. $\beta$ increased, while hits and false alarms decreased with time on the task regardless of coupling condition or sense modality. $D'$ did not change over time for all conditions. Correlations between visual and auditory detections and between auditory and visual $d'$ values were insignificant but correlations between auditory and visual false alarms, latency, and $\log \beta$ measures were significant. Mean latency increased over time for the auditory tasks. Hatfield and Soderquist discussed their results in
terms of an observing response model. The authors found high $p$-values (322.25) that was attributed to blurred or distracted observing. Hatfield and Soderquist expected loosely coupled tasks to allow more blurred observing with time on the task (increased $p$ over time) but did not find this result. The authors had difficulty in operationally defining observing state as did Jerison.

Metzger, Warm, and Senter (1974) performed an experiment using a visual task very similar to Jerison's to study background event rate and critical signal amplitude except they used limited hold signals. Metzger, et al., reported a decline in overall performance efficiency with time, a background event rate effect, a critical signal intensity effects, and an interaction between event rate and critical signal intensity. Unlike Jerison and Pickett (1964), Metzger, et al., found a vigilance decrement independent of signal intensity or background event rate. The event rates used by Metzger, et al., were 6 and 21 events/min. whereas Jerison and Pickett used 5 and 30 events/minute. This disparity in event rates may have accounted for the differences in results in the two investigations. Metzger, et al., rejected expectancy for a signal as an explanation for their results since inter-signal intervals and number of signals were the same for all experimental conditions. Instead, they explained their findings as supportive of Jerison's two-stage elected observing rate response model.

A study by Guralnick (1972) was previously cited as being not supported of Holland's operant theory was investigated within Jerison's two-stage model. Guralnick investigated easy versus difficult signals and two payoff matrices. Signals were available on a limited hold
basis for 5 seconds. Guralnick found that percent hits and false alarms decreased with time. A signal difficulty effect was found but not a payoff effect. Frequency of observing responses increased significantly over time. The observing response (OR) rate (key-pressing rate) was higher for the easy signal condition. D-prime decreased and log β increased over time. Guralnick interpreted these changes as reflecting increases in blurred observing. He interpreted the increases in observing responses over time as being under the control of different variables than those controlling the decrement function. Guralnick suggested that increases in β indicate that subjects have initial high expectations for signals but adjust these expectations as the session progresses. Guralnick does not indicate how these high expectations initially arise. It is not clear that differential treatment effects occur in training and experimental sessions with the exception that KOR was provided in the training session. Guralnick had a bonus of $25.00 for the subject with the best performance and this overall payoff may have determined the distribution of ORs over time.

Jerison's model is an improvement in the observing response theory of Holland. Holland's approach would build a theory based on all the effects observed in each research task and setting. Holland's theory was simply a catalog of observed effects in operant settings—there were no notions of reductionism or generalization. Jerison's model posed more questions than it answered but this is a value of his theory and not a criticism. Many features of Jerison's model require validation and validation may be extremely difficult at the present. Jerison's model is based on responses to well-defined events—it was clear
when an event occurred. Jerison had three classes of discrete events: (1) nonsignal events, (2) signal events, and (3) empty intervals between events. The nonsignal events in the present research were not well-defined. Nonsignal events were continuous with no breaks between them except for an occasional deflection of one of the meter needles. It is difficult to speak of background event rate in the present research. Jerison's model generalizes or speculates far beyond his data. His data said that percent detections was better for a background event rate of 5-minute than for a background event rate of 30-minute while the number of signals was held constant for both conditions. His results could also be explained in terms of changes in signal probability. His event rates and the numbers of signals presented to subjects were considerably higher than those used in the present research. Another difference concerns signal type, Jerison used "unlimited hold" signals whereas the present research employed "limited hold" signals. Jerison's later interests involving attentional states and concomitants of attentional states; namely physiological measures is more closely related to the focus of the present research. However, he has published little in the area of physiological states associated with observing responses. The most significant difficulty with Jerison's model is its inability to operationally define alert, blurred, and distracted observing. This is a characteristic of research by others, as well, who have used Jerison's model to explain $\beta$-values that are high and change over time.
Signal Detection Theory (SDT)

Signal detection theory arises out of the work of Green and Swets (1966), Swets (1964), and Swets, Tanner, and Birdsall (1971). It also had its origins in the statistical decision theory of Wald (1950) and the work of L. L. Thurstone in the 1920s.

The signal detection model was an attempt at solving a problem that existed in classical psychophysics, that is the problem of controlling or specifying the criterion used by observers in making perceptual judgments. The model reflects a two-stage process; one stage is sensory and the other is cognitive. Implicit in the theory is an activation process (Atkinson, Bower, and Crothers, 1965) that allows sensation to occur.

The SDT model is normative in that it prescribes how an ideal observer should behave. The model is descriptive in that is describes behavior in terms of SDT model parameters. The SDT model assumes an observer has an internal, cognitive model of a detection situation. It is assumed that the internal model is nearly identical with the classical hypothesis-testing decision model from statistics. Sensory information is represented by two overlapping, equal variance normal distributions, called the noise (N) and signal plus noise distributions (S+N). The noise distribution represents a combination of environmental noise and a person's internal variability. Stimulation may consist of background noise or background noise plus the signal. Four decisions are possible in a signal detection task, correct rejection (CR), correct detection (hit), error of omission (miss), and an error of commission (false alarm--FA).
Several variations of the basic signal detection task exist: (1) yes-no, (2) rating, (3) forced-choice, and (4) undefined observation intervals (Green and Swets, 1966; Swets, 1964). In the yes-no procedure, an observer is presented with a fixed interval; the interval is defined and he must decide whether the interval contains signal or noise. The observer usually is given knowledge of signal probabilities and payoffs. The rating scale task has an observation interval as in the yes-no task except the observer must make judgments about the input such as "definitely a signal, possibly a signal, possibly noise, or definitely noise." An advantage of such a procedure is that it is more efficient—the entire ROC curve can be developed in a single experiment whereas the yes-no procedure requires several experiments with variations in prior probabilities or reward matrices. Comparisons of the yes-no experiment and the rating scale experiment produce similar results (Sheridan and Ferrel, 1974). The forced-choice procedure presents the observer with two or more observation intervals. The observer must choose which interval contains the signals. In a two-alternative forced-choice (2 AFC) task, the observer is presented with two intervals, one of which always contains the signal, and he must choose one of the intervals as containing the signal. Again, there is close agreement between the 2 AFC and the yes-no method (Sheridan and Ferrell, 1974). The undefined observation interval introduces uncertainty into the research task. The introduction of uncertainty into the detection task allows a closer approximation to real-world monitoring tasks. The method is often called the "method of free response."
The use of undefined observation in SDT is the procedure that is most related to the vigilance task.

The measures of importance in SDT are percent detections and false alarms. If these are known and the assumptions of SDT are met, then two other measures can be derived, $d'$ and $\beta$. $d'$ is the index of discriminability and $\beta$ is the observer's decision criterion. Presumably, these two measures are independent. $d'$ is the distance between the means of the noise distribution and the signal and noise distribution in standard deviation units. It is assumed that the mean of the noise p.d.f. is zero and the variances of both normal distributions are one. The sensitivity measure $d'$ is defined as $d' = z(S/n) - z(S/s)$ when $z(S/n)$ are $z(S/s)$ are z-scores corresponding to the conditional probabilities $Pr(FA)$ and $Pr(HIT)$, respectively. The criterion or observer bias is defined as the likelihood ratio at the decision cutoff (c) point on the evidence axis. It is the height (ordinate) of the $S+N$ p.d.f. at the decision cutoff point divided by the height of the noise p.d.f. at the decision cutoff point, i.e. $\beta = f_{S+N}(c)/f_n(c)$. Generally, it is assumed the observer stores in memory the N and $S+N$ distributions and payoff matrix. He then observes a stimulus, encodes the stimulus, calculates the likelihood $l(x)$ associated with the observation (evidence) and then compares $l(x)$ with the criterion $\beta$. If $l(x) > \beta$, he says "yes"; if $l(x) < \beta$, he says "No." Generally, it is assumed the $d'$ is relatively invariant for an observer and $\beta$ can be influenced by manipulating signal probabilities and payoff matrices.

Actual signal detection performance is often compared with the behavior of an ideal observer. An ideal observer is one who establishes
a decision criterion in such a way that his payoff is optimized according to some decision rule. The decision rule mentioned most often in the literature is the maximization of expected value (maximizing gains and minimizing losses). If an observer maximizes expected value then $\beta_{\text{ideal}}$ should be as follows:

$$\beta_{\text{ideal}} = \frac{\Pr(n)}{\Pr(sn)} \cdot \frac{(C_{FA} + V_{CR})}{(C_{MISS} + V_{HIT})}$$

where $\Pr(n)$ and $\Pr(sn)$ are the prior probabilities of noise and signal plus noise. $C_{FA}$, $V_{CR}$, $C_{MISS}$, and $V_{HIT}$ are the values of a false alarm, correct rejection, miss and hit respectively. These are all absolute values. Other decisions rules are mentioned in the literature (Green and Swets, 1966; McNicol, 1972) e.g. maximizing the number of correct responses, minimizing false alarms, or satisfying the Neymann-Pearson criterion (maximize the hit rate for a fixed false-alarm rate).

An receiver operating characteristic (ROC) curve is a plot of $Pr(\text{HIT})$ versus $Pr(\text{FA})$. This curve is often called an isosensitivity curve. The shape and symmetry of the ROC curve can yield information about whether the underlying assumptions of SDT are met. The isosentivity curve for the observer can be plotted on double normal probability paper. An average $d'$ value can be determined from this plot. It is equal to $z(s/n)$ at the point on the ROC curve where $z(S/s) = 0$, i.e. it is the vertical distance from the positive diagonal ($d' = 0$) to the isosentivity line. If the variances of the distributions are equal, the slope of the isosentivity line will be 1. If the slope < 1, the variance of the $S+N$ distribution will be greater than the variance of the noise distribution. If the slope > 1, then the
variance of the noise distribution will be greater. If the isosensitivity curve is not a straight line, the underlying distributions are not Gaussian. In the unequal variance Gaussian distribution case, the reciprocal of the slope of the isosensitivity curve is equal to the standard deviation of the signal distribution. The unequal variance case does not pose any serious analysis problems. A measure of sensitivity \( \Delta m \) is defined as was \( d' \). If the assumptions of normality are not met, nonparametric measures for sensitivity and criterion have been developed. Davies and Parasuraman (1981, pp. 45-51) catalog a large number of parametric and nonparametric indices of detectability and response bias.

One of the objectives of a vigilance theory is to describe and explain the "vigilance decrement." A common finding in vigilance experiments exhibiting a decrement and analyzed in a SDT paradigm is that \( \beta \) increases over time. Swets (1977, p. 706) discusses the vigilance decrement,

"In most experiments, what had been defined as the "vigilance decrement," that is, a decrease over time of the proportion of true detections, was accomplished by a decrease in the proportion of false detections, and was seen to result from a change in the observer's decision criterion (\( \beta \)) rather than from a change in his sensitivity (\( d' \)). Alternatively, or in addition, the observer may have tended toward a stricter criterion because he regarded the value of a true-positive response as declining relative to the value of a true-negative response, and/or the cost of a false negative response as declining relative to the cost of a false-positive response."

The increase in criterion over time is often referred to as the "vigilance increment" (Williges, 1976). The increase in \( \beta \) over time indicates an increasingly more conservative response criterion. Williges (1976, p. 188) says in regards to the shift in decision criterion,
"... it is reasonable to expect a decrease in detection probability as documented by the vigilance decrement because detections and false alarms are too high at the beginning of the session for an ideal observer. What may appear to be degraded detection performance at the end of the session may actually be near optimal decision performance."

Davies and Parasuraman (1982) claims, "an interpretation based on expectancy theory seems to provide the best explanation of changes in the criterion and hence of the vigilance decrement which results from a criterion increment (p. 79)."

The decision criterion $\beta$ in vigilance has been found to be influenced by signal probabilities, payoff structures, time at work, and instructions. Swets (1977) thinks that variation in payoff structures affect $\beta$ the greatest amount whereas Davies and Parasuraman (1982) believe that the interaction of signal probability and time on the task have the greatest influence on criterion shifts. Time on the task generally increases $\beta$ whereas increases in signal probabilities tend to lead to a decrease or a relaxing of the decision criterion.

Occasionally, $d'$ will shift in a vigilance experiment. Swets (1977, p. 707) discusses shifts in $d'$, "The exceptional experiments--showing a decline in sensitivity, usually in addition to a progressively stricter criterion--were experiments with a relatively high event rate, requiring almost continuous attention to the display. The exceptional experiments, moreover, presented visual displays, which have been described as loosely 'coupled' to the observer relative to auditory displays, meaning that a warning of attention could result in events (signal and non-signal) being missed altogether."
Davies and Parasuraman (1982) reviewed the literature on experiments that have reported a shift in $d'$ over time. They found evidence that suggests that event rate interacts with task variables to bring about a decline in $d'$, for example high event rates in sensory tasks seem to bring on a decline in $d'$. Low signal intensity, time-sharing on the same sensory channel, stress and adverse environmental conditions also seem to decrease sensitivity.

Another measure besides the probability-based measures that have been mentioned in the literature is response latency. Detection latencies in vigilance are generally longer than reaction times. The level of signal detection has often been found to decrease with time on the task. If vigilance is a central process as signal detection theory assumes, then one would expect a negative correlation between percent detections and detection latencies; detection latency would be expected to increase with time. Parasuraman and Davies (1976) performed two 45 minute visual monitoring experiments and recorded latencies associated with correct detections, correct rejections, false alarms, and misses. The first experiment manipulated, signal probability and the researchers found that correct detection and false alarm latencies increased whereas correct rejection and omission error latencies decreased; criterion increased. The second experiment manipulated event rate. The investigators found that $d'$ decreased over time. They found that increased event rates increased latency of response to signals but not to nonsignals. Davies and Parasuraman (1977) found that detection latency increased with decreases in signal probability and increases in
event rate. FA latencies increased with decreases in signal probabilities but were unaffected by event rate increases. Correct rejection latencies decreased with decreases in signal probability. Miss latencies decreased with increases in event rate.

Pike (1973) discusses two models for response latency in signal detection experiments; one model is based on the notion that latency in detection is some inverse function of distance from the criterion; the other model assumes that observers make multiple observations on a trial and a count of these observations determines the response and its latency. Broadbent (1971) also suggested a line of thinking similar to that involved in Pike's second model.

Buck (1966) reviewed the response latency literature with reference to a hypothesis that reaction time and detection rate are correlated indicies of vigilance. He found that response time in many studies increased over time. Rest pauses tended to restore response time to its original level. Signal detectability (intensity) has the effect of offsetting deterioration of detection rate and response time. There seems to be a critical level of intensity below which a decrement will not occur. Signal rate effects have been reported in the literature; whether an effect is found seems to depend on the range of signal rates employed. Signal irregularity tends to increase response time. Knowledge of results has, in general, a favorable effect on response time. Evidence concerning number of signal sources, choice versus simple responding, and monetary reward is less conclusive.

Buck also found some evidence that response time was a more sensitive measure of performance in a task than detection rate; in some
instances response time increments have been reported in the absence of
detection rate decrements. In general, detection rates decrease with
increased response times. Buck found in the studies he reviewed rank
order correlations between response times and detection rates of be­
tween 0.20 to 0.83.

Buck reviewed studies that related EEGs to response times. A
study by Lansing, Swartz, and Lindsley (1959) found shorter response
times when a visual light signal occurred during α-wave blockage pro­
duced by a warning signal. If light signals appeared during the in­
terval required to produce blockage, response times increased. One
study reported by Haider (1963) indicated that subjects monitoring
infrequent irregular light signals produced a negative correlation
between reaction time to detected signals and EEG frequency below 12
cups in the 1-second interval preceding a signal. The correlations in
the preceding second and third seconds were smaller but in the same
direction. The mean value of wave counts in periods preceding
unobserved signals was significantly lower than for periods preceding
observed signals, a positive correlation between detection rate and EEG
frequency was inferred from this.

Buck developed a model relating vigilance state, signal detection
rate, reaction time and time on the task. Detection rate and reaction
don't covary in a simple way; rather a critical level is set by signal
intensity and duration. The critical level is high if signal intensity
and duration are low. The critical level is low if signal intensity
and duration are high. Vigilance fluctuates up and down over time but,
in general, the trend is downward. Where the critical level is established, will determine the covariation between reaction time and detection rate.

Davies and Tune (1969) and Davies and Parasuraman (1982) reviewed the literature on signal detection and response latencies. Davies and Parasuraman argue that response latency is an inverse function of the distance the current observation is away from the criterion. This assertion should be viewed more as a hypothesis than a fact. Interestingly, little information has been reported in regard to the variability of response time over time. Parasuraman and Davies (1977) performed a taxonomic analysis of variables that seem to affect sensitivity. They reviewed 33 studies that reported decreases in d'. Quality checks reduced this number to 27. Their results are reproduced in Figure 2. Decrements in d' or a similar measure occurred for both auditory and visual tasks. "Coupling" of the task to the observer via sense mode does not seem to be as an important determinant of changes in d' as had previously been thought. Reliable decreases in sensitivity do seem to be affected by high event rates and successive discriminations. Successive discrimination would have the observer detect a change in the intensity of intermittent light flashes whereas simultaneous discrimination would have the observer detect a specified configuration in a complex pattern of letters. The only study involving a multi-source task with a high event rate did result in a decrement in d'. Decrements in d' may be more ubiquitous than previously thought. The classification scheme also suggests areas where research might be done to determine task influences on sensitivity, e.g., high-rate, multi-source, auditory tasks.
Figure 2. Classification of Vigilance Studies According to the Presence (Y) or Absence (N) of a Reliable Decrement in Sensitivity Over Time (From Parasuraman and Davies, 1977).
Several reports have dealt with the applicability of SDT to vigilance (Davies and Parasuraman, 1982; Swets, 1977; Broadbent, 1971; Jerison, 1967; Broadbent and Gregory, 1963; Craig, 1977; Long and Waag (1981); Davies & Tune, 1969). A major concern of Jerison (1967) was that vigilance experiments reporting decrements had very high $\beta$ values that were psychologically meaningless. A concern expressed by Elliot (1960) had to do with the high signal rates used in laboratory vigilance studies as compared to low event rates observed in operational tasks. This is a criticism of vigilance research in general. Craig (1977) expressed concern over determining SDT parameters on the basis of grouped data--individual data may be quite different than group data. Many of the studies reported using SDT do not test the assumptions of SDT. The assumed internal model may be too complicated a model--the model assumes a certain level of "estimation" sophistication on the part of the observer. A serious shortcoming of SDT applied to vigilance is that many studies report no false alarms. This prevents a researcher from deriving the SDT parameters. Davies and Parasuraman (1982, pp. 56-57) discuss a method for estimating false alarm probabilities when there are no reported false alarms. This is given by Max $(p(s/n)) = 1-2^{-1/t}$ where $t$ is the number of nonsignal trials. Although estimation is one method that enables a SDT interpretation of data when no false alarms are present, confidence limits around $d'$ and $\beta$ may be large. Furthermore, this method seems to result in an overinterpretation of data. Swets (1977) says, "...there is more to vigilance than discrimination and decision, so that SDT cannot offer a total explanation of vigilance effects (p. 715)." SDT can offer much
to understanding vigilance but in the present research it does not seem an appropriate model for analysis or understanding. One fundamental purpose of the present research was to examine changes in cortical, behavioral, and physiological changes within a session. Allied to this purpose was to determine the effect of introducing a limited number of signals in a session (3-5 signals). The limited number of fairly distinctive signals, undefined observation intervals, lack of false alarms, and the physiological and cortical focus of the research all speak against a SDT interpretation of results.

Probability Matching

Loeb and Alluisi (1980) link probability matching to expectancy and classify it as a psychophysical model along with signal detection theory. A paper by Thomas and Legge (1970) proposed the "matching hypothesis" to describe the behavior of poorly informed subjects in signal detection experiments. Many SDT experiments don't provide feedback on each trial to subjects. Consequently, subjects fail to adjust their decision-making behavior optimally. The matching hypothesis was formulated to describe the detection behavior of subjects who lack the complete knowledge of an ideal observer. With SDT it is assumed that an ideal observer knows the probability density functions of N and S+N and that he can calculate the likelihood function \( l(x) = \frac{f_{S+N}(x)}{f_N}(x) \) and compare it with his criterion \( \beta \). \( \beta \) is chosen so that his expected payoff is maximized. Seldom do observers act in this way. Moreover, Legge and Thomas assert that he cannot act as an ideal observer in the absence of feedback because he cannot construct adequate internal models. A statement of the matching hypothesis is: \( P(YES) = \)}
q. The probability of reporting a signal is matched to the signal frequency, q. The probability matching model of Thomas and Legge was contraindicated by the data of individual subjects (Dusoir, 1974). Consequently, Thomas (1975) revised the model of Thomas and Legge (1974) to account for earlier model inadequacies. Deviations from the simple matching model are consistent with a model that assumes that criterion shifts on each trial depend on the discrepancy between sensory information and the criterion on the current trial. Shifts are in the direction of sensory information. Thomas assumes that when no feedback is given, criterion adjustments are based only previous responses and sensory information.

Craig (1978) raises the question, "Is the vigilance decrement simply a response adjustment towards probability matching?" Many vigilance experiments have shown increases in the SDT parameter β over time. Craig reviewed 100 signal detection studies on vigilance and inspection. He rejected studies that reported only latency measures, did not report false-alarm rates, reported only derived measures d' and β, involved asymmetric payoffs, involved drug and stress effects in the absence of control data, or that failed to report the proportion of signals presented. Craig reviewed the remaining 30 studies for response: signal (R:S) ratio biases over experimental sessions. Probability matching would suggest that the R:S ratio may be high initially for naive subjects and asymptote towards 1.00 as subjects gain experience. The decline in detections and false alarms over a session reflect an adaptive downward shift of R:S as the subject gains more experience rather than a shift in β. Craig found that studies with an
initial positive bias, i.e. $R:S > 1.0$ initially, did asymptote towards 1.0 at the end of a session. Studies without an initial bias did not show changes in $R:S$ over a session. A probability matching model would suggest that the initial $R:S$ ratio and its decline are both inversely related to signal probability. Craig found higher initial $R:S$ ratios for low signal probabilities ($R:S = 1.6$ for $P(S) < 0.02$). For intermediate signal probability values ($0.02 < P(S) < 0.10$), he found $R:S = 1.2$ and for high values of probability ($P(S) > 0.10$) he found $R:S = 0.9$. Craig suggests that a probability matching model may describe the vigilance decrement—particularly in the cases where observers are inadequately trained and are confronted with a low signal probability task.

Loeb (1978) presented a rejoinder to Craig which questioned the validity of using only declines in $R:S$ ratios to account for the vigilance decrement. He says, "It is not enough to tabulate response rate without tabulating signal detection and false positive rates, unless it can be shown that the reduction is both within and across sessions is similar (p. 450)."

The probability matching model is a model that combines the notions of expectancy with signal detection theory. Unlike SDT, the model allows the observer to change his decision goal from trial to trial. The model postulates that subjects progressively adjust their response rates downwards to match their experiences with signal rates and this accounts for the vigilance decrement. The discrepancy between response rates to a signal and signal probability is greatest for low probability signals. The Craik study found nearly half of the 100
studies reviewed reported sessions with R:S ratios less than 1, initially, and remained less than 1 throughout the entire session. If subjects really are matching response rates to signal rates, one would expect R:S to asymptote to 1. The model, just as SDT, is not particularly applicable to this research for several reasons: (1) a primary measure in this research was response latency—not percent detections; this measure is not useful for a probability matching model; (2) probability matching theory does not address false alarms and therefore is too limited, and (3) the assumptions underlying SDT and probability matching cannot be verified empirically in this experiment—the detection data per subject was too limited—a maximum of 22 detection responses were made per subject across all runs.

Arousal/Activation Theory

The concept of arousal has a role in most vigilance theories. In some instances arousal is central to a particular vigilance theory e.g. inhibition and habituation theories. In other instances arousal theory only has a supporting role, e.g. in expectancy theory. Baker (1963c) an advocate of expectancy theory said: "There are a number of vigilance data which prevent us from accepting arousal as anything more important than 'background supporting action' (p. 149)."

Duffy (1972) has defined activation as "... the release of energy into various internal physiological systems in preparation for overt activity (p. 578)." Since energy is released to physiological processes, Duffy believed activation could be measured by measuring changes in various physiological processes. Duffy's view of arousal was very similar to the view put forth by Yerkes and Dodson (1908).
Yerkes and Dodson found experimental evidence in research with rats for an inverted U-shaped function that related performances level to arousal. Performance was best for a moderate level of arousal and suffered if arousal were too high or too low. The optimal level of arousal for difficult tasks was thought to be lower than for easy tasks. These relationships are often referred to as the Yerkes-Dodson law. Evidence that Duffy embraces this view can be found in the statement, "... a wide variety of measures of physiological processes show relatively consistent changes with changes in what appear to be the energy requirements of the situation. The direction of the change in these measures seems, in general, to be consistent as the individual goes from the sleeping to the waking state, from waking relaxation to work on easy tasks, and from work on easy tasks to frantic effort or extreme excitement (p. 22)." Although Duffy favors the above view she acknowledges that activation is both general and specific. A person has a general level of activation that is present in the homeostatic condition. The patterning of activation, however, must change to meet the demands of specific situations that arise.

Duffy (1962, 1972) and Eysenck (1982) focus more on the psychological aspects of arousal; particularly as related to performance. Another view of activation is represented by Scott (1969). He says, "At the present time it appears more fruitful to conceive of activation as the degree of excitation of the brain stem reticular formation. (p. 354)." The reticular formation (RAS) lies deep within the brain stem and controls alertness, sleep, locomotion, the cardiovascular system, respiration and the gastrointestinal system. The RAS is connected to
other parts of the brain such as the cortex. It is believed that the EEG is the summation of excitatory and inhibitory activities in the nerve cells of the cerebral cortex. Many researchers believe the EEG or some time-locked averaged measure of the EEG is indicative of the state of consciousness.

Discovery of the arousal properties of the reticular activating system (RAS) has been recent. Moruzzi and Magoun (1949) found that direct stimulation of the brain-stem core reproduced EEG effects associated with a wakeful, alert state. The recency of this discovery has allowed only a vague and imprecise discription of the anatomy and function of the RAS. It is thought that upward, rostral discharges from the RAS produce the EEG in the cerebral cortex. Downward caudal discharges seem activate the neuroendocrine system, viscera, and muscles. According to Meldman (1970, p. 45): "It now appears that general arousal may be mediated by the more caudal portion of the reticular formation... this is consistent with what is known about startle and arousal responses, they tend to arise quickly, are somewhat evanescent, and adapt quickly to a repeated stimulus... specifically alerting reactions of the waking animal appear to be controlled by the rostral portions of the reticular formation in conjunction with non-specific nuclei of the thalamus and the diffuse thalamo-cortical projection system."

The early work of Mackworth (1950) indicated that vigilance performance could be maintained at its initial level of administering benzedrine to subjects. Subjects were "blinded" in clock test 8 and did not know whether pills were dummy tablets or benzedrine. Mackworth
called this a "pharmacological effect." This early work of Mackworth certainly had its effect on suggesting the presence of a physiologically-based arousal mechanism that affects vigilance.

Recent work in the area of neurophysiology and neuropharmacology tends to support the notion that arousal is mediated by chemical changes taking place in the entire brain but particularly in the reticular formation. Jasper (1969) discusses changes in brain chemistry that seem to be related to arousal. Neurotransmitters in the brain are chemical substances that are inhibiting or excitatory. These exist and functions at the neuron level. Neurons communicate with other neurons across specialized synaptic gaps. This is accomplished by the release of a single transmitter substance across the synaptic gap. A neuron may, however, have postsynaptic receptors for many substances. In addition, clearing mechanisms exist at each synapse to inactivate the transmitter substances. Hernández-Peón (1969) suggests the existence of certain parts of the RAS that have a chemical specificity to certain neurotransmitters. He also suggests that exhaustion of norepinephrine in the vigilance system may lead to overactivation. Other substances also seem to be related to arousal. Jasper (1969) reports the rate of release of certain amino acids (taurine, aspartic acid, glutamine, glutamine acid, glycine, and lysine) and acetylcholine from the surface of the cortex parallel reticular stimulation. The rate of release of gamma amino-butyric acid (GABA), on the other hand, was higher during sleep than during wakefulness. Jasper suggests that chemical correlates of activation are presently under study and hard to sort out.
The purpose of the foregoing discussion emphasizes that much is unknown about brain chemistry and the anatomy and function of the reticular formation. It is perhaps this lack of knowledge concerning brain neurophysiology that has resulted in a vague and general psychological theory of arousal. Despite the lack of specificity in arousal theory, it is intuitively appealing because of our individual experience with various arousal states. Moreover, nearly every vigilance theory appeals to the concept of activation in some form. The largest problem with arousal theory is that there is no single specific theory one can identify.

What evidence is there for arousal theory as an explanation of vigilance performance? The performance improving pharmacological effect of benzedrine found in the Mackworth (1950) study has already been mentioned. Jerison (1967) makes several points concerning activation theory and long-term performance: (1) idiosyncratic activation patterns make it difficult to specify activation patterns across individuals or to estimate generalized behavioral activation from physiological measures, (2) activation or arousal plays a significant role in vigilance performance, (3) activation is generally viewed as a nonspecific effect that can be directed in many ways, and (4) activation theory generally predicts that detection latencies are low and detection rates high under optimally activated performance—the reverse is true under less activated conditions. Jerison showed that subjects exhibited superior performance when he used low background event rates; however, response latencies were higher than for high event rates. Performance decrements in some vigilance tasks are not the result of
understimulation but overstimulation. He says: "It is an implicit yet strong criticism of activation theory that is not embarassed by this shift in viewpoint. Overactivation, like underactivation, can be handled by activation theory to predict performance decrements. Alternatively, an activation theorist could point out that the high rate of unreinforced stimulation would produce habitation of the activation system. The theory is so broad that almost any result seems to be explainable after the fact (Jerison, 1967, p. 380)!

Lacey (1967) also discusses the need for revision of activation theory. Lacey suggests there are three different forms of arousal: (1) electrocortical, (2) autonomic, and (3) behavioral. These systems are complex, imperfectly-coupled interacting systems that may or may not act simultaneously. Traditional activation theorists assume that electrocortical, automatic, and behavioral arousal all occur simultaneously in response to stimulation. According to Lacey, these three forms of arousal often do occur simultaneously but limitations of present knowledge make it impossible to predict the frequency and conditions under which these three forms occur simultaneously. The need for a revised and broadened concept of activation theory is based on several research results. First it has been shown that somatic and behavioral arousal can be disassociated experimentally by the use of certain drugs or surgery. This suggests that somatic and behavioral arousal are mediated by separate neural mechanisms. Lacey cites a study by Malmo (1966) that reported performance declines in a divided attention task but no changes in EEG, EMG, EKG, respiration or palmar conductance. Second, traditional activation theory would predict
strong intercorrelations among physiological variables said to measure arousal. Correlations within autonomic nervous system (ANS) measures and between CNS and ANS measures are generally low. Thirdly, Lacey refers to stimulus specificity or situational stereotypy. This means that different activities produce different somatic patterns. For example, warm stimuli produce one activation pattern whereas cold stimuli produce another. Anger directed inwards produces a different pattern than anger directed outward. The researcher is faced with the same problem as with a theory based upon an operant conditioning paradigm--the theory becomes an endless cataloging of all stimulus situations. Clearly, this is an undesirable position from the viewpoint of the theorist. Fortunately, a generalization has come out of Lacey's work. This generalization is called the intake-rejection hypothesis and concerns interactions between cardiovascular activity and the brain. Lacey has argued that sensory intake or attention to the environment results in decreased cardiac output. Sensory rejection or the focusing of attention inward results in increased cardiac output. Lacey referred to this task-specific response of somatic functions as directional fractionation of responses. Lacey and Lacey (1974) have shown that heart activity affects cortical function. Elevated cardiovascular activity tends to be inhibiting and decrease EEG activity. Conversely, decreased cardiovascular activity tends to increase EEG activity. Traditional activationists would predict decreases in slow brain activity and increases in fast activity with increases in stimulation. Similarly, heart rate would increase with increases in stimulation. Intake-rejection theory would predict just
the opposite cardiac-EEG pattern to external stimulation. The controversy continues and there is evidence cited by both camps to support their position.

Arousal theory is replete with conflicting results, ambiguities and untestable hypotheses. When discussing arousal theory, one should really speak of arousal theories, i.e. which and/or who's theory? Arousal has been used to interpret information processing, performance, motivation, attention, anxiety, and effects of environmental stressors—it is clear that an understanding of arousal will also provide an understanding of many other processes. This makes interpretation of research results in terms of arousal theories both easy and difficult. Interpretation is easy because the theory is broad and general enough to predict nearly any results. Interpretation is difficult because it is hard to make any real sense out of the muddle. Eysenck (1982) has tried to impart some order to arousal theories and has a thorough review of the literature. Eysenck dismisses a Yerke-Dodson unidimensional view of arousal as simplistic, limited and in many instances erroneous. Eysenck does say, "... it is nonetheless of interest that it does describe reasonably adequately findings involving extremely heterogeneous arousers (p. 173)." The work of Lacey, previously mentioned, did much to make people aware of the need for a revision of arousal theory—Lacey argued for three systems of arousal: behavioral, autonomic, and cortical. Pibram and McGinness (1975) also argue for three separate neural systems that control arousal. One system is based in the amygdala and controls arousal—it is associated with phasic physiological responses to stimuli. A second
system is located in the basal ganglia of the forebrain and is related to tonic physiological readiness to respond—it controls activation. Another system is located in the hippocampus and serves the function of coordinating arousal and activation processes (note the distinction between arousal and activation—many theorists make no such distinction). Eysenck (1982) argues for two arousal systems. One system is a passive system that is related to performance in a task. A second compensatory, cognitive control system monitors the first and alters resources when the highly aroused or under aroused subject perceives his performance falling below an acceptable level. This structure is not unlike Broadbent's (1971) theory of upper and lower interrelated arousal mechanisms. There seems to be some evidence for upper and lower arousal mechanisms. When two arousers affect the same arousal mechanism, performance is impaired more than if two arousers affect the different arousal mechanisms. Wickens (1980) has done some work in this area with attentional resources; tasks that draw from the same resource pool adversely influence performance more than if the task draws from different resource pools. The view of Eysenck is that the passive system does not exert a direct influence over behavior but that its effects are indirect and mediated by the central, cognitive, compensatory system. Both systems affect the level of arousal indicated as by physiological measures. Some physiological support exists for a two component arousal system—particularly in the area of sleep deprivation research. An outline of the two-component system is shown in Figure 3. Sleep deprivation seems to affect the passive arousal system—arousal is decreased when a person is sitting passively or
Motivating task, some stressors

Decreased arousal A-effect

Sleep deprivation, monotonous, paced tasks

Increased arousal B-effect

B-effects tend to offset A-effects

Active, compensatory, cognitive, task-specific arousal system, compensates for sub-or supra-optimal performance of the passive system.

Passive Arousal System

Motivating task, some stressors

Figure 3. Eysenck's Two Arousal Systems.
engaging in a nonstimulating task. If subjects are provided suitable motivation, the compensatory, cognitive system becomes active and arousal increases. A unidimensional one-stage activation system would predict larger changes in autonomic activity than occur. It appears that a compensatory arousal mechanism augments a passive arousal mechanism. Frankenhaeuser (1975) showed that higher catecholamine excretion levels seemed to be related to better performance in a variety of tasks such as vigilance, arithmetic, CRT and learning. She interpreted this result as an indication of an active system. Paced tasks seem to overdrive the compensatory system and prevent recovery in the passive arousal system. Unpaced tasks seem to allow a person to recover and compensate for inefficiencies. Barbituates, which are cortical depressants seem to affect unpaced tasks more than paced tasks. Tranquilizers which affect reticular formation activity unpaced tasks more than paced tasks. Tranquilizers which affect reticular formation activity seem to affect paced tasks more than unpaced tasks--this was also true for sleeplessness (Mirsky and Rosvold, 1960). This would suggest at least two different arousal systems, one lower and one higher.

There appears to be at least two broad theoretical views of arousal. One view can be considered to be the classic unidimensional--view of Duffy. This view assumes that activation follows an inverted U-shaped curve. Low levels of arousal result in suboptimal performance as do high levels of arousal. Overstimulation results in performance much like that observed in the understimulated observer; however, arousal is higher and load stress results in perceptual narrowing. This perceptual narrowing manifests as missed or ignored cues and
delays in the processing of signals. Duffy (1962) has summarized the
views of traditional activationists: (1) the effect of stimulation is
to cause a greater degree of excitation in a wide variety of organs and
systems, (2) evidence suggests that measures of arousal are correlated
and not independent, and (3) it seems to be justifiable to think of
activation of the observer as a whole and to attempt to measure it (p.
110).

Another theoretical view of arousal has been advanced by Lacey
(1967) on the basis of a considerable number of findings which did not
conform to the traditional view. Lacey views arousal as consisting of
three forms: autonomic, cortical and behavioral (Duffy's view empha-
sized behavioral arousal.). Lacey has argued that changes in the
components are often not well-correlated. Recently, there is some
evidence to suggest at least two aspects of brain arousal, a higher and
a lower. Lacey also claims that different patterns of autonomic activ-
ity are task-specific. This view put theorists in a difficult posi-
tion. A theory based on each type of task does not provide any kind of
generalization. However, Lacey observed that general autonomic pat-
terns are associated with internally-directed attention, e.g. problem-
solving, reasoning, mental arithmetic, etc. Externally-directed
attention yields a different pattern of autonomic activity that is
characterized by directional fractionization, i.e. heart rate deceler-
ates while cortical activity may increase. Lacey also found that
reaction times (a measure of behavioral arousal) may be correlated to
different fractions of autonomic arousal. Reaction times varied
depending upon where the observer's heart was in its cycle. Fastest
reaction times were observed just before the onset of heart contraction i.e. during the P-wave. Slowest times were associated with the QRS-complex during the onset of systole. Lacey's theory would predict fast reaction time early in the cardiac cycle and slow reaction times late in the cardiac cycle. One problem that arises from this result is that one would need to know when a stimulus is presented in relation to the phase of the cardiac cycle in order to make any predictions about RT.

Predictions concerning behavior may be quite different depending upon which view of arousal theory is accepted. Currently, the trend is to accept the Lacey view of arousal as a complex-multidimensional process. However, there is evidence that does not support Lacey's view. For example, Surwillo (1976) hypothesized that HR would be lower when signals were detected than when they were missed. His results did not confirm the hypothesis and Lacey's theory. Although Lacey's theory has empirical support, so does the traditional view. There are also many studies which don't support either view. Arousal theory is still emerging and being modified and that is one of the difficulties in attempting to evaluate research in terms of these constructs.

**Information Processing/Channel Capacity Models**

There are a number of information processing models that relate to vigilance. The most well-known is that of Broadbent (1958, 1971). Broadbent's model is not only a model for vigilance phenomenon but a general model that describes information processing. Broadbent's model has been classified as a model of attention by Davies and Tune (1969) and an information processing model by Loeb and Allusi (1980). This illustrates an equivalence of the terms "information processing" and
"attention" that is often found. Broadbent's model is often referred to as a model of "selective attention" or a "filter model" or a "limited capacity" model. All of these descriptions refer to functional characteristics of the model. Broadbent theorized that an organism must be selective to stimuli because the nervous system cannot handle the total volume of stimulation that impinges on it at any given moment. Furthermore, the organism must selectively respond to information because, as often has been observed, adequate responses to one set of stimuli may be incompatible with adequate responses to another set. Broadbent's (1971) most recent model is similar to his model proposed in 1958 with a few modifications. Stimuli from the environment are sensed and stored briefly (about 1 second) in a buffer (short-term memory-STM). A filter selects stimuli on the basis of certain priorities. Stimuli that are novel, biologically important or intense receive preference for further processing along a single channel. A store of conditional probabilities of past events (long-term storage) is thought to affect the operation of the filter. Information can be passed back to STM after passage through the limited capacity channel. This information can be held until the period of decay runs out. Information can be continuously recycled in this manner but at the expense of channel capacity. It is believed that visual and auditory stimuli are recycled and held in STM by a process of vocal or sub-vocal rehearsal, i.e. the original stimulus is recoded and returned to STM in a different form and location. This process can continue unless the limited capacity channel receives a priority interrupt and has to process new more relevant information. It is believed that limits on
this recycling process is determined by a fixed number of events rather than a fixed time. Message errors, unreliability and internal and external noise may corrupt the original message. This may be combatted by using error-correcting codes. Three types of filtering occur that are governed by three kinds of rules. First, environmental stimuli pass through a filter and give rise to an evidence state about the stimulus. The evidence state may not be an exact representation of the original stimulus because of noise or errors or limited feature extraction. The original stimulus, in the form of evidence, passes through the single channel and gives rise to a category state. Evidence states are linked to category states by pigeon-holing. Pigeon-holing refers to the manipulation of a set of categories which must be used when any message arrives--some category states may have a bias in their favor whatever the nature of the evidence. Pigeon-holing involves no input selection but is a form of response bias. For example, an evidence state may not contain any information about a certain stimulus feature but through the process of pigeon-holing a feature may be assumed. Thus, an evidence state is linked to a category state and permits a certain response. A third kind of filtering termed "categorizing" occurs. "Categorizing" links stimulus events to category states. A given set of stimulus features are thoroughly analyzed and matched to category states. These processes are shown in Figure 4. Categorizing is both output and input-selective. Filtering is input selective but influences outputs. Pigeon-holing, according to Broadbent is only output-selective but this cannot be entirely true since certain features may have been already selected during filtering and certain
Figure 4. Broadbent's Processes of Categorizing, Filtering, and Pigeon-Holing (From Broadbent, 1971).
assumptions have to be made about stimulus features. Arousal enters Broadbent's model and may cause filtering and pigeon-holing to be more extreme. Motivation is thought to affect only pigeon-holing and not filtering—certain pigeon holes may have a greater tendency to be used due to motivation. Noise has the effect of making a person distractible, i.e. efficiency is impaired by causing a shift of the filter.

The vigilance decrement is the result of shifting rules that link stimuli to evidence states at the filter. The repeated application of the stimulus results in reduced novelty and a restructuring of priorities; some stimulus events gain priority. Broadbent explained Mackworth's early results in terms of a loss of stimulus novelty; spontaneous recovery is explained in terms of a restoration of stimulus novelty. Introduction of a new stimulus such as a telephone message will restore novelty to the original task stimulus. Filter theory predicts that a vigilance decrement is more likely with paced tasks and tasks with short signal durations and high signal and/or event rates. The use of redundant displays should facilitate a pronounced decrement. Broadbent (1971), himself, reviewed the literature and could find only weak support for his theory. Jerison and Wallis (1959) found no vigilance decrement with redundant displays. There is evidence that detections increase with signal rate (Jenkins, 1958). Baker (1963) did lend some support to filter theory, albeit weak. Baker found decrements for signal durations of 0.2, 0.3, 0.4, 0.5, 0.6, and 0.8 seconds but the decrement for a duration of 0.8 seconds was less; moreover, detection rates increased with signal duration. Vigilance decrements, however, have been found in both paced and unpaced tasks. Jerison's work with
background event rate does support Broadbent's model; although Jerison interpreted the larger vigilance decrement observed with a high event rate as due to inhibition. Another aspect of Broadbent's model is that it would predict an increase in response time and response time variability. A consensus of recent opinion is not particularly supportive of Broadbent's 1958 or his revised 1971 model as an explanation for vigilance phenomenon. Loeb and Alluisi (1980) state, "In some ways filter theory might be viewed as a 'black box' analog of habituation theory, as indeed, might the extinction theories of learning (p. 602)." Warm (1977) says, "... it is often difficult to develop clear ad hoc predictions from the model and many of the effects incorporated by the filter model can also be incorporated under other approaches (p. 635)." Davies and Parasuraman (1982) summarize their views of filter theory, "The explanation of the vigilance decrement given by Broadbent's filter theory thus appears to receive only weak empirical support and in any case it is difficult to distinguish the theory from some form of observing response theory in a clear and convincing fashion, since the failure of the filter to select information from the task would seem equivalent to a failure to observe the display on which the task is presented (p. 16)."

Other channel capacity models have been reported in the literature in various forms. For example, Elliot (1960) developed a model referred to as the "Elliot perceptual model." The model is very similar to Broadbent's model in its elements, functions, and predictions. Kahneman (1973) outlines a capacity model for vigilance as well as information processing in general. Kahneman's model incorporates a
Yerkes-Dodson notion of arousal into his model with some of Easterbrook's ideas (Easterbrook, 1959). Easterbrook's hypothesis was that perceptual narrowing occurs under high arousal. This is also similar to Broadbent's idea that stress would cause filtering to be more extreme. Still other capacity models have been reported, e.g. Norman and Bobrow (1975) have reported on resource-limited and data-limited processes. Resource limited processes and transitional processes are in line with the ideas of time-sharing and capacity models. Such models might be appropriate to high signal and event rates and overarousal whereas data-limited processes may describe the vigilance situation and underarousal. This model may explain why some vigilance decrements have been observed at high event rates and low signal rates. Fisk and Schneider (1980) performed three experiments that examined vigilance performance in terms of automatic and controlled processing. They found significant decrements in controlled processing tasks but not in automatic processing tasks. Controlled processing tasks are demanding of attentional resources. These are usually high load, serial, easy to alter once learned, limited-comparison rate tasks. Automatic processing tasks are relatively well-learned in long-term memory, demanding of attention only when a target is detected, parallel rather than serial and difficult to alter, ignore or suppress once learned. They are virtually unaffected by load. In the second experiment they found that the vigilance decrement decreased if the number of channels was decreased and memory load was increased. The vigilance decrement was maximal when subjects were required to continually and consistently
allocate controlled-processing resources to display locations. The authors rejected the neural habituation hypothesis of Mackworth (1969).

Broadbent's model and capacity models are not particularly relevant to the present research because of the nature of the models and the nature of the research. Broadbent's model was used to describe vigilance results that Mackworth's model failed to predict, namely that performance improved with higher signal frequencies. Stroh (1971) suggests that monitoring tasks with high signal rates are not truly vigilance tasks. The present research used low signal frequencies. Moreover, Broadbent said little about activation in 1958 but did acknowledge it as a necessary process in 1971. He said little about the nature of the process or how it enters into vigilance performance.

**Expectancy Theory**

Several different theories already discussed have used an expectancy construct to explain results. For example, Mackworth (1950) explained incomplete extinction on the basis that KOR received during training trials was replaced by expectancy reinforcement. Craig's (1978) probability matching model can be considered a variation of expectancy theory. Deese (1955), an advocate of the expectancy position argues: (1) vigilance is an excitatory process not an inhibitory process as suggested by Mackworth, (2) the observer's expectancy or prediction about future events is shaped by his previous experience with the task, (3) the observer's level of expectancy determines his vigilance level and his probability of detection, and (4) feedback from the task determines what the observer expects from further participation in the task—his vigilance will be proportional to his
expectations. According to Deese, detection for a given signal will be determined by the average of all intersignal intervals preceding the present one. The observer's ability to estimate error is free of constant error and shows relatively low variable error. Expectancy should be low immediately after a signal and high around the mean intersignal interval. The model assumes the observer is continually averaging previous data to update current expectations. Deese equated motivation to participate in a task with individual differences. Inhibition was not considered a viable explanation for vigilance performance which sometimes rises or fluctuates instead of falling. Moreover, end spurt sometimes occurs in visual search tasks if the observer can estimate the task's end. Deese postulated an arousal-like mechanism in his expectancy theory. He asserted that the waking center in the hypothalamus must be maintained at a high excitatory level in order to maintain performance. There is also a place for personality measures in Deese's theory. Much of Deese's outline of expectancy theory in 1955 was not empirical but based on speculation. The value of his outline is that it suggested plausible hypotheses to test.

Baker (1959, 1960a, 1960b, 1961, 1963a, 1963b, 1963c) performed a number of experiments to test and outline to a greater extent expectancy theory. Baker (1959) investigated the effect of signal regularity on vigilance performance. He defined signal regularity in terms of the range of intersignal intervals. Regular signals had a mean ISI of 2 minutes with a range of 36-196 seconds. Baker performed five experiments using this schedule. A sixth experiment used an irregular schedule with a mean of 2.5 minutes with a range of 45-645 seconds. In
four of the first five experiments there were no decrements in time. The sixth experiments with irregular signals did display a marked performance decrement. Baker concluded that observers must have had expectancies which were not reinforced which resulted in performance dropping to a low level; this accounted for the decrement. One problem with this research is that Baker does not offer a reason for expectancies being initially high. Furthermore, he mentions nothing about instructions or practice trials.

Baker (1960a) performed an experiment with artificial signals and KOR as experimental conditions. Artificial signals are signals that are identical to real signals. The signal was a 2 mm dot of light lasting 0.6 seconds. It appeared in one of the four corners of a 4 inch square in random order of sequence. Baker used a Mackworth signal schedule for a control condition. The task duration was 1½ hours. He superimposed the artificial signals on the Mackworth schedule with ISIs of 2½, 1-3/4, 1½, 2, and 2½ minutes in that order as many times as necessary to fill in the temporal gaps between real signals. Signal frequency varied from 58-63 over a session. KOR in the form of "correct," or "missed one," was provided after each event. Performance was significantly better for the experimental condition than for the control condition. The condition KOR plus artificial signals was compared with Mackworth's KOR results and found to give superior performance. However, the result could not be checked statistically and KOR was confounded with the effects of artificial signals and signal frequency. The experimental condition contained a greater number of signals than
the control condition. One cannot conclude that KOR caused an effect—
it may have been from differential signal frequencies in the two tasks.

Baker (1961) performed an experiment in which performance on a
central 1½ hour vigilance task was maintained at a high level by giving
KOR on a secondary task. Subjects monitored visual signals (2 mm light
spot, 0.6 second duration) for the main task and changes in ambient
lighting levels or background noise for the subsidiary task. Feedback
in the form of response latencies for the subsidiary task were provided
to experimental groups. He used a Mackworth signal schedule for the
main task randomized for each subject. Five secondary signals were
inserted during periods when there was sufficient time between central
signals to ensure detection of the subsidiary signal. Second signals
were not closer than six minutes and they came at different times for
each subject. Baker found significant differences between the experi­
mental and control conditions (no KOR). The implication is that obser­
vers can better structure expectancies for signals in a central task
when provided feedback on a secondary task. Details of how this occurs
were not provided.

Baker (1963a) performed an experiment which has been previously
mentioned with regards to signal duration. He hypothesized that the
shorter the signal in a vigilance task, the poorer the initial detect­
ion performance and the steeper the decrement in detection perform­
ance. Baker used a clock test where a signal was the brief stoppage of
the clock hand. The hand rotated once every minute. Baker used signal
durations of 0.2, 0.3, 0.4, 0.6, and 0.8 seconds. He found a signi­
ficant decrement and significant difference between signal conditions
which he said expectancy theory would predict. It is not clear from reading this report what expectancy theory means to Baker and what links exist between signal duration and expectancy theory. It is not clear why you need an expectancy theory to predict that an observer will be more likely to detect a signal the longer it is displayed.

Baker (1963b) performed an experiment to study the consistency of performance of subjects within and across visual vigilance tasks. He used two visual tasks, one was a clock test and the other a needle-deflection or dial test. A signal on the clock test was defined as a 0.4 second hesitation of the hand. A signal on the meter was a brief, .25 second excursion of the needle to the right from a reading of 100 to a reading of 125. The needle quivered around a value of 100 ± 10 units except for occasions when a signal was presented. Experiment I examined the reliability of performance for the two tasks within a run and across days. He found significant correlations between total signals detected in half-hour run segments for both tasks. Baker found consistency within a particular task but this consistency was not strong across days. He also found that you could not predict performance on the clock test from the needle deflection task. Baker also found that response time was not correlated with detections or ISI. According to Baker, reinforcement theory would predict that the probability that an observer will predict a particular signal depends upon the length of time that has elapsed since the immediately preceding signal was detected. The expectancy position would argue that the probability of detection of a signal depends more on the prevailing mean signal rate.
Baker (1963c) presented some new data and summarized earlier data to outline a theory of vigilance based on expectancy. Baker (1963c, p. 128) defines expectancy as follows:

... the expectancy hypothesis is a statement that the probability of detection of a signal in a vigilance task is greatest when the signal occurs after an interval which is equivalent to the mean of the intersignal intervals preceding the interval in question: detection probability is low immediately after a signal, increases as the mean of the intersignal interval of the preceding series is approached, and if not reinforced by the occurrence of a signal, again decreases ... when observers are exposed to a temporal series of signals, they expect the next signal in the series to occur at or about the mean temporal interval separating the past signals. Such expectancies do not necessarily imply any conscious formulations by observers: the degree to which they are aware of expectancies is unknown. With respect to the mean intersignal interval, observers are not precision-time measuring instruments. The consequence is that the peak of the expectancy curve is flattened to some degree, i.e. there is an interval of uncertainty (in addition to a constant error). Because of this uncertainty the decrease in expectancy after the mean intersignal [interval] has elapsed is a more gradual phenomenon than the rise in expectancy before the intersignal interval has been reached.

Baker performed a study that replicated earlier work done by Mower (1940) for the purpose of describing the expectancy function. Baker used a series of 20 signals with constant ISI equal to 10 seconds. The 21st signal had an ISI of 2, 5, 20, 25, or 30 seconds on separate trials. A plot of mean RT versus time of signal appearance shows RT to be at a peak for an ISI of 2 seconds and then declines to the mean of the series at an ISI of 10 seconds. Deese's (1955) version of expectancy predicted that expectancy increased up to the value of the mean ISI and beyond, i.e. probability of detection would be below average when an ISI was less than the mean interval and equal to or greater than the average probability of detection when the ISI was equal to or
greater than the mean. An interesting aspect of his data that is not stressed is that with ISI's greater than 10 seconds, a definite increasing trend of RT versus ISI appears in the data.

Baker (1963c) reported another study where he had subjects repeatedly reproduce a 2 minute interval over a 1 hour watch. There was no correlation between the observer's estimate of the length standard interval and the actual length of standard interval. The interval was consistently overestimated in every instance (N = 11). Estimates ranged from 2½ minutes to 7 minutes. Baker asserted that variability was small, however, and observers could generate a series of signals with fair regularity over time. The data presented by Baker suggested just the opposite. Standard deviations of interval estimates ranged from 11 seconds to 120 seconds. Baker's data suggested that observers did not estimate a 2 minute time interval very well.

Baker, in the same report, discussed the results of an experiment whereby he tried to determine where in a series a person begins to average ISI's to structure expectancies. He used the Mackworth series as well as several less variable versions of it and asked subjects to inject two signals at the end of the series that predicted the arrival of the next signals. The mean ISI's for the series were either 150 seconds or 100 seconds. Observers overestimated the longer interval by about 3%. Subjects tended to underestimate the shorter interval by about 12%. These results are a reversal of the usual time estimation findings. In addition to this analysis, Baker correlated the estimated mean interval lengths produced by observers with past intervals--much
like autocorrelation coefficients. Although he did not find signifi-
cant individual correlation coefficients, he did report a trend in
correlation coefficients that went from positive to negative the fur-
ther back in the series that the intervals went. He concluded that
this indicated that observers were extrapolating to the future based on
5-7 preceding intervals. Baker's conclusions don't appear to be justi-
ﬁed by the data since none of the autocorrelations were signiﬁcant.

Baker (1963c) summarized the effects of variables that have an
effect on vigilance performance in terms of expectancy theory. Baker
claimed the following: (1) increases in signal rate generally improves
performance because observers have a larger sample of data on which to
base expectancies and they can observe shorter intervals with greater
precision, (2) less variable intersignal intervals mean a more probable
conﬁrmation of expectancy on the part of the observer, (3) signals
that are large, intense, or long tend to make conﬁrmation of expec-
tancy more probable, (4) KOR serves to establish and conﬁrm the se-
quential nature of a series, (5) environmental factors, e.g. heat,
noise, etc. tend to distract the observer and compete for attention
with the signal, thereby lowering the perceived signal frequency and
performance, (6) a task that involves visual search (spatial uncer-
tainty) results in unconfirmed expectancies through the mechanism of
lowering apparent signal frequency; performance is lowered, (7) periods
of rest result in a spontaneous recovery and performance returns to an
initial level--the more irregular a series, the more difﬁcult it is to
remember; the spontaneous recovery is due to forgetting of a series
that was never well-learned, (8) motivation level will determine ini-
tial performance level and may expedite or postpone the vigilance
decrement; however, expectancy theory is valid regardless of a person's
motivation, (9) artificial signals added to a task increase interval
regularity and improves performance, and (10) performance on a central
vigilance task can be improved by giving KOR on a peripheral vigilance
task—Baker interpreted this effect as due to a purely motivational
factor, KOR gives the observer feedback about his alertness thereby
increasing motivation to perform better on the central task.

Discussants at the vigilance symposium took Baker to task on
several points: (1) forgetting was not accepted as a plausible expla-
nation of spontaneous recovery with rest, (2) expectancy theory expla-
nations of the vigilance decrement in terms of a "vicious circle"
explanation was considered superficial and doesn't explain situations
where the vigilance decrement does not occur, (3) expectancy theory
does not explain initial performance levels and (4) expectancies are
reshaped, according to expectancy theorists, as the results of missed
signals and this results in the vigilance decrements; however, a vigi-
lance decrement has been observed with one signal presented over a
watch.

Bergum and Lehr (1963) tested the hypothesis that knowledge versus
no knowledge of the length of a vigil should yield differential results
in the form of an end-spurt for the informed group. The experimenters
did find a significant end-spurt effect for the experimental group.
The authors suggested the results could be interpreted in three dif-
ferent models: (1) motivation-expectancy, (2) activation, and (3)
goal-gradient. The authors favored the last model which suggested that observers were reinforced to perform better by anticipation of the end of the task.

Faulker (1962) performed an experiment in which he found that dummy signals which occurred at semi-regular intervals were more effective in reducing subject response time variability than those which occurred at nonregular intervals. These results would be predicted by expectancy theory.

Wilkinson (1961) found results that were not supportive of expectancy theory. Wilkinson studied four different kinds of tasks: paced regular, paced irregular, unpaced, and continuous display. Wilkinson found a vigilance decrement in all tasks but not differential effects. Expectancy theory would predict better performance for regular tasks.

McFarland and Halcomb (1970) examined the effects of signal probability in a training task on a subsequent experimental task. Subjects received pretask signals with a probability of $p = 0.18$ or $p = 0.02$ in an auditory detection task. The experimental task was a visual task with the same signal probabilities. The authors used a split-plot factorial analysis of variance and found significant effects. Results suggest that an expectancy set is created on the pretask and that this set carries over to the experimental task to enhance performance. The results are supportive of expectancy theory.

Colquhoun and Baddeley (1964) also examined the role of the pretest expectancy on the vigilance decrement. Signal probabilities in a practice task were the same as used by McFarland and Halcomb (1964). Colquhoun and Baddeley tested three hypotheses suggested by expectancy theory.
theory: (1) the vigilance decrement is due to a drop in the S's expectations about initial signal probabilities, (2) pretest practice at a task sets the observer's expectancies concerning signal probabilities, and (3) the degree of decrement is a function of the difference between pretest and test signal probabilities. The experimenters found support for their hypotheses which they concluded supported expectancy theory.

Boulter and Adams (1963) sought to replicate Baker's finding concerning the effects of ISI variability on performance. Boulter and Adams had three different degrees of signal uncertainty in signal schedules: high uncertainty (ISI = 15, 15, 30, 30, 60, 120, 120, 300, 420, 600, and 900 sec.) medium uncertainty (ISI = 120, 120, 220, 220, 220, 220, 220, 220, 270, 270, 270, and 270 sec.) and low uncertainty (ISI = 220 sec. for all 12 signals). The mean ISI was the same for all conditions. Subjects had to detect a two-digit number which appeared on a display for 5 sec. over a three-hour vigil. The authors found a significant decrement (increases in response time) across all conditions but not differential effects. They analyzed the data using a Lindquist Type I (A x B design) where temporal uncertainty was a between subjects variable and trials a within-subjects variable. This was a split-plot factorial design, split on temporal uncertainty. The authors also found that response time increased with ISI for the high uncertainty condition and decreased with ISI for the medium uncertainty condition. The finding of response time as a decreasing function of ISI for the medium uncertainty condition supported Deese and Baker's expectancy hypothesis. However, increasing response time as a function
of ISI does not support an expectancy hypothesis. The authors concluded, "... the expectancy function in a monitoring task is no single relationship, and that presently hypothesized functions for expectancy are oversimplifications ... a family of functions would seem clearly indicated from our data (p. 208)."

It should be noted and emphasized that the increasing function of response latency found in the present research is similar to Boulter and Adam's function under the high uncertainty condition. Moreover, up to about 300 seconds, their signal schedule is similar to the one used in the present research.

Expectancy theory is a non-physiologically based, strategic theory (as opposed to a state theory such as arousal theory) that explains the vigilance decrements in terms of monitors expecting a much higher signal rate than they encounter; the discrepancy between expectancy and reality produces the decrement, i.e. expectancies are not reinforced. Some of the notions are fuzzy and not well-defined. Expectancy assumes that subjects are fairly precise estimators of time intervals; the evidence is equivocal or even counter to this idea. Neveryitch (1982) has shown that when Ss try to estimate intervals of 60, 120, and 300 seconds they overestimate the shorter interval and underestimate the longer intervals (N = 16 for each condition, \( \bar{x}_{60} = 62, SD_{60} = 15.8; \bar{x}_{120} = 113.4, SD_{120} = 35.6; \) and \( \bar{x}_{300} = 268.5, SD_{300} = 67 \) seconds)

There is a suggestion that observers estimate shorter intervals with more precision than longer intervals.

Davies and Parasuraman (1982) recognize that pre-task signal probabilities may be a significant determinant of performance in a
subsequent experimental task. The authors assert that when signal frequencies are high on a training task and not on a subsequent experimental task, Ss revise their criterion (b) upwards towards greater strictness. Krulewitz, Warm, and Wohl (1975), however, have shown that the shift from a high to a low event rate enhanced the probability of signal detection while the shift from low to a high event rate depressed performance on the experimental task. Krulewitz, et al., did not, however, calculate criterion shifts.

Expectancy theory, like arousal theory is an intuitively appealing theory. It is easy to believe that observers can learn to expect the arrival of a signal under some set of conditions. The conditions under which this occurs and the exact relationships that exist between variables is not known at this time. It is likely that expectancy theory in some form will be a part of a comprehensive vigilance theory.

Habituation Theory

Habituation theory is a neurological or state model closely related to arousal theory. Until recently, habituation has been the concern primarily of investigators interested in brain research. For example, Callaway (1975) discusses habituation in terms of brain (EEG, AEP) responses to stimuli and recovery cycles. Horn (1969) discusses neural mechanisms that underlie habituation at a neuronal level. Lynn (1966) discusses the orientation reaction. It is the reaction that accompanies a novel stimulus and is often called the "what is it?" reflex. The response is complex and a host of physiological, cortical and behavioral changes accompany this reflex. If a stimulus is repeated a number of times the orientation reaction subsides. This
process is called habituation. Previously, when filter models were discussed, a process like habituation occurred with the filter; as signals were repeated and lost their novelty, other more novel signals gained priority in the gating mechanism. Sharpless and Jasper (1956) were some of the first researchers to suggest that a strong stimulus (such as a telephone call) applied to a drowsy person could restore alertness. The effects of this arousing stimulus may last for several hours. Mackworth (1968a, 1968b, 1969), one of the early advocates of habituation theory, discussed vigilance performance and compared observing behavior, signal detection, expectancy, arousal and habituation theories. Later, Mackworth (1969) in her book Vigilance and Habituation: A Neurophysical Approach outlined, to a greater extent, the basics of a theory based on habituation. Mackworth (1968b) put forth a thesis regarding habituation and vigilance:

... the decrement in these learned responses is related to habituation of two kinds of innate neural response: (a) the arousal response (or alpha block) occurring in the spontaneous rhythms of the brain, and (b) the evoked potentials produced by the repetitive, continuous, or unchanging stimuli that constitute the background events of the task. The subsidiary part of the thesis suggests that (a) as a result of habituation of the arousal response, sensitivity decreases, while (b) as a result of habituation of the evoked potential, positive motor responses to both the background events (false alarms) and to the signals (correct detections) will decrease. It is believed that the first process (a) is more important in tasks in which the background event rate is very fast or continuous, while the second process (b) is more important in tasks in which the background event rate is discrete and fairly slow (two events or fewer per second) (p. 308).

Mackworth (1968b, 1969) has drawn much of habituation theory as it relates to vigilance from Thompson and Spencer (1966), Lynn (1966), and
a Russian, Sokolov (1963). Groves and Thompson (1970) summarize a considerable amount of research on habituation, discuss the theories that have been used to describe habituation and propose a dual-process theory that they claim satisfies the requirements for a theory. Their dual-process theory assumes that habituation consists of two basic processes, habituation and sensitization. "The strength of the behavioral response elicited by a repeated stimulus is the net outcome of the two independent processes. . . (Groves and Thompson, 1970, p. 442)." Habituation is the decreased response to repeated stimulation. Sensitization refers to long-term increments in response. Sensitization is often treated as disinhibition but disinhibition occurs to a habituated response whereas sensitization need not. Sensitization is an increased responsivness. Habituation is thought to occur in the S-R pathway whereas sensitization occurs in the collection of pathways, systems, and regions that determines the general level of responsiveness of the organism. It is believed that habituation is a central process that involves interneuron processes. Sensitization is believed to be a similar process. In addition, it is accomplished by increases in motorneuron excitability. Studies have shown there are three types of interneurons: a nonplastic interneuron showing no changes in response and two plastic interneurons; one specific to habituation (type H) and one specific to sensitization (type S). It is not clear at this time exactly how the synapses function during habituation and sensitization.
A theory of habituation and sensitization must explain:

1. parameters of sensitization/habituation
   a. time course and recovery
   b. effect of repeated series
   c. effect of stimulus frequency
   d. effects of stimulus intensity
   e. stimulus generalization
   f. below-zero habituation
   g. dishabituation of response sensitization
   h. habituation of dishabituation of response sensitization.

2. occurrence of sensitization

3. independence of habituation and sensitization

4. incremental stimulus intensity effect

5. "missing stimulus" effect

6. gross changes in state such as sleep.

Several theories have been advanced to explain the above requirements. The models tend to be grouped into two broad categories: one-stage models and two-stage models. One-stage models did not satisfy many of the requirements for a theory and so two-stage models have been advanced recently.

Some of the models that have been advanced to explain habituation of the orienting response are: (1) classical conditioning (Pavlov), (2) Sharpless and Jasper selective habituation, (3) Gastaut's habituation rate model, (4) Roitbak's non-specific thalamic reticular system theory, (5) Sokolov stimulus-model comparator model, (6) Hernández-Péon afferent neuronal inhibition model, (7) Moruzzi central activating and inhibitory systems model, (8) Grastyan model, (9) Jorvet model, (10) Carlton cholinergic inhibition model, (11) Pibram limbic forebrain function model, and (12) Grove and Thompson dual-process theory.
The Sokolov and Grove and Thompson two-stage models are the ones that have provided the greatest amount of input to vigilance theory. Features of the Grove and Thompson model are summarized below (Grove and Thompson (1970) provided empirical support for their model—these references are not included here. References to habituation models are given by Lynn (1966) and Grove and Thompson, (1970)).

1. Every stimulus that evokes a behavioral response has two properties: it elicits a response and influences the "state" of the organism. The S-R pathway is the most direct route through the central nervous system from stimulus to discrete motor response. State is the general level of excitation, arousal, activation, tendency to respond, etc. of the organism—state need not be a unitary entity. The discrete motor pathways and general state of excitability interact to produce a behavioral response.

2. Repetition of an effective stimulus results in an inferred decremental process in the S-R pathway which is termed habituation.
   (a) During habituation training, habituation develops exponentially and reaches an asymptotic level.
   (b) Rate of development and relative degree of habituation are directly related to stimulus frequency and inversely related to stimulus intensity. Amount of absolute habituation is directly related to the intensity of the habituation stimulus. Frequency has a strong effect and intensity a weak effect on habituation.
   (c) Upon cessation of the habituation stimulus, habituation decays spontaneously (spontaneous recovery).
   (d) Repeated series of habituation training and spontaneous recovery result in progressively more habituation.
(e) Response habituation will exhibit generalizability to a test stimulus to the extent that the habituating and test stimuli activate common habituation elements.

3. Presentation of an effective stimulus results in an inferred incremental process in a state of excitation or tendency to respond of the organism which may be termed sensitization.

(a) The process of sensitization occurs in state systems, but not in S-R pathways.

(b) During habituation training, sensitization first grows and then decays.

(c) The amount and duration of sensitization are directly related to stimulus intensity. At higher intensities, sensitization is directly related to stimulus frequency. At low frequencies there may be little or no sensitization.

(d) Upon cessation of a stimulus that has produced sensitization, sensitization decays spontaneously.

(e) Repeated presentations of a sensitizing stimulus results in progressively less sensitization, that is, sensitization decreases or habituates. Habituation of sensitization below initial control level may be due to interaction with elements within the S-R pathway or may be a property of state systems. The extent to which sensitization habituates below control level may also depend in part on the initial baseline state of the organism and the conditions of measurement.

(f) Response sensitization will exhibit generalization to a test stimulus to the extent that the sensitization elements.

(g) Dishabituation, the increase in a habituated response following presentation of a stimulus other than the habituation stimulus, is simply an instance of sensitization.
(h) Under certain circumstances (strong stimulus presented regularly at relatively slow rate) temporal conditioning of sensitization of state may occur.

4. The two processes of habituation and sensitization occur and develop independently of one another but interact to yield the final response output function. Habituation may be primarily "phasic" in its action on response output, while sensitization may be primarily "tonic" (pp. 440-441).

The work of Grove and Thompson has a base in psychophysiology. The empirical base has evolved, primarily, out of animal experiments. However, Sokolov has studied the orienting response and habituation in humans. Mackworth has attempted to marry psychobiology and psychophysiology to vigilance. The habituation model is based on observations of S-R behavior as well as state behavior. The model addresses the question of "How do observers respond to repeated stimulation?" Novelty of a stimulus has a central part in the theory. A novel stimulus, of sufficient strength, invokes the orientation reaction. Repeated application of the same stimulus leads to habituation--a tendency to weight less the value of the stimulus. The habituation model is like Broadbent's filter model in this respect--novel stimuli have favor with the filter. The model goes beyond Broadbent's model in that it attempts to specify and detail the mechanism by which a stimulus loses novelty and under what conditions. Some researchers believe that habituation and sensitization are processes that have evolved that are keyed to survival.

One problem that occurs when trying to use habituation theory to describe vigilance behavior has to do with the low between session correlations of performance that have been observed in vigilance
Sokolov (1963) has postulated that a model of repetitive stimuli is formed in the cortex. Incoming stimuli are compared to the model and if the signal matches the model, the reticular formation is inhibited and responses strength decreases. If a novel stimulus arrives; a mismatch results and the reticular formation is activated. This suggests that cortex activity mediates reticular activity and that higher and lower brain functions have intimate feedback mechanisms. The model for repetitive stimuli creates expectancies for future stimulation. These expectancies act as stimuli and arouse the observer. It is believed that the contingent negative variation wave (CNV) discovered by Walter et al., (1964) sensitizes the cortex to expected stimuli. Habituation theory encompasses arousal (orientation reaction), habituation to repeated stimuli, and expectancy. It is not possible, however, at this time to unravel the three basic theories.

Mackworth (1968b, 1969) has summarized the relationships and lack of relationships between habituation theory and vigilance performance:

1. Habituation depends on repetition of a stimulus or series of stimuli. Whether a decrement occurs depends on the nature of the task. For example, complex tasks and visual scanning tasks may provide arousal that offsets decrement. Decrements may also be masked by learning.

2. Vigilance decrement rates are much slower than habituation rates. Mackworth (1969) suggests that this results from stimuli in vigilance tasks having signal properties and expectancy.
(3) The rate of change of habituation is a function of background event rates. Decrement is more pronounced with higher background event rates. Jerison (1967a) also found performance independent of signal probability at low background event rates but at high background event rates, higher signal probabilities resulted in slightly improved performance. Improvements, however, never achieved the level of performance that was observed with low background event rates.

(4) Presentation of another strong stimulus results in recovery of the habituated response, e.g. a telephone call or KOR. Irrelevant stimuli such as noise may improve performance initially (novelty effect) but detract from it later (habituation effect). This may explain some of the conflicting research results with vigilance and noise.

(5) There are marked individual differences in behavioral, cortical and behavioral responses in habituation patterns and vigilance performance. The evoked potential has been correlated with detections and EEG alpha and theta rhythms, electrodermal responses, and catecholamine production has been correlated with vigilance performance.

(6) The rate of habituation of the arousal response may be decreased by amphetamines and increased by depresssant drugs. Mackworth asserts that decreases in $d'$ are related to habituation of the arousal response while increases in $\beta$ are related to habituation of the evoked potential. The evidence is not conclusive. Amphetamines appear to halt declines in $d'$ (since detection rate and false alarm rates do not change) for high-event rate successive discrimination tasks. Amphetamines have been shown to prevent increases in $\beta$ for low-event rate successive discrimination tasks. Depressant drugs have been shown to either have no effect or to impair performance. In some instances the false alarm rate has increased but not the detection rate. This would suggest a decrease in $d'$. The exact mechanism by which this occurs is not known.

(7) Changes in the pattern of regular repetitive stimulation may result in dishabituation. Habituation research has shown that even missing signals in a regular sequence of signals can produce an orientation reaction. Rest pauses produce dishabituation as do novel stimuli.

(8) Sleep loss has an effect on habituation. False alarms tend to decrease with moderate sleep loss but increases with more severe sleep loss. Noise and KOR tend to counteract sleep loss effects.
(9) Temperature effects exist with habituation. Temperatures may provide distractions when they depart from accustomed values. Temperature effects may interact with sleep loss and the effects cannot be entirely explained in terms of arousal. Temperature effects are complex and not completely understood.

(10) Mackworth suggests that tasks requiring active head movements or that require difficult or complex decisions may show less decrement than tasks that feature single-source signals and require simple decisions. Adams and Boulter's (1962) study with tasks requiring head and eye movements and those not requiring such movements did not provide support for this notion. Jerison's three clock task did not show a decrement, as did his one clock task. Complexity of a task may be a determinant of habituation.

Mackworth (1969, p. 198) states, "... it would seem that at this time it is better to determine as widely as possible the various concomitant changes in the physiological and psychological responses of the subject, rather than trying to distinguish between one theory and another." To add to Mackworth's statement, it may not be possible to distinguish between habituation, arousal and expectancy because the exact mechanisms and processes have not yet been worked out. A comprehensive theory of vigilance will require a thorough understanding of research results on a number of fronts, e.g. brain research, psychobiology and psychophysiology and psychology.

Motivation Theory

There are few explicit motivational theories in vigilance research. However, nearly all vigilance theories have an implicit motivational component. For example, a person's motivation is represented by his criterion in SDT. Where he places his criterion can be influenced by payoff structures. Jerison theory had the observer attending to the
display in three states: alert, blurred or distracted. The quality of observing was determined by the observer's utility matrix which could change over time. Classical conditioning theory viewed motivation as reinforcement that was received during training trials. Holland's observing response theory also viewed reinforcement as motivation. Detected signals had intrinsic motivational properties which upheld behavior. Expectancy theory viewed anticipation of an event as motivating. Arousal theory equated arousal with motivation. Habituation theory viewed the repeated signal as having properties that reduced performance. Broadbent viewed drives and motives as having an effect on pigeon-holes; they did not affect filtering but did affect the tendency to use certain pigeon-holes.

Eysenck (1982) claims:

While no satisfactory theory of the effects of incentive on performance is currently available, we can at least identify some of the major factors which are involved . . . the effects of incentive on task performance can usefully be regarded as depending on four classes of variables: (1) the nature of the incentive, (2) the processes required by the task, (3) the aspects of performance selected for measurement and (4) individual differences in mood or state and in semi-permanent personality characteristics. Such a statement is so non-controversial as to be almost banal, yet its implications have been almost totally ignored. So far as the author is aware, there has never been an empirical study in which all four factors were manipulated simultaneously; in fact, it is usual for only one or two of them to be considered within the confines of a single experiment (p. 87).

Wiener (1975) echoes Eysenck's statements but in reference to inspection tasks: "There has been little investigation of motivational variables in inspection, which is surprising, in view of the popularity that motivational constructs enjoy in industry. There is almost
nothing which a critical reader would regard as an adequately con-
trolled study (p. 110)."

The vigilance task has little to offer in the way of intrinsic
motivation. The task is usually monotonous, unstimulating and provides
little to hold an observer's attention or interest. Because of the
lack of intrinsic motivation, some researchers have performed studies
to manipulate extrinsic motivation. These studies have examined KOR,
pree-task instructions practice, rewards and punishment, financial
incentives, attitudes towards experimenters and laboratory versus field
comparisons.

Several reviews are available of motivational theories and the
effects of motivational variables on vigilance performance (Loeb and
Below is a summary of some of the effects of motivational variables
that have been found:

1. Knowledge of results (KOR)--KOR abolishes the vigilance
decrement (Mackworth, 1950). False KOR prevents an
increase in response latency (Loeb and Schmidt, 1963).
This suggests that KOR effects are more motivational
than informative. Davies and Parasuraman (1982) have
reported that verbal KOR is superior to nonverbal KOR.
Auditory KOR tends to be superior to visual KOR. Par-
tial reinforcement is as effective as complete reinforce-
ment if provided at least 50% of the time (Loeb and
Alluisi, 1977 conclude just the opposite; complete KOR
is more effective than partial KOR). Amount of KOR
seems to differentially affect decision responses;
complete KOR given for hits, FAs, misses, and CRs pro-
duces more correct detections and fewer FAs than does
partial KOR. Some beneficial carryover effects have
been found from one day to the next.

2. Practice--Two types of effects have been reported in the
literature. First, changes in vigilance performance
across sessions have been reported (Binford and Loeb,
1966; Wiener, 1962). Binford and Loeb reported hits to
be stable or increase slightly across sessions whereas
d' increases and β decreases others have found the decrement to still exist across sessions despite practice efforts. Wiener found that FAs decreased across sessions. The other effect that has been reported is usually interpreted in terms of expectancy theory; experience gained in a practice task affects performance in a subsequent task. Performance is enhanced if the practice task is similar to the experimental task; however, relationships between pretask probabilities and experimental task performance may break down for low signal rates (Krulowitz, et al., 1975).

3. Rewards and payoffs--The results are mixed. Wiener (1969) suggests, "... the experimental results [of financial incentives] have been unimpressive (p. 628)." Davies and Tune (1969) conclude that financial incentives have an effect if sufficiently large. Withdrawal of rewards on a task after being given in a previous task leads to impaired performance. Combinations of reinforcement (money) and punishment (mild shock for errors) are effective in maintaining performance (Bevan and Turner, 1965). The effects of financial incentives are not entirely clear; some studies have shown positive effects; while others have shown no effects. Levine (1966) found that d' was unchanged with increases in the costs of errors and performance declined (β increased). Performance was unchanged by changes in the payoffs for correct detections. Others have found similar results.

4. Laboratory versus operational settings--vigilance decrements tend to be a laboratory finding. Nachreiner (1977) found a vigilance decrement in a study where subjects were told they would be participating in a vigilance task. Other subjects were given the same task but instructed that they were screening for a job; no decrement was found. Laboratory subjects are generally untrained. Laboratory task durations without interruption are generally longer than for operational tasks. Signals in military settings are not usually weak. Moreover, signals in operational settings are inferred from patterns of information representing a system's status. Probability of a signal is usually fixed in the laboratory but variable in operational settings. Wiener (1973, 1974) and Wiener and Keeler (1975) have worked with adaptive vigilance tasks where signal magnitude is adjusted up or down depending on a subject's performance. One finding is that subjects require an increasing signal magnitude to maintain a stationary detection rate. Vickers, Leary and Barnes (1977) reported on the adaptation by subjects to decreasing signal probability. The researchers found a change in β but not d' over a session. The probability of reporting a hit and FA
increased over time whereas the probability of misses and CRs declined. Target dimensionality is usually unidimensional in the lab and multi-dimensional in applied settings. Responses tend to be more complex in field settings. Field settings usually do not have observers working in isolation.

Motivational variables have been shown to have a significant affect on vigilance performance. In fact, some researchers have attributed the vigilance decrement to the absence of operational setting variables in laboratory research, i.e. the vigilance decrement is an artifact of laboratory research. There is some support for this notion. Teichner (1974) examined 37 vigilance studies that reported a vigilance decrement. Only two studies closely simulated a radar task. Neither of these studies showed a decrement. The decrement function, according to Teichner, is more presumed than real.

Because of some of the above observations, some theorists have been inclined to favor a strong motivational theory of vigilance performance. Two theorists have formulated theories based on motivational constructs: (1) Smith's motivational theory--interaction of monotony and motivation and (2) McCormack's inhibition--motivation theory.

Smith (1966) defined motivation as: "...a state of goal-directed behavior induced by certain internal and external influences... its degree may be inferred from an analysis of performance data; its type from an analysis of independent variables (Smith, 1966, p. 8)." Smith asserted that Ss make FAs because they want the experimenter to think they are fully participating in the task; other observers make errors because they reject the experiment altogether. Signal rate is thought to improve detection rate by decreasing monotony.
Smith believed that monotony was related to task duration and the number of noncritical events. Vigilance tasks are always monotonous, according to Smith, because they demand few if any higher mental acts and because they are prolonged and repetitive. Smith implies that FAs are the result of intrinsic subject motivation variables and not task factors. Smith's theory has some difficulty handling such factors as the effects of drugs which are perhaps best explained by arousal theories. Other factors such as noise, temperature and secondary task effects are not easily explained in terms of interactions between monotony and motivation.

An early motivational theory was expressed by McCormack. McCormack's theory was based on a limited set of experiments and can be considered to be an ad hoc theory. McCormack's theory says nothing of individual differences. McCormack's theory had a reinforcement/inhibition dimension and a motivational dimension. His theory evolved from a series of experiments that had subjects depress a microswitch whenever a 15 watt light bulb came on. The light was on for 100 msec. with ISIs of 30, 45, 60, 75, and 90 sec presented to all Ss every 5 min. in a random order. McCormack's dependent variable was RT and his independent variables were ISI, KOR, and the sudden injection of an extraneous stimulus into the task. These IVs were studied over seven different experiments. McCormack postulated a relationship that was: Reaction time = f(CD - IN). Reaction time is a function of C; an individual difference variable; D, a variable reflecting a stimulus regularity dimension and IN, an inhibition construct. D is considered a function of stimulus variability and other numerous variables, as yet unknown.
Similarly IN is a function of a number of unknown variables as well as time on task, amount of rest, ISI and intensity of the extraneous signal. It can be seen that McCormack's theory is narrow and ad hoc. The theory is not rich enough to account for the various phenomenon outlined at the beginning of the chapter. McCormack's theory runs into trouble on the same grounds as Mackworth's classical conditioning theory. McCormack found no significant change in RT with ISI (McCormack, 1960). Inhibition theory would suggest on opposite result, RT should be fastest just after a signal and then decline as ISI increases. McCormack's model has never enjoyed much acceptance in the literature.

The conclusion that can be drawn from a review of motivational variables and motivational theory is that such variables and constructs deserve a place in a comprehensive theory of vigilance performance. However, at this time two little is known of their role and influence.

Davies and Tune Model (1969)

Davies and Tune (1969), after a comprehensive review of the vigilance theory literature, outlined an information flow model. The model is shown in Figure 5. The numbers above the arrows indicate the sequence of the flow of information. The model is an electric model that has many of the features suggested by the results from many researchers. Davies and Tune (1969, p. 232) viewed the vigilance process as consisting of several different stages: (1) extracting information about the stimulus event, (2) processing this information, (3) comparing the central representation of the stimulus event with a
Figure 5. Davies and Tune (1969) Minimum Requirement Information-Flow Model of Vigilance Performance (From Davies and Tune, 1969).
stored representation of the signal, (4) determining the discrepancy between the two representations, (5) deciding whether the discrepancy is sufficiently small for the stimulus event to be called a signal, and (6) making the appropriate response. The model assumed a Sternberg (1967) two-operation analysis of signals and non-signals. The first operation extracted and abstracted signal and non-signal features from the stimulus event and the second operation compared these patterns with a stored representation which resulted in a match or mismatch; this in turn determined the response. Sternberg has suggested the same amount of time is required for processing signals as nonsignals. The same amount of time should be required for correct detections as for correct rejections. Parasuraman and Davies (1976) have, however, found differential response times for correct detections and correct rejections. The authors found RTs for CRs and misses to be lower than for CDs and FAs. This would support Neisser's (1963, 1964) contention that a target stimulus is analyzed initially for gross features and further analyzed if the initial test is passed. This would have the effect of placing an arrow between the non-signal and signal template. This suggests the existence of two different templates or further processing within a single template for signals. However, Jerison's findings of a decrement for high event rates but not low event rates (number of signals held constant) and lower response times for higher event rates poses some questions about the nature of this matching process. Davies and Tune have indicated that paths 5a, 5b, 6a, and 6b are alternate paths that stimuli may take. The authors have suggested that the
vigilance decrement may be a forgetting process; increases in RT over time may be due to a fading of the signal template. Motivation and expectancy are assumed to influence decisions to observe or not to observe. Davies and Parasuraman (1982) make no mention of this earlier model. Lack of knowledge about attentional mechanisms (observing mechanisms) and details of the internal model used by the decision-maker are perhaps, reasons for little or no mention of this flow model in the literature. The model does not really add much more to our knowledge of vigilance than has already been reported.

Stroh IAF Model 1971

Stroh (1971) has outlined an IAF (Inhibition-Arousal-Filter) signal detection model based on his review of the research literature. The model is shown in Figure 6. Stroh (1971, p. 66) comments, "after examining the already existing theories, the present author came to the conclusion that there was a little truth in all of them, but that none of them was strong enough to stand on its own." The vigilance decrement is attributed to filter deviations or shifts of attention. Attention may be directed internally or externally (although Stroh does not mention it, one might expect concomitant, RT, heart rate and EEG changes). Intersignal interval characteristics influence the extent of filter deviations—a large mean or range of ISIs will make it difficult for observers to estimate when the next signal will occur. This notion embodies some of the influences of signal regularity, expectancy and time estimation ability. Some of the notions of habituation are also included at this point. Stroh includes in his model inhibition and
Figure 6. Stroh IAF Signal Detection Model (From Stroh, 1971).
extinction. If there are few signals to detect, watching for signals is not sufficiently reinforced. Apparently, Stroh is assuming that signal detections are intrinsically motivating. This results in extinction of observing behavior. At the same time, according to Stroh, the observer is required to watch a large number of repetitive events which lead to inhibition. Signals lose their novelty and the filter becomes more selective. Stroh asks "Why do filter deviations occur?" He attributes these deviations to extinction and inhibition. To speak of filter deviations implies a deviation away from some set point. Stroh does not indicate what set point he has in mind or what it means. Rather than speaking of deviations of the filter, one would think that the filter gets fixed at a certain point. As stimuli becomes repetitive the filter prevents them from being processed other than superficially. Inhibition affects internal filter deviations, which leads to daydreaming and these in turn predispose a person towards low levels of arousal. Externally influenced filter deviations are referred to as task-irrelevant external events. Stroh does not define these--it is not known if this is noise, temperature extremes, etc. The model does not mention where increases in arousal enter the model and yet some monitoring situations are characterized by heightened arousal. Time error can either improve or detract from performance by having the observer estimate intersignal appropriately or inappropriately. This aspect of the model is similar to the notion of expectancy; however, it is different than Deese's or Baker's view of expectancy. The last part of the model is "increased cautiousness."
More signals are missed and fewer false alarms are emitted. This corresponds with a shift in criterion that has often been observed in signal detection tasks. Stroh is not explicit concerning how increased cautiousness leads to both a vigilance decrement and vigilance improvement.

Many features of other models are incorporated into the Stroh model. However, one failing of Stroh is that he does not cite the empirical data which have led him to postulate different elements of the model. Moreover, the functioning and interactions of the various model components are not specified. It appears that intersignal interval variations plays too large a role in the model. This variable describes temporal uncertainty and signal regularity but many other signal characteristics are not included which are known to affect vigilance performance, such as signal intensity, signal duration, etc. Little provisions of motivational effects exist in Stroh's model unless it is reflected in interal filter deviations or increased cautiousness. The model considers too few signal, task, subject or environmental variables. The model, like many other models is somewhat descriptive but too general. Furthermore, the model and its components need to be operationalized if it is to be tested or capable of prediction.

Status of Vigilance Theories Today

In reviewing the literature, one gets the impression that 10-15 years ago, the thing to do was publish an article and explicate a model or theory of vigilance. Perhaps this is an exaggerated view of previous research. However, recent publications seem to be more concerned with descriptions of vigilance performance and changes that occur in
the observer during a watch and less on modelling. Another observation, is that current research is focused in the direction of trying to better understand existing models and what they mean at a very fundamental level. For example, there seems to be a strong emphasis on the psychophysiology of arousal and basic brain research. Current interest seems to be focused on attentional mechanisms, i.e. why, how and under what conditions does an observer attend to a display? There are some concerns for the validity of vigilance measures and the vigilance decrement. Attempts at consolidating research results seem to be more empirically-based and less speculative than in the past.

Davies and Parasuraman (1982), state:

Theories of vigilance have addressed themselves to two main questions: "What determines the decline in performance during a vigilance task? and what determines overall level of performance? However, much greater emphasis has been placed on the development of answers to the first question than to the second and most theories of vigilance are devoted exclusively to an explanation of the vigilance decrement (p. 9).

... the vigilance decrement is a widespread and persistent phenomenon for which no single theory ... can provide a completely convincing explanation. The reasons for this state of affairs are first that traditional measures of the vigilance decrement are inadequate and do not accurately reflect the changes in performance occurring during a vigilance task and second, that with a few exceptions, the influence of task factors upon the decrement has been relatively neglected in theoretical accounts of vigilance performance (p. 24).

Loeb and Alluisi (1980) reported the following statement in the proceedings of the human factors society--25th annual meeting:

There really are no theories or models specific to vigilance; in general they have reflected the psychological fashions of their time ... interestingly, almost all of these descriptive and theoretical frameworks describe and predict data almost equally well, and no one of them explains all vigilance phenomena (p. 600).
Corcoran, Mullin, Rainey and Frith (1977, p. 645) divide the vigilance theories into two categories: "It is useful to consider theories of vigilance to fall into two categories--state theories and strategic theories."

A detailed review of vigilance theory was reported by Loeb and Alluisi (1970, p. 360). They said:

... There are several reasonable theories, but the data are seldom if ever such as to corroborate one model and refute others exclusively ... the problem is not so much that none accounts for the data completely, but rather that all of them can do so reasonably well, sometimes with only minor modifications. There has been an unfortunate general tendency, doubtless shared by the writers, to make a prediction based on a given theory, confirm the prediction, and claim that the theory is thereby supported (if not completely proved), ignoring the fact that most of the other theories also make the same prediction! There has been an equally unfortunate tendency to set up complex situations in which more than one complex mechanism is sure to operate and on the basis of the data collected to argue that one mechanism must be ruled out completely rather than limited in applicability.

Loeb and Alluisi updated their 1970 literature review with another in 1977 (Loeb and Alluisi, 1977, p. 749). They summarize their views in the following way:

The current status of vigilance theories in the mid-1970s is summarized as follows: (1) recent research, like previous research, has failed to confirm any one theory exclusively, (2) the data available continue to cast doubt on the prospect of any current theories being able to account adequately for all established vigilance phenomena, (3) the differentiation of "cortical arousal" may provide a basis for a useful advance in an arousal-theory explanation of some monitoring phenomena, especially as related to certain brain-wave activities, and (4) other factors not encompassed by any of the theories are known to affect vigilance, some of them to appreciable extents.

Broadbent (1971, pp. 20-21) has the following to say concerning vigilance theory:
Broadbent (1958) divided the various theories of vigilance into four (inhibition theory, expectancy theory, arousal theory, and filter theory). As it was clear at the time, these were not necessarily mutually exclusive, and it might well be that more than one of them was true. Without undue forcing, one can classify the work done on vigilance since 1958 into the same four groups, since various investigators have planned their research in the interval in the light of their enthusiasm for one theory or another. Most investigations can be pushed into relevance to one or other of the main lines of thinking.

Warm (1977, p. 644) pointed out that research goes through an evolution consisting of four phases:

First is one of enthusiastic discovery; second, the accumulation of detailed information; third, a phase of deeper understanding and exploration; and fourth, an attempt at synthesis. After a quarter of a century of intensive investigation, it seems to me that vigilance research is in the third phase.

Task Influences on Performance

A task is any activity that a person performs, and is characterized by a transfer of information from input to output. Information processing between input and output is generally considered to consist of stages that require time. These stages transform the information in some way and make the information available to the following stage. A stage may be considered or likened to a subtask. The importance of tasks to information processing is, perhaps, best explained by Teichner's assertion: "A theory of human performance . . . must be a theory of subtask functions and relationships. (Teichner, 1974a, p. 2)

Warm (1977) is more specific concerning the nature of vigilance tasks:

In spite of wide variation in the nature of vigilance tasks, they all have several features in common which, in sum, are unique to the watchkeeping situation: (1) the task is prolonged or continuous, (2) the response of the observer typically has no effect upon the probability of occurrence of critical events, (3) the signals to be detected are usually
clearly perceivable by the observer when is is alerted to them, and (4) the signals to be detected occur infrequently and aperiodically. (p. 625)

A number of attempts have been made to understand human performance in terms of task components. A task taxonomy is a classification of behavior involved in task performance. A classification system is considered the sine qua non for science. The purpose of developing a task taxonomy is similar to the purpose of the periodic chart developed by Mosely for classifying the atomic structure of the basic elements. Once the task taxonomy is established, it can be used to predict human behavior and to generalize the results to other similar situations in terms of task components. Fleishman (1967), for example, discusses the need for a task taxonomy for complex performance. Fleishman argues that for many years psychologists have studied learning and performance under many different task and environmental settings and collected much data and yet when new systems are designed, little of this vast amount of data is applicable and the performance problems must be re-studied from scratch. There seems to be no set of unifying dimensions. Two problems exist because of the lack of an empirically derived task taxonomy: (1) principles cannot be generalized from one operational system to another, and (2) findings from the laboratory cannot be generalized to operational tasks. Fleishman (1967) says: "Most learning theory is devoid of any concern about task dimensions and it is this deficiency which, many of us feel, makes it so difficult to apply these theories in the real world of tasks and people. What we need is a learning and performance theory which ascribes task dimensions a central role." (p. 350)
Other researchers have studied the problem of task taxonomies. Bergum (1966) developed a conceptual framework for tasks based on the activation level of the human. The conceptual framework is designed to include a range of research tasks from monitoring to production-line performance. Table 1 summarizes Bergum's task classification scheme. Three general sources of stimulation are considered: (1) relevant task stimulation (stimulation arising from the task being performed), (2) mediation stimulation (the degree to which memory and central processes are reflected in output responses), and (3) reaction (the characteristics of overt responses). The three general sources of stimulation are broken down into three additional dimensions. For example, relevant task stimulation is broken into complexity, frequency, and periodicity of stimulation. Table 1 provides examples of tasks that range from a low to high degree of stimulation on the CNS. Bergum suggests that the classification paradigm of Table 1 can be transformed into a matrix to score various tasks. A particular task is classified by scoring the task from 1 to 3 along each of 9 dimensions and totaling these numbers. Total scores that range from 9 to 27 are possible and should roughly indicate the type and amount of performance to be expected from a particular task. Application of this scoring technique to vigilance tasks yielding the classical decrement results in scores of between 9 and 12 while high loading tasks (production-line type tasks) yield scores of around 24. The approach appears to have the merit of providing organization for research on continuous performance as well as enabling one to predict the performance to be expected under various task conditions. If the initial source stimulation is low and some variable that
Table 1. Matrix of Continuous Performance Tasks.

<table>
<thead>
<tr>
<th>Stimulus</th>
<th>Low (1)</th>
<th>Moderate (2)</th>
<th>High (3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source</td>
<td>(Vigilance Effort)</td>
<td>(Efficient Performance)</td>
<td>(Lapses and High Variability)</td>
</tr>
<tr>
<td>Relevant</td>
<td>Simple Frequency Low Aperiodic</td>
<td>Complex High Frequency Periodic Groups</td>
<td>Continuous-Compound Continuous Periodic</td>
</tr>
<tr>
<td>Mediation</td>
<td>Transduction Low Reactivity Low Activation</td>
<td>Choice Making Normal Moderate</td>
<td>Combination Over-Reactivity High Activation</td>
</tr>
<tr>
<td>Reaction</td>
<td>Simple Low Motor Infrequent</td>
<td>Chained Moderate Motor Frequent</td>
<td>Complex Chained Heavy Motor Continuous</td>
</tr>
</tbody>
</table>

Note: From "A Taxonomic Analysis of Continuous Performance" by Bruce O. Bergum, Perceptual and Motor Skills, 1966, 23, 47-54.
contributes to overall stimulation is added, an increase in performance is expected. However, if the initial level of stimulation is high and a variable is added that increases the level of stimulation, lowered performance might be expected. For example, if a person performs a monitoring task with a single display, the addition of other displays might increase the overall level of stimulation and instead of observing a vigilance decrement, no change in performance might be observed. Bergum's task classification follows the inverted y-function specified by the Yerkes-Dodson Law.

More recent efforts have been made to demonstrate the value of developing a task taxonomy. Levine, Romasho, and Fleishman (1973) classified 53 studies in the vigilance literature in terms of four abilities required for task performance—perceptual speed, flexibility of closure, selective attention, and time-sharing. Perceptual speed refers to the time required to compare patterns for identity or degree of similarity. Flexibility of closure refers to the ability to extract a specified target from a more complex stimulus field. Sustained attention studies were selected for classification by the researchers because the range of tasks is not as great as in other areas, a reasonably large number of studies are reported in the literature, and a common performance metric—detection accuracy—is used in most vigilance studies.

The study was designed to determine if performance in vigilance tasks could be distinguished on the basis of abilities and if generalizations about the effects of the independent variables could be
improved because of such a task classification. Studies were selected for classification that had signal rate, sensory-input mode, or knowledge of results (KOR) as independent variables. These manipulated variables show consistent effects in the literature. Performance is generally improved by increasing signal frequency, use of multiple modality monitoring, and KOR.

The researchers plotted median percent correct detections at each 10-minute interval up to three hours for studies falling within each category of task and for the three independent variables. Median percent correct detections showed a decrement for tasks in which perceptual speed was dominant. The decrease in median percent correct detections was immediate but asymptoted to about 60-65 percent after 50-60 minutes. Tasks in which flexibility of closure was dominant exhibited a U-shaped performance curve. Performance began at about 80 percent correct detections and declined for the first hour. After 60 minutes, median percent correct detections began to increase until a level was reached that approximated initial performance. Studies in which selective attention or time-sharing was the dominant ability were too few to interpret functional relationships. When the effects of independent variables on performance were plotted, different functional relationships were also found depending on the abilities required by the tasks across the vigil.

The work of Levine, et al. is important for several reasons: (1) classification of tasks according to abilities required for the task resulted in different relationships between performance and time on the task, (2) when task performance was partitioned by levels of three
independent variables, differences in functional relationships also emerged for the ability categories, (3) improved generalizations concerning vigilance performance resulted from the task taxonomy and (4) relationships were revealed that previously had been obscured.

More recently, Parasuraman and Davies (1977) developed a taxonomic analysis of vigilance performance. Parasuraman and Davies surveyed the vigilance literature with the view of developing a task classification scheme that would enable researchers to predict performance as a function of selected independent variables. Nine dimensions of tasks were discussed that had previously been reported in the literature: sense modality, source complexity, response type, coupling, signal duration, time course of events, attention requirement, stimulation value, and task abilities. The researchers conclude that only sense modality, source complexity, time course of events and task abilities (perceptual speed and flexibility of closure) are relevant to the classification of vigilance tasks. Twenty-seven studies reported in the literature were classified according to these four dimensions. Each study was reviewed to determine whether or not a sensitivity decrement was reported. The task classification scheme developed by Parasuraman and Davies did impose a degree of order on the various results reported in the literature. Studies that reported a sensitivity decrement showed a definite "clustering". Studies that exhibited a single source of complexity, that were classified as requiring perceptual speed, and had a high rate of stimulus presentation (time course of events) showed a sensitivity decrement in most instances. Sense modality did not seem to be important. A sensitivity decrement appeared about equally in both auditory
and visual tasks. It should be noted that a sensitivity decrement appeared in only one multi-source complexity task. The agreement was unanimous for other classification categories; there was no decrement in performance over the watch.

The development of a taxonomy for vigilance tasks is necessary to a theory of vigilance performance. A vigilance task classification scheme enables the research literature to be systematized so that improved generalizations can be made when extending data from the laboratory to operational settings. An improved predictive capability results from such a taxonomy. Presently, a comprehensive vigilance task taxonomy does not exist and it is precisely for this reason that a diversity of seemingly conflicting results are reported in the research literature. Until a validated task taxonomy exists, vigilance research will continue to be uncertain and task performance unpredictable.

Time Perception and Vigilance Performance

An important question arises concerning the relationship between protensity and vigilance performance. If people can reliably estimate time intervals, then they can improve vigilance performance providing the signals are regularly spaced or have some discernable temporal pattern. Such a finding would have a bearing on task design in sustained attention research. A search of the Psychological Abstracts was done for the period 1967-1981 for human time estimation.

McGrath and O'Hanlon (1967) conducted a three-stage experiment that examined the influence of time perception on the performance of a
vigilance task—one of the few that relates time perception to vigilance performance. They found no significant correlation between the means and standard deviations of productions and reproductions of five-second and 30-second intervals and ability to detect signals in a vigilance task. Measures of time duration estimation ability and detection ability were developed in two different stages of the research. The researchers also examined a person's ability to estimate a 10-minute interval measured concurrently with performance of a vigilance task. Individual differences were large and ranged from one subject who perceived an hour to pass in 22 minutes to another who perceived it to pass in 109 minutes. Overall the subjects detected the highest rate of detections of signals when they estimated duration to be somewhat slower than clock time.

A follow-up study to the McGrath, O'Hanlon study was performed by Coltheart and von Sturmer (1968) that employed the method of reproduction to estimate empty intervals. Ten male subjects were required to make 30 consecutive reproductions of a 10-second interval by tapping a Morse Code key after the initial presentation of a standard interval. In a second experiment, 10 different subjects made 10 consecutive reproductions of a 60-second interval. A trend analysis of the data was performed that indicated successive reproductions of either a 10-second or 60-second standard tended to grow longer and longer. Coltheart and von Sturmer conclude that their results contradict the findings of McGrath and O'Hanlon. They suggest the relationship between objective time and subjective time is quadratic and not linear as reported by McGrath and O'Hanlon. The two studies use different
standard interval values for their estimation tasks. A range effect may exist and the results from the two studies may not be comparable.

von Sturmer (1968) performed a study that related time perception and vigilance in terms of statistical decision theory. von Sturmer found a serial effect in continuous time estimation tasks, i.e., he found a progressive lengthening of judgments over the task. The serial effect is analogous to the vigilance decrement and can be manipulated by varying task pay-off structures according to von Sturmer. The vigilance decrement may be more general than previously believed; identical processes may underlie the vigilance decrement and the serial effect in the view of von Sturmer. The task used by von Sturmer had subjects estimate short intervals (8-10 seconds).

Hawkes and Sherman (1971) extended the work of von Sturmer and investigated longer durations in a time estimation task. The researchers used estimation intervals of 0.5-, 1.0-, and 5-minute duration. Hawkes and Sherman used the method of verbal estimation and the method of reproduction to investigate time duration ability with a stimulus filled background in a vigilance task. Hawkes and Sherman analyzed their data using ANOVA procedures and found significant stimulus duration and method effects. They also found a significant method-by-blocks interaction which indicated that estimated durations got longer with time on the task. The finding of a judgment drift supports the earlier findings of von Sturmer.

Roeckelein (1972) performed an extensive review of time perception research and concluded that people tend to overestimate short durations and underestimate long ones. According to Roeckelein, the findings are
valid for both filled and empty intervals although filled intervals have been less studied. The research reviewed by Roeckelein used traditional time estimation paradigms and his conclusions conflict with the findings of von Sturmer (1968) and Hawkes and Sherman (1971) who coupled time estimation with a vigilance task. The findings of McGrath and O'Hanlon suggest that people that underestimate time duration perform better in a vigilance task. Only the work of McGrath and O'Hanlon studied a person's ability to estimate long intervals and this work indicated that estimation performance is characterized by wide individual differences. It seems that people in vigilance tasks have difficulty in estimating time intervals in the range of even less than one minute and that their performance gets worse over the watch in the direction of overestimation. There is little information concerning a person's ability to estimate time durations longer than one minute in vigilance tasks.

Task Duration and Time on the Watch Effects

Session duration needs to be considered in the design of a sustained attention research task regardless of whether this is an independent variable to be studied or a factor to be fixed. The task duration must be long enough to include any performance changes that might occur. Mackworth (1950) found the percentage of signal detections to decline during the first 30 minutes of a 1.0, 1.5, or 2 hour watch. Teichener (1974b) reviewed 37 vigilance studies to determine the functional relationship between the detection of a simple visual signal and time on the watch. Teichener grouped these studies according to their reported or estimated initial percentage of detection,
90-99%, 80-89 percent ..., or 30-39%. The largest performance decrement, about 35%, is associated with the 90-99% initial detection level. The average loss in performance is about 10% across all studies and is generally complete in about 50 minutes. The two exceptions are the extreme groups. The 90-99% group never asymptotes and the 30-39% group shows no performance decrement. In most cases the loss in detection is complete in 20-35 minutes, and at least half of the final loss is complete within the first 15 minutes. Singleton (1953) showed a performance decrement over the first few minutes of a perceptual-motor task. The literature suggests the vigilance task need not be extremely long to show a performance decrement. This is true for tasks that use percentage of detections as a performance measure. It is unclear from the literature what relationships exist between other performance measures such as response latency and time on the watch.

Intersignal Interval Effects

Related to temporal orientation or time perception is intersignal interval (ISI). Signals may be spaced over a watch regularly or irregularly. One would hypothesize that subjects will develop expectancies for the occurrence of regularly spaced signals. Expectancy would be low immediately after a signal, increase to a high point and then decrease to a low level. Baker (1959) found the percentage of signals missed to be higher in a task where the ratio of the shortest intersignal interval to the longest was 1:5.4 than when the ratio was 1:14.3. Warm, Epps, and Ferguson (1974) found signals were detected more rapidly when regular as compared to irregular intersignal intervals were used. Response times tended to decrease with time on watch.
when regular ISI's were used and to increase over time when ISI's were irregular. Loeb and Alluisi (1970), reviewed the variables that influence monitoring performance and suggested that signal regularity may not be as important as originally supposed and that it is not as important as variations in spatial uncertainty. An update of findings regarding vigilance by Loeb and Alluisi (1977) states the case more strongly:

We find no evidence that signal regularity has been found to affect detections or false alarms generally. Also, later studies have reported the effects of signal regularity to be less important than those of signal density (Smith, Warm and Alluisi, 1966) or spatial uncertainty (Adams and Boulter, 1964). Apparently signal regularity is less likely to have an effect at longer intersignal intervals. (p. 723).

Davies and Tune (1969) reviewed the literature on the effect of intersignal interval variability on response time and concluded:

"Response time appears to be a decreasing function of inter-stimulus interval in low variability conditions, is invariant in medium conditions of variability and is an increasing function of inter-stimulus interval in high variability conditions" (p. 79). However, the authors acknowledge the relationship between response latency and the length of the intersignal interval is a little confused.

**Signal Duration Effects**

The design of a sustained attention research task needs to consider signal duration. The longer a signal is displayed the greater is the likelihood for detection (Baker, 1963a) found that a range of signal durations of 0.2-0.4 seconds seemed to be critical to detection
performance. Durations beyond 0.4 seconds have little effect on performance.

Artificial Signals

Secondary tasks are often used to study human capabilities in human factors research. Secondary tasks are also used to maintain alertness in prolonged monitoring tasks. For example, the "dead man control" is used in locomotives to maintain the alertness of the engineer. In the nuclear industry, consideration is being given to using a subsidiary task presented to control room operators on a computer-interfaced CRT to maintain alertness. One serious disadvantage of using subsidiary tasks to maintain monitor alertness is the monitor may become too engrossed in the subsidiary task and miss critical plant signals. One solution to the problem is to provide a subsidiary task does not distract the operator. The use of artificial signal is a signal which is not discriminable from a real signal that is injected into the monitoring task. Usually knowledge of results (KOR) about detection performance with regard to these artificial signals is provided to the operator.

Baker (1960a) tested subjects in a vigilance task to determine whether artificial signals coupled with knowledge of results improved performance. The control condition consisted of a Mackworth sequence of signals. The experimental condition consisted of the same sequence of signals but with artificial signals interspersed in the sequence and KOR provided. Baker reported a difference between the control and experimental group that was significant at the 1 percent confidence
level. One serious drawback to Baker's study is that the experimental design confounded KOR with artificial signals. Consequently, it is now known from Baker's study whether performance improved because of artificial signals or because of KOR.

Wilkinson (1964) performed a study to separate the effects of artificial signals and knowledge of results. The detection task was an auditory detection task. Subjects had to detect real signals that lasted 0.37 seconds from a train of 0.5 second tones against a background of white noise. An event occurred every 0.3 seconds and the entire task lasted one hour. Eight real signals were presented per hour. Five conditions were compared:

1. real signals only
2. 40 identical artificial signals—no KOR
3. 40 identical artificial signals—partial KOR
4. 40 identical artificial signals—full KOR
5. 40 different (0.66 sec. duration) artificial signals—full KOR.

Wilkinson found that condition 3 produced a significantly higher detection rate than did condition 1 or 2. Wilkinson concluded that artificial signals produce better results only when coupled with KOR. Wilkinson used signal durations in a range in which percent signals detected was found to be very sensitive by Baker (1963b). Wilkinson's results may be confounded with signal duration effects.

Murrell (1975) performed an experiment that used a visual detection task to examine the effects of artificial signals with and without KOR on the detection of real signals. The subjects were 54 members of the navy. These subjects were divided into three groups of 18: a control group with real signals only (RSO), a group with artificial
signal injection (ASI), and a group that received both artificial signals and knowledge of results (ASI + KOR). Each subject viewed an oscilloscope display in a darkened booth. The oscilloscope display consisted of an array of 16 lines that changed at the beginning of each nine second trial. Nonsignal line length was normally distributed with a mean length of 16.5 mm and a standard deviation of 1 mm. A signal consisted of a fixed increment of 3.3 mm added to the length of one of the lines. Artificial signals were identical with real signals. Real signals occurred randomly on one-in-ten of the trials whereas artificial signals occurred randomly on one-in-twenty of the trials. KOR was provided only in trials with artificial signals by having a circular symbol appear on the screen next to the signal during the last two seconds of a trial. An analysis of variance was performed for each of four performance measures: (1) percent detection, (2) percent of false alarms, (3) sensitivity ($d_a$), and (4) natural logarithm of $b$. Murrell found that on the average, the addition of KOR had no significant effect on performance and that the use of artificial signals resulted in significantly ($p < 0.001$) more detections and false alarms than the control group. The experimental groups also exhibited a significantly more lax mean criterion level as shown by $\ln \beta$ ($p < 0.001$). This would explain the increase in the mean percent of false alarms. An analysis within sessions (first half of session versus last half) shows that artificial signals and KOR did little to prevent a sensitivity decrement. Murrell suggests that KOR was ineffective because of the inherent confusability of signals and nonsignals; subjects had to learn
distributions of stimulus values rather than single values of signals and nonsignals.

The research literature concerning artificial signals is equivocal and often conflicting. However, the use of artificial signals increases signal density. Increasing signal density, in general, improves monitoring performance. Because of the conflicting results reported in the literature concerning artificial signals, further research is suggested.

**Warning Signals**

Warning signals that announce some impending condition generally reduce response latencies. The warning signal may be visual or auditory. An auditory alarm is superior to a visual alarm when immediate action is required. Since the ears are omnidirectional, attention is attracted more quickly than through the eyes (Huchingson, 1981, McCormick, 1976). If a warning signal is too close to the condition that requires an action, the warning signal may interfere with the action. If the warning signal is too distant from the condition that requires a response, warning signal credibility may suffer. McCormick (1976) uses a rule of thumb and suggests the warning interval should not be less than 0.5 seconds.

Warning signals, like artificial signals have been used to maintain alertness during the watch. No studies, however, have been found that compare the effectiveness of warning signals with artificial signals.
Signal Intensity Effects

One problem that exists concerning research on the effects of warning signals and artificial signals on monitoring performance is that of signal intensity. If differences are found in monitoring performance when a warning signal is presented as compared to effects found when artificial signals are presented, the researcher cannot be sure the effects are not brought about by differences in signal intensity. Methodologically, this can present some problems to the researcher. The perceived intensity of auditory warning signals can often be measured if the sounds are not too complex or masked by other signals. Measurement of the perceived intensity of visual signals can also present methodological problems. For example, if an operator is monitoring several meters that vary about some mean value, does the operator perceive signal intensity to be the mean value or the variability about the signal? Assuming that signal definition problems can be overcome, the technique of cross-modality matching can be used to equate different modality signal intensities. (Stevens 1975)

Recent research suggests that signal intensity variations may not affect performance as much as previously thought. Lisper, Kjellberg, and Mellin (1972) performed an auditory vigilance task in which signal intensity was at threshold, 34, 48, and 88 dB. They found increases in RT across the vigil that were largest for threshold intensities. These results are counter to earlier results found by Mackworth (1950) and Adams (1956). Thurmond, Binford, and Loeb (1970) also found that signal intensity in an auditory vigilance task did not influence detection rates, false alarm rates, d' or b. Van der Molen and Keuss (1979)
report that simple RT decreases with increases in signal intensity whereas RT in a choice RT task is related to signal intensity by an inverted y-shaped curve. Davies and Tune (1969) in a review of the vigilance literature, reported that within narrow limits the Bunsen-Roscoe law describes the effects of signal intensity, i.e., RT decreases with increases in signal intensity. This conclusion is supported by the findings of Van der Molen and Keuss. Indeed, the usual finding concerning the effect of signal intensity is that it reduces RT. However, the several studies cited here suggest that signal intensity effects may not be simple and in some instances don't exist. Moreover, the effects of changes in signal intensity are unknown with regard to physiological measures.

**Performance Measures in Sustained Attention Research**

Traditionally, sustained attention research reports cite the following performance measures: (1) frequency of observing responses, (2) the probability of errors of omission or commission, (3) timing aspect of behavior, i.e., response latencies, (4) changes in the threshold sensitivity, i.e., changes in the decision cut-off point, and (6) changes in signal detection theory parameters (Warm, 1977).

In addition to these traditional measures, other measures have been suggested. For example, fifteen years ago, Buckner and McGrath (1963) made the following assertion: "If the arousal hypothesis is going to be useful in explaining vigilance behavior, it is fairly obvious that we need an independent measure of arousal. I suppose a physiological measure is the most likely candidate" (p. 124).
More recently, researchers are recommending broad-based approaches in the study of sustained attention, that it, research should integrate subjective, behavioral, and physiological data (Gale, 1977). Mackie (1977) supports this notion of a multilevel approach to research: "We are not engaged in an agreement whether physiology or psychology is more basic to vigilance behavior, but rather about how physiological changes of all kinds take place along with subjective changes, behavior changes, and so on." (p. 15) Recently, Mulder (1979a) has followed this approach and organized literature according to vigilance state versus various criteria, i.e., electro-physical, psychophysiological, and psychological characteristics of the operator. His results are presented in Table 2.

The effect of vigilance state on various criteria is not as certain and clear cut as Mulder's table suggests. For example, Mulder asserts that in a hypovigilance state, a person's mean reaction time increases as well as reaction time variability. Davies and Tune (1969) on the other hand, claim that little is known about the relationship between reaction time and detection rate during vigilance task performance. Although latency measures have been used to describe vigilance performance, usually only detection latencies are recorded and latencies associated with correct rejections and false alarms are not recorded (Parasuraman and Davies, 1976). Parasuraman and Davies (1976) performed a decision theoretic analysis of response latencies for correct detections, false alarms, correct rejections, and omissions in
Table 2. Efficiency as a Function of Vigilance Level.

<table>
<thead>
<tr>
<th>Stages</th>
<th>Hypervigilance</th>
<th>Optimal level</th>
<th>Hypovigilance</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>I. ELECTROPHYSIOLOGICAL CRITERIA</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1) EEG general aspect</td>
<td>Desynchronized</td>
<td>normal</td>
<td>synchronized</td>
</tr>
<tr>
<td>dominant frequency (in c/m)</td>
<td>2 (18 - 30)</td>
<td>(1 (15 - 20)</td>
<td></td>
</tr>
<tr>
<td>mobility and complexity (M &amp; C)</td>
<td>eventual θ rhythms</td>
<td>small α</td>
<td>slower &amp; higher α</td>
</tr>
<tr>
<td>amplitude (in uV²)</td>
<td>10-20</td>
<td>20-40</td>
<td>50-70</td>
</tr>
<tr>
<td>activity (A² in μV²)</td>
<td>A&lt;</td>
<td></td>
<td>A&gt;</td>
</tr>
<tr>
<td>power spectral density</td>
<td>&lt;</td>
<td>average</td>
<td>&gt;</td>
</tr>
<tr>
<td>in μV²/8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coherence (0-1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2) E.C.G. (frequency rhythm)</td>
<td>slight tachycardia</td>
<td>normal</td>
<td>bradycardia</td>
</tr>
<tr>
<td>sinusal arrhythmia</td>
<td>sinusal arrhythmia</td>
<td>normal</td>
<td>sinusal arrhythmia</td>
</tr>
<tr>
<td>decreased</td>
<td>increased</td>
<td>increases</td>
<td>increases</td>
</tr>
<tr>
<td>(3) R.R. (Respiratory rhythm)</td>
<td>Quick</td>
<td>regular</td>
<td>Low, more regular</td>
</tr>
<tr>
<td>small amplitude</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(4) E.O.G. (Amplitude rhythm)</td>
<td>increases</td>
<td>normal</td>
<td>decreases, generally accelerated</td>
</tr>
<tr>
<td>accelerated</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(5) E.M.G. (Amplitude)</td>
<td>Lesser</td>
<td>normal</td>
<td>Larger</td>
</tr>
</tbody>
</table>
Table 2. (continued).

<table>
<thead>
<tr>
<th>Stages</th>
<th>Hypervigilance</th>
<th>Optimal level</th>
<th>Hypovigilance</th>
</tr>
</thead>
<tbody>
<tr>
<td>II. PSYCHOPHYSIOLOGICAL CRITERIA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensorial afferencies</td>
<td>filtered</td>
<td>optimal regulation</td>
<td>defective filtering</td>
</tr>
<tr>
<td>Perceptive area</td>
<td>focalized</td>
<td>well adjusted</td>
<td>'lapses'</td>
</tr>
<tr>
<td>consequences in detection</td>
<td>'false alarms'</td>
<td>possibility of anticipation and subception</td>
<td>free associations; omissions</td>
</tr>
<tr>
<td>Psychomotricity</td>
<td>mean reaction time (m) shorter</td>
<td>normal</td>
<td>longer variability increases stability</td>
</tr>
<tr>
<td></td>
<td>standard deviation (σ) increases</td>
<td></td>
<td></td>
</tr>
<tr>
<td>III. PSYCHOLOGICAL CHARACTERISTICS OPERATOR</td>
<td>Tension</td>
<td>optimal alertness</td>
<td>relaxed, wandering</td>
</tr>
<tr>
<td></td>
<td>irritability</td>
<td>attentiveness</td>
<td>attention</td>
</tr>
</tbody>
</table>

a 45 minute visual monitoring task. They found that latencies associated with correct detections and false alarms increased; whereas latencies for correct rejections and omissions decreased. Mulder's classification is too general and misleading.

The effect of vigilance state on heart rate measures is also not clear. A number of studies under controlled laboratory conditions have shown that by increasing the difficulty of a controlled perceptual task the variability of heart rate tends to decrease (Kalsbeek and Ettema, 1963, 1965). These researchers used a paced binary choice task to assess heart rate variability. Surwillo (1976), on the other hand, found no relationship between heart rate (HR) and the detection of signals. O'Hanlon (1972) found that HRV increased markedly with driving time and that HRV dropped substantially after the occurrence of events which realerted the drivers. O'Hanlon also found that HRV recovered after rest and did not seem to be related to traffic event frequency. Laurell and Lisper (1976) found HR to progressively decrease over a two hour driving task. These results confirmed results found in an earlier study by Lisper, Laurell, and Stening (1973). Manuck and Schafer (1978) demonstrated reproducible task-related heart rate changes across experimental sessions one week apart. They used a problem solving task. Gaillard and Trumbo (1976) found mean HR and HRV increases in a three-hour serial reaction test that they interpreted to mean that autonomic activity decreased.

Firth (1973) suggests there are two levels at which cardiac arrhythmia may be interpreted:
(1) Changes in HRV that occur over a finite period of time (or alternately a fixed number of R-R intervals) can be used to estimate task difficulty.

(2) Second-by-second changes in instantaneous heart rate can be related to specific task components. Sayers (1973) suggests that simple measures of sinus arrhythmia may not be adequate to assess or resolve complex workload responses.

There are many ways to score heart rate measures. Opmeer (1973) found at least 30 methods reported in the literature to score HRV. Different scoring techniques for HRV can lead to different conclusions concerning the effect of task loading on HRV. Some researchers have begun to analyze heart rate measures in the frequency domain with positive results (Charnock and Manenica, 1978, Hyndman and Gregory, 1975). Mulder, (1979b) summarizes the work in this area with reference to the measurement of mental workload. Spectral analysis of the R-R intervals reveals peak activity in three energy bands: a peak around 0.10 Hz associated with the fundamental frequency of the system, a peak between 0.20 and 0.40 Hz associated with respiration effects, and a peak between 0.42 and 0.60 Hz associated with task loading.

Although many researchers suggest the concept of sinus arrhythmia is well understood, this is not the case; understanding is really at a nascent stage (Kalsbeek, 1973).

Mulder's table indicates the EOG amplitude and rhythm vary with vigilance state. In a hypovigilant state the monitor's EOG shows increases in the amplitude and an accelerated rhythm. In a hyper-vigilant state, the monitor's EOG shows a decreasing amplitude and an
accelerated rhythm. One difficulty with Mulder's classification is he does not operationally define hypervigilance, normal, and hypovigilance for all listed measures.

There are good reasons to study eye movements in a monitoring task. A considerable amount of information processed in a monitoring task is visual. Rockwell (1972) estimates that vision accounts for over 90 percent of information input in a driving task—a form of monitoring task. Since there is heavy reliance on vision in many tasks, it makes sense to examine visual performance in sustained attention tasks.

Kaluger and Smith (1970) investigated the effects of prolonged driving and sleep deprivation on eye movement patterns—a task much like a vigilance task. Subjects were sleep deprived for 24 hours and then asked to drive nine hours. Data collected under this condition was compared with data collected under a baseline condition where a subject was not sleep deprived. The researchers found that the mean focus of eye movement patterns shifted two degrees to the right and two degrees down over the course of the baseline run. In the sleep deprived state this shift came immediately at the outset of the run. The dispersion of the eye movements during the sleep deprived runs was greater than for the non-deprived condition. The authors interpreted these two results to mean that subjects were allocating more focal viewing time to areas normally monitored peripherally and that this indicated perceptually narrowing.

Kaluger and Smith also found that mean pursuit eye movement duration and the volume of pursuit eye movements increased under the sleep
deprived condition as compared to the baseline condition. This pattern is very similar to patterns observed for inexperienced drivers which suggests that some regression takes place due to sleep deprivation.

An early study by Mackworth, Kaplan, and Metlay (1964) investigated eye movements using the corneal reflection technique during a vigilance task. The study attempted to investigate the phenomenon of fixation without a signal processing in a signal detection task where either one dial with a revolving pointer was present or two dials. The subject's task was to detect 0.5 seconds irregular pauses in the revolving pointer. The signals were presented at rates of either 1.8 signals/minute or 7.6 signals/minute. Detection probabilities for two dials was about half that for one dial. Analysis showed that every missed signal in the one dial case was fixated without resulting in a report. In the two dial situation in contrast to the one dial situation, the largest proportion of missed signals were not even fixated. Nearly as many signals were partially fixated and about one quarter of the unreported signals were fixated for the full duration of the 0.5 second signal.

Schroeder and Holland (1968) performed an experiment that investigated operant control of eye movements during vigilance task. Sixteen subjects monitored a four-dial display that occupied an area of 4° by 4°. Subjects were credited with a detection if they pressed a button within 2.5 seconds after the onset of a pointer deflection. A Mackworth eye movement camera (a corneal reflection method) was used to collect data. Sessions were 40 minutes duration and signal rates of 10, 1 or 0.1 signals per minute were used. The researchers found some
interesting results. Eye movement rate (an activity measure) decreased as signal rate decreased. Individuals with higher overall eye movement rates detected many more signals. The interpretation was that the speed of shifting fixation was an index of alertness. It was also found that the rate of looking without seeing or reporting was sensitive to signal rate and time on watch. The lower the signal rate and the longer a person had been in the task, the greater was the tendency to gaze at the target signal but not report it. Schroeder and Holland also found that eye movement response rates become more variable as the signal rate decreases. The authors conclude that eye movements can be viewed as observing responses that are controlled by signal detection; i.e., signal detections act as reinforcers.

Baker (1960b) performed a study to investigate observing responses and general activity (restlessness). Subjects were seated in a slightly unstable chair and given the Mackworth clock test. Baker unobtrusively photographed the subjects once per second for an hour and then analysed over 64,000 photographs. Observing responses were defined as eye fixations in the direction of the display. Baker found that observing responses increased with restlessness over the watch but did not parallel the frequency of signal detections. Baker's results suggest that observing responses are not necessarily eye fixations since the looking without seeing phenomena occurs.

Gaarder (1966) found that minature involuntary eye movements (micro-saccades), fine fast tremor, and slow drifts had a reliable characteristic pattern during inattention as compared to a state of
attentiveness. Inattentiveness produces oscillatory fine eye movements with infrequent flicks and an unstable pattern that is different from eye movement patterns in an attentive state. Gaarder suggests that a stable closed loop feedback system may be operating during attention whereas during inattention the loop is opened and results in an unstable oscillation.

Hall and Cusack (1972) performed a comprehensive review of eye movement and blink literature. They discuss an existing need to perform research in the area of both short- and long-term visual responses to stress. Hall and Cusack state: "There is evidence that attending arousal and sustained concentration are related to stress; tunnel vision or shrinking of the effective field of view is often associated with stress. Analysis and study of eye movements might lead to the development of special displays and information formatting techniques which would compensate for the effects of stress." (p. 36)

Recently, a number of researchers have tried to relate EEG with vigilance states. Mulder suggests that theta rhythms are characteristic of a hypervigilant state, small alpha rhythms a normal state, and slower and higher alpha a hypovigilant state. This suggests that cognitive states can be related to the power in various frequencies of the EEG. Indeed, this seems to be true. The power is determined for each frequency and usually plotted versus frequency. This gives some idea of the way variability is distributed over frequencies.

The EEG is not time-invariant very long and as a result researchers generally perform short epoch analysis. Frequencies that have been
associated with psychological states and are reported in the literature are summarized in Table 3.

Kappa, lambda, mu, gamma and V waves are not particularly common waveforms and therefore not as useful as alpha, beta, theta, and delta waveforms since they are not found in a large percentage of population.

Results reported in the literature tend to support Mulder's summary shown in Table 2. Beatty et al. (1974) found the performance decrement in a vigilance test was associated with the presence of theta wave activity. Milosevic (1978) found decreases in the amplitude of alpha and theta activity accompanied a decrease in sensitivity of human observers. Gale (1977b) found that performance was poorer in persons who exhibited greater amounts of theta activity.

The literature searches did not turn up a single study that related response latency, eye movements, EEG, and EKG to sustained attention. Some studies examined the relationships between one or two these outputs but not all of them. For example, Carriero (1977) studied GSR, EKG, and EMG as correlates to accuracy measures in a long duration visual task. Carriero had subjects scan a 4 x 4 letter array and to keep track of letters that appeared more than once. He found no significant main effects but did find significant accuracy x times effects. Carriero did not examine relationships between dependent measures. Gale, Davies, and Smallbone (1977) studied EEG correlates of signal rate, time in task and individual differences in reaction time during a five-stage sustained attention task. Twenty subjects
Table 3. Brain Signal Frequencies and Associated Behaviors

<table>
<thead>
<tr>
<th>Frequency Component</th>
<th>Frequency</th>
<th>Amplitude</th>
<th>Associated Psychological State</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alpha</td>
<td>8-13 Hz</td>
<td>20-60 MV</td>
<td>Relaxed state</td>
</tr>
<tr>
<td>Beta</td>
<td>14-30 Hz</td>
<td>2-20 MV</td>
<td>Excitement</td>
</tr>
<tr>
<td>Theta</td>
<td>4-7 Hz</td>
<td>20-100 MV</td>
<td>Drowsy, displeasure, frustration</td>
</tr>
<tr>
<td>Delta</td>
<td>0.5-3.5 Hz</td>
<td>20-200 MV</td>
<td>Sleep</td>
</tr>
<tr>
<td>K-complex</td>
<td>14 Hz</td>
<td>200 MV</td>
<td>Thinking</td>
</tr>
<tr>
<td>Lambda</td>
<td>--</td>
<td>20-50 MV</td>
<td>Visual exploration</td>
</tr>
<tr>
<td>Mu</td>
<td>9-11 Hz</td>
<td>50 MV</td>
<td>Tactile stimulation</td>
</tr>
<tr>
<td>Vertex (V)</td>
<td></td>
<td>Often exceeds 25 MV</td>
<td>Alertness</td>
</tr>
<tr>
<td>Gamma</td>
<td>35-40 Hz</td>
<td>--</td>
<td>Seldom encountered</td>
</tr>
</tbody>
</table>
were tested at each of five signal probability conditions, 10, 20, 30, 40 and 50 percent in a 112 minute number identification task. The researchers used EEG abundance in frequency ranges of 7.5-13.5 Hz and mean dominant alpha frequency (m.d.f.) as measures of cortical activation. Results indicate that EEG abundance for the lower measured alpha frequencies increased as the task progressed and m.d.f. decreased, EEG abundance for the higher measured alpha frequencies increased as a function of signal ration, and subjects with higher EEG abundance and lower m.d.f. were faster than subjects with lower abundance and higher m.d.f. The researchers interpret their findings to be compatible with arousal theory. Surwillo (1965) examined the relationship between alpha amplitude and heart rate in a reaction time task and failed to find an inverted U-shaped relationship such as was found by Stennet (1957). Luborsky, Blinder, and Mackworth (1964) examined the relationship between eye fixations and heart rate measures in a memory recall task. They found the more a person was accelerated in his response via heart rate or fixation rate, the greater his recovery of images. Sandman, McCanne, Kaiser and Diamond (1976) examined visual perception as a function of heart rate. Subjects were told that stimuli would be presented at very high rates on a screen and that two seconds after the stimulus went off, a light would come on. Their task was to inform the experimenter if the light came on. The experimenters used electronic circuits to initiate the stimulus on predetermined high, low, or middle heart rates. Enhancement of perceptual accuracy was found to be related to atrial contraction. Salzman and
Jaques (1976) also performed an experiment to relate reaction time to heart rate and found no relationship.

**Literature Summary and Research Questions**

The intent in discussing the latter few experiments is to give a sampling of the kinds of research reported in the literature concerning physiological and behavioral measures and to make the point again that no comprehensive studies have been performed that relate task and signal variables to physiological and behavioral changes. Moreover, there have been no studies of the present scope that have attempted to examine temporal relationships between dependent measures.

Specifically, this study addresses the following questions:

1. Does the time at which a signal is presented during an experimental run have an effect on response time or eye, heart, or brain activity?

2. Does the presence of an audio warning signal or an artificial signal affect response time or eye, heart, or brain activity?

3. Does the length of the interval between a warning signal or an artificial signal and a subsequent signal to be detected affect response time or eye, heart, or brain activity?

4. Is a warning signal more effective than an artificial signal?
3. METHODOLOGY

Scope

This section provides a description of the experimental procedures, including the selection of the subjects, equipment, experimental design, data collection, and an outline of the analysis.

Selection of Subjects

The individuals who participated in this study were paid volunteers. They were solicited by means of a posted flyer that requested subjects for eye movement research. The volunteers were all undergraduates at Ohio State University. There were four subjects, three men and one woman. The subjects were screened for vision problems, none wore glasses or contact lens. The subjects ranged in age from 19 to 22. Initially, subjects were recruited through the use of notices posted on bulletin boards in Baker Systems Engineering Building. The notices advertised for subjects with good vision. Subjects that responded were told that the experiment would span the months from June, 1979 to September, 1979, and that they would be expected to be available for the entire period. Subjects were told that they would earn $3.00 per hour and perhaps more due to incentives.

Instrument and Measurement System

The sustained attention monitoring experiment was conducted as a real time experiment. A PDP-8 computer controlled the presentation of
display signals to the subject and recorded psychological and physiological responses during the experimental session. System block diagrams show the experimental set-up and the functional relationships in the data acquisition system in Figures 7, 8, and 9.

The experimental system performed six basic activities: (1) the presentation of signals, i.e., the monitoring task, (2) recording of eye movement data, (3) recording of EEG data, (4) recording of EKG activity, (5) synchronization of the video tape, and (6) the recording of subject response times.

**Computer-Data Collection and Experimental Supervisory Control**

A Digital Equipment Corporation PDP8/L12 bit minicomputer was used to supervise the experiment, provide timing pulses to synchronize various measuring equipment, and to record and store experimental data in a real-time experiment.

Figure 10 shows the computerized experimental control and data flow within the measurement system. Central to the system was an assembler language program that supervised the experiment. A general Fortran IV program was written that generated an input file. The input file controlled the timing of the signals displayed on the meters, what physiological and psychological data would be collected, and when the data were collected. The output file contained all the responses of the subject collected during an experimental session. These responses in turn were screened and pre-processed in order to get subject data, such as a heart rate, and eye movements ready for analysis.
Figure 7. Physical Arrangement of the Display, Subject Station and Instruments Used in the Real Time Sustained Attention Experiment.
Figure 8. Block Diagram of the Information Flow, Laboratory Logic and PDP-8 Computer Interfaces in the Sustained Attention Experiment.
Figure 9. Functional Measurement and Analysis System for Response Time, Brain, Eye Movement, and Heart Signals.
Figure 10. Computerized Experimental Control and Data Flow Within Measurement System.
Display Panel and the Presentation of Signals

Two standard ammeters, each with a range of 0-100 UA, were mounted in a plywood display board. The display board was 60 inches long and 32 inches wide. The meters were separated by a distance of 21 inches. This distance assured that the meters were located outside the effective range of subject peripheral vision (> ± 10°) but within the range of the eye movement measurement system limitations (± 20°). The meters were driven by a Wavetek--Model 102 function generator which was controlled by the PDP-8 computer. A Fortran program generated a file of signals and non-signals which was stored in the PDP-8 computer. Figure 11 shows an example of a non-signal and a signal. The top of the figure shows the meters fluctuating around an average value of 50. The bottom part of the figure shows an example of a signal. The meter on the left is swinging about an average value of 45 while the meter on the right continues to swing about an average value of 50. The pointer settings for non-signals were uniformly distributed with a mean of 50 at the twelve o'clock position of the meter and a variance of 8.3. Introduction of a signal causes the mean of one of the meters to shift by five points (the meter mean value shifts to either 45 or 55). The direction of the shift of the meter mean value is randomized. The shift in the mean does not affect the variability.

Eye Activity Video Recording System

A capability was built into the display board for recording eye activity on video tape. A 200 mm Pentax telephoto lens was mounted on the front of a Shibaden Model Al 4101 one inch vidicon. The lens and
Figure 11. Signal and Non-Signal Representation.
camera were connected by a specially machined one-half inch adapter coupling. The entire assembly was mounted on a bracket behind the display board to pick up and record the movement of the subject's right eye. The vidicon and telephoto lens could be adjusted vertically and horizontally slightly to accommodate subjects with different physical characteristics. The lens-vidicon system were adjusted so as to be on a horizontal axis through the subject's right eye. The vidicon was connected to a Thalner Model 405 Counter/Timer and a Sony AV-3600 video recorder. This closed loop video system enabled the experimenter to record the movements of the subject's right eye and blink behavior. In addition, the time at which visual events occurred was also recorded on the video tape during an experimental session. The close circuit television system was designed into the experimental system for two reasons: (1) to validate visual events recorded by the Biometrics eye movement monitor and (2) to measure pupil size. The latter purpose was never realized because it was learned that the extreme technical requirements for pupil size measurement could not be provided by this particular measurement system.

**Eye Movement Recording System**

The subjects eye movements were measured on-line by a Biometrics, Inc. SGHV/2 eye movement monitor. The Biometrics system has a capability of measuring horizontal and vertical eye movements. The measurement range for horizontal movements is ± 20°. with a resolution of approximately one-quarter of a degree. Only horizontal movements were studied in this experiment. The resolution can be improved to one
minute of arc with rigid head mounting (chin rest or bite board). In this experiment the subject used a chin rest and a head restraint system. The Biometrics monitor employs a non-contacting photoelectric technique. The outputs are analog voltages that can be digitized for recording on a PDP-8 computer. The subject is required to wear special spectacles that have infra-red light sources and photodiode transducers. The photocells detect the changes in reflected light between the white sclera and the left and right sides of the iris. The output voltage of the Biometrics monitor was sent through an A/D converter and recorded on PDP-8 RK0-5 computer disk.

Cardiac Recording System

The cardiac monitoring system consisted of a modified Taber Telecardio model 202-4 amplifier. The heart signal was picked up by standard disposable leads. The leads were placed in a bipolar configuration (two active electrodes and a reference electrode, usually designated by convention as RA, LA, and RL). The cardiac signal or waveform is shown in Figure 12. A complete cycle of the heart typically consists of PQRST waveforms. The highest amplitude wave is the R-wave of the QRS segment of the cycle. The QRS segment is associated with ventricular depolarization. The amplifier was modified to pick up the continuous raw heart signal as well as a digitized version of the heart signal. In this experiment only the digitized version of the signal was recorded. The amplifier had an adjustable threshold that was set low enough to consistently pick up the R-wave but high enough so that a digitized heart pulse would not be triggered by a P- or T-wave.
Figure 12. Components of the Heart Signal Waveform During a Cardiac Cycle, Threshold Setting and Temporal Characteristics.
If the system would trigger on P or T Waveforms, spurious results would occur, i.e., the calculated heart rate would be erroneous as well as the heart rate variability. The raw waveform was only used to aid in the experimental set up. If difficulty was encountered in the set-up for heart signal recording then the raw signal could be viewed on an oscilloscope to determine the characteristics of the signals that were causing difficulty.

The configuration of leads that were found to result in the largest P-wave is shown in Figure 13. One electrode (RA) is at the level of the second intercostal space just to the right of this sternum. The other electrode (LA) is near the sixth or seventh intercostal space along the left mid-clavicular line. The RL electrode is a reference electrode and its position is not critical.

The R-wave is the largest and most dominant portion of the heart signal and as a result is used to determine heart rate and heart variability. Heart rate is generally calculated by determining the time between successive R-waves. In this research the important heart signal parameters were calculated off-line.

**Recording Brain Wave Activity**

Brain wave activity (EEG) was measured for a total of fifteen seconds around each signal presented to a subject; brain activity was measured five seconds prior to a signal onset, five seconds during the presentation of a signal, and five seconds after the cessation of a signal. The electrical signal associated with brain wave activity was sensed by a Biosone II professional quality alpha/theta monitor (Edmund
Figure 13. Monitoring Points on Chest Area of Subject.
Scientific). The alpha/theta monitor has two jacks that enable the researcher to record either the raw brain signal or a filtered version of either the alpha or theta frequency content of the brain signal. The brain signal monitor utilized three external electrodes: (1) an ear lobe electrode, (2) back of head occipital lobe electrode, and (3) a forehead frontal lobe electrode. The system of electrode placement used is the system recommended by the International Federation of Societies for Electroencephalography and Clinical Neurophysiology, namely the 10-20 System. The 10-20 System is based upon measurements from four standard points on the head: the nasion, the inion, and the left and right preauricle points. The electrode placed on the back of the head was placed up from the inion in the region of the scalp designated as the occipital-parietal region. Specifically, the rear electrode was placed over the O2 region which is directly over the occipital cortex.

These three electrodes were held in place by a rubber adjustable headband. Good contact was maintained between the scalp of the subject and the monitor electrodes by using standard electrode creme. The subject wore the electrodes throughout the experiment.

The monitoring system utilized filters to limit interference from 60 cps noise. In this research the filtered raw brain signal and the theta logic signal were sent to the PDP-8/L computer for recording and storage. The raw analog brain signal was sampled at a maximum rate of 100 times per second and this sampled value was stored in the computer for later analysis.
Research Design

The research design for this exploratory research study is shown in Figure 14. On day one, two baseline runs are presented to each S. A baseline run is an experimental run in which S is presented with three signals for detection. No alerting signals precede the signal. One day two, visual 1 and auditory 1 experimental sequences are presented to the S. Visual 1 sequence consists of five visual signals that require a detection response. Two of the visual signals are preceded by visual signals in close proximity. Auditory 1 sequence consists of three visual signals presented for detection. Three of these signals are presented for detection. Two of these signals are preceded by an auditory alerting signals. On day three, the S is presented with visual 2 and auditory 2 experimental runs. These runs are similar to the runs presented on day two.

The independent variables in this study were:

1. Type of alerting signal
   a. None
   b. Auditory warning signal
   c. Visual artificial signal

2. Nominal alerting signal or artificial signal warning interval
   a. 5 seconds
   b. 60 seconds
   c. 250 seconds
   d. 315 seconds
NOTE: Up arrow indicates an alerting signal; down arrow is a signal presented for detection.

Figure 14. Experimental Signal Presentation Schedule (Seconds Into Run).
3. Time of signal presentation in an experimental run
   a. $T_1$ (early)
   b. $T_2$ (middle)
   c. $T_3$ (late)

The dependent variables in this experiment were:
1. Eye movement activity and fixation duration
2. Mean heart rate and heart rate variability
3. Brain wave amplitude and frequency measures
4. Response time

The experimental design is a repeated measures design. This particular design allows for fewer subjects and the partitioning of subject variability. The partitioning of subject variability results in a smaller error term but with fewer degrees of freedom. If individual differences are large as they usually are in behavioral research, blocking on subjects can result in a powerful and efficient design.

One of the usual problems with this kind of design is carryover effects. Carryover effects are residual treatment effects that linger after a treatment is administered and affects responses to subsequent treatments. Carryover effects can often be controlled by counterbalancing and maintaining large time intervals between treatments. The purpose of this research, however, is to investigate the carryover effects of previous signals.

A serious deficiency of this design is the lack of balance that arises because of unbalance. The seriousness of this problem was not
anticipated at the outset of the research. Lack of balance arises because of missed signals and an unbalanced initial design. The results of an unbalanced design is the confounding of effects, difficulty in estimating effects, and nonestimability of effects. Missing data and unbalanced designs force the researcher to develop ad hoc analysis strategies. Validity of conclusions are threatened due to the low statistical power of tests. Low statistical power arises because of smaller sample sizes that come about in "piecemeal" analysis. In general, the unbalanced design is less efficient than a balanced design. These design aspects will be further discussed in subsequent sections of this dissertation.

Experimental Procedures

The experiment consisted of three phases: (1) collection of personal and skill performance data, (2) collection of baseline data (pre-testing), and (3) collection of experimental data under various treatment conditions. Ss were given a vision test, an embedded figures test (EBT), and reaction time tests (RT) under 0, 2, 4, and 8 bits of uncertainty.

After personal data and skill performance data were collected, Ss were either dismissed from the study or scheduled for baseline testing.

Baseline Testing. Each S was tested twice under baseline conditions on the first day. Baseline conditions consisted of testing each S using the signal presentation schedule shown in Figure 14. Subject instructions were recorded on a tape recorder and presented in
the same manner to each S. Subject instructions are contained in the Appendix. During baseline testing a visual signal was presented to each S at 315, 635, and 955 seconds to which the S was instructed to respond upon detection. A visual signal was defined as the occurrence of the two meter needles swinging around different average values. Each signal had a duration of 5 seconds. Each S was instrumented for brain, heart, eye movement, and detection response signals. The PDP-8 computer was used for both experimental control and data acquisition. Data was initially recorded on an RK05 disk (later data was transferred to floppy disks, non-standard magnetic tapes, and standard magnetic tapes).

Experimental data collection was similar to baseline testing except the S's instructions were modified and the following treatment variables were introduced: (1) type of warning signal (none, audio or visual--the visual warm-up signal was similar to the signal it preceded), (2) alerting signal warning interval (0, 5, 60 seconds), and (3) elapsed time in the experimental run (early, middle, or late). As with baseline testing, brain, heart, eye movement, and detection response signals were recorded. A computer-driven audio buzzer was installed on top of the display board to provide an audio warning signal to the S. The signal presentation schedule was shown in Figure 14. Signal duration for all signals was 5 seconds. Eye movement and heart data were collected for the entire twenty-minute run whenever an event occurred. Brain waves, however, were sampled every 0.01 seconds--five seconds prior to a signal, five
seconds during a signal, and five seconds after a signal. This restriction on brain wave sampling was necessary due to the control and memory limitations of the PDP-8 computer. Each S was given the experimental runs in the following order: (1) visual sequence 1, (2) auditory sequence 1, (3) visual sequence 2, and (4) auditory sequence 2. Each run lasted 20 minutes. Owing to the extensive set-up and calibration time required for a run and the stressful nature of the experiment, Ss were asked to perform visual sequence 1 and auditory sequence 1 on day 2 and visual sequence 2 and auditory sequence 2 on day 3. A simple response analysis program was written in Fortran IV to analyze S responses in order to pay subjects according to their performance. All other analyses were done off-line on other computers.

**Analysis Strategy**

Figure 15 presents a flow diagram of the general research analysis strategy used to examine and test the research hypotheses.

Raw data events were collected and stored on a computer disk. This means that every time a heart beat occurred or the eyes made a change from one display segment to another or a detection response occurred, the time of that event was recorded. Whenever the computer was instructed to acquire brain signal data, the signal was sampled every 0.01 seconds.

Raw data must be pre-processed to varying degrees depending upon the particular signal. At this stage of analysis, policies are established concerning the definition and treatment of artifacts, outliers, and data errors.
Figure 15. General Research Analysis.
The next stage of data analysis involves the analysis of response time data. This involves a pre-processing of data from the raw data file, i.e., scanning the raw data file for response and relating these responses to the signal generation input file (see Figure 10 which outlines experimental control and data flow in the measurement and analysis system) which controls the monitored display. This response data is analyzed to determine whether a response is a miss, false alarm (FA), or a hit. The type of response made by the subject will dictate the extent of analysis that can be performed at a subsequent stage. For example, if a FA is made, it is impossible to relate this response to a display signal. Furthermore, if a FA response is made it cannot be related to brain signals since the brain signal acquisition system is not turned on until five seconds prior to the presentation of a display signal. Consequently, the only meaningful analysis that can be performed is an eye movement and heart signal analysis. In the case where a correct detection is made (a hit) a complete analysis can be performed for all data collected. At this stage of analysis, decisions concerning all other analyses are keyed to the response time analysis and the response data is in an usable form for further detailed analysis.

The previous discussion outlined the need to key the analysis to the subject's behavioral responses. There are two basic kinds of analyses: (1) signal promixity analysis, and (2) long term performance analysis.
Signal proximity analysis refers to analysis of performance measures around the time that a signal is presented. In the case of brain signal analysis, descriptive and inferential statistics are developed for a before-after-during signal analysis. The effect of the independent variables on the performance measures are analyzed. In some instances, where it is meaningful, a time-series analysis is performed. Signal proximity analysis can be considered to be microscopic analysis.

Long term performance analysis is a study of changes that take place over the entire experimental run in the performance measures. One might consider such analysis as a macroscopic analysis. One can then look for differences or similarities that might occur in the various measures early, intermediate, or late in the experiment. Again, one can use descriptive and inferential statistics to aid the analysis task. Signal analysis in both the time and frequency domains might be useful at this point especially for heart and brain performance measures. A combined analyses of dependent measures is a part of this stage of the analysis. This means that the interrelationships of signals and performance measures are examined along a common time base.

This dissertation documents the more detailed analyses strategies used in the research in the chapter entitled "Results."
4. RESULTS OF THE RESEARCH

Introduction

The results from this research are presented in five sections: (1) response time analysis results, (2) EEG analysis results, (3) eye movement analysis results, (4) heart rate analysis results, and (5) combined analysis results. The analysis follows the broad research strategy outlined in Figure 15. Detailed analysis strategies are outlined in terms of a flow diagram in each analysis results section.

A caveat should be mentioned with regards to the inferential analysis. Sufficient planning of the experiment must occur prior to the implementation of a research design. What is sufficient planning? The research planning process should recognize limitations in resources but at the same time recognize the need to collect enough observations under controlled conditions to make inferences. Insufficient planning may conserve resources during the data collection stage but costs may increase in the analysis stage due to the ad hoc analyses required to cope with the design. The design may cause several problems: (1) special analyses are required to cope with problems of confounding and unbalance, (2) unbalanced designs generally require increased computer use to estimate effects, and (3) effects may not be estimable due to the confounding that results from unbalanced designs. All research has deficiencies. Deficiencies do exist in the present research and the research design has dictated and imposed severe restrictions on the analysis.
Response Time Analysis Overview

The research strategy used to analyze response time data is outlined in Figure 16. The response time data was read from the original data file and used to set up a file of only response time data for analysis. Descriptive analyses were performed which consisted of calculating descriptive statistics (means and variances), graphical analyses and an AID3 analysis. A Kolmogorov-Smirnov one-sample test was used to test the response time cumulative distribution against a cumulative normal distribution. The descriptive analyses were followed by inferential analyses. The inferential analysis consisted of the application of general linear models to the data. Prior to the general linear models analyses, analysis of variance assumptions were checked to determine the suitability of this kind of analysis for the data.

The response time data was a set of eighty observations coded according to five independent variables. The independent variables and the levels of each variable used in this analysis are in Table 4. It should be noted that for earlier analyses such as AID3, X3 (inter-signal interval) was coded in a slightly different manner than is shown in Table 4. For that analysis, X3 had only three values: X3 = 1 (any signal that did not have a prior signal or any signal that had a prior signal greater than 60 seconds), X3 = 2 (signals preceded by an auditory signal or visual artificial signal with an ISI of 5 seconds), and X3 = 3 (signals preceded by an auditory signal or visual artificial signal with an ISI of 60 seconds).
Figure 16. Response Time Analysis Strategy.
Table 4. Independent Variables and Variable Levels Used to Obtain Response Time Analysis Results.

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>Variable Levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signal presentation time (X1)</td>
<td>1 = 305 seconds</td>
</tr>
<tr>
<td></td>
<td>2 = 315 seconds</td>
</tr>
<tr>
<td></td>
<td>3 = 570 seconds</td>
</tr>
<tr>
<td></td>
<td>4 = 625 seconds</td>
</tr>
<tr>
<td></td>
<td>5 = 635 seconds</td>
</tr>
<tr>
<td></td>
<td>6 = 895 seconds</td>
</tr>
<tr>
<td></td>
<td>7 = 955 seconds</td>
</tr>
<tr>
<td>Type of prior signal (X2)</td>
<td>1 = no warning signal; no artificial signal</td>
</tr>
<tr>
<td></td>
<td>2 = auditory warning signal</td>
</tr>
<tr>
<td></td>
<td>3 = visual artificial signal</td>
</tr>
<tr>
<td>Inter-signal interval (ISI), (X3)</td>
<td>1 = 5 seconds</td>
</tr>
<tr>
<td>(X3) - X3 is derived from X1 by</td>
<td>2 = 55 or 60 seconds</td>
</tr>
<tr>
<td>differencing.</td>
<td>3 = 250 or 255 seconds</td>
</tr>
<tr>
<td></td>
<td>4 = 305 seconds</td>
</tr>
<tr>
<td></td>
<td>5 = 315 seconds</td>
</tr>
<tr>
<td>Subject (X4)</td>
<td>1 = S1</td>
</tr>
<tr>
<td></td>
<td>2 = S2</td>
</tr>
<tr>
<td></td>
<td>3 = S3</td>
</tr>
<tr>
<td></td>
<td>4 = S4</td>
</tr>
<tr>
<td>Order of run (X5)</td>
<td>1 = Run 1 (Baseline 1)</td>
</tr>
<tr>
<td></td>
<td>2 = Run 2 (Baseline 2)</td>
</tr>
<tr>
<td></td>
<td>3 = Run 3 (V1)*</td>
</tr>
<tr>
<td></td>
<td>4 = Run 4 (A1)*</td>
</tr>
<tr>
<td></td>
<td>5 = Run 5 (V2)*</td>
</tr>
<tr>
<td></td>
<td>6 = Run 6 (A2)*</td>
</tr>
</tbody>
</table>

*See Figure 3.9 for definition of these designations.
A review of Figure 14 shows some inconsistencies in the signal schedule. Sometimes the intended ISI of 60 seconds is only 55 seconds and likewise the intended ISI of 250 seconds is sometimes 255 seconds. However, these errors in the signal schedule design do not seem to create any serious analysis or interpretation problems.

Response Time Data. Response time latencies are shown in Table 5 by run, subject, and time at which a signal occurs. Prior signals are shown by type of signal. These are either auditory warning signals or visual artificial signals. Auditory warning signals did not require a button-pushing response whereas visual artificial signals did require a button-pushing response. Missing signals and outliers are also indicated in the table.

AID3 Analysis. The AID3 program was developed by Sonquist, Baker, and Morgan (1973). The computer program is a part of the OSIRIS software system and stimulates the strategy that a researcher might use in searching for predictors that explain the variance of a dependent variable. The procedure is a multivariate procedure that has certain advantages over other techniques such as regression analysis, analysis of variance, or discriminant analysis. AID3 eases the assumptions concerning linearity and additivity. The procedure uses the least squares technique in the algorithm and the focus is on the importance of relationships in the data rather than the statistical significance of relationships. The AID3 algorithm starts with the predictor variable that explains the greatest amount of variation in the dependent variable and then makes a dichotomous split on this
Table 5. Summary of Response Time Data for Each Subject and Run.

<table>
<thead>
<tr>
<th>Run</th>
<th>Subject</th>
<th>305</th>
<th>315</th>
<th>570</th>
<th>625</th>
<th>635</th>
<th>895</th>
<th>955</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (B1)</td>
<td>1</td>
<td>4.03</td>
<td>1.18</td>
<td>3.31</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1.49</td>
<td>2.85</td>
<td>2.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>4.87</td>
<td>1.23</td>
<td>1.09</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>3.13</td>
<td>3.53</td>
<td>2.04</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 (B2)</td>
<td>1</td>
<td>1.11</td>
<td>1.37</td>
<td>1.62</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2.13</td>
<td>2.52</td>
<td>4.78</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.96</td>
<td>5.68**(x)**</td>
<td>2.55</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>----</td>
<td>1.15</td>
<td>4.27</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 (V1)</td>
<td>1</td>
<td>2.28</td>
<td>2.00**(x)**</td>
<td>1.76</td>
<td>2.41**(x)**</td>
<td>5.28**(x)**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1.98</td>
<td>3.50</td>
<td>1.83</td>
<td>1.52</td>
<td>2.56</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>2.79</td>
<td>1.74</td>
<td>1.18</td>
<td>2.41</td>
<td>2.67</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>2.82</td>
<td>1.85</td>
<td>2.69</td>
<td>1.58</td>
<td>2.30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 (A1)</td>
<td>Auditory</td>
<td>----</td>
<td>Auditory</td>
<td>----</td>
<td>----</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Warning</td>
<td>2.16</td>
<td>Warning</td>
<td>2.57</td>
<td>1.86</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>2.08</td>
<td>1.64</td>
<td>4.72</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 5. (continued).

<table>
<thead>
<tr>
<th>Run</th>
<th>Subject</th>
<th>305</th>
<th>315</th>
<th>570</th>
<th>625</th>
<th>635</th>
<th>895</th>
<th>955</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 (V2)</td>
<td>1</td>
<td>1.63(*)</td>
<td>3.13</td>
<td>1.95(*)</td>
<td>1.27</td>
<td></td>
<td>4.86</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2.88</td>
<td>1.45</td>
<td>2.24</td>
<td>1.32</td>
<td></td>
<td>2.71</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>----</td>
<td>1.39</td>
<td>2.34</td>
<td>1.98</td>
<td></td>
<td>----</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>1.37</td>
<td>1.62</td>
<td>1.67</td>
<td>1.17</td>
<td></td>
<td>1.89</td>
<td></td>
</tr>
<tr>
<td>6 (A2)</td>
<td>1</td>
<td>2.05</td>
<td></td>
<td></td>
<td></td>
<td>Auditory</td>
<td>1.13</td>
<td>Auditory</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>3.76</td>
<td></td>
<td></td>
<td></td>
<td>Warning</td>
<td>1.54</td>
<td>Warning</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>----</td>
<td></td>
<td></td>
<td></td>
<td>↑</td>
<td>1.80</td>
<td>↑</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>2.01</td>
<td></td>
<td></td>
<td></td>
<td>↓</td>
<td>1.28</td>
<td>↓</td>
</tr>
</tbody>
</table>

(*) Visual Artificial Signal.

(**) Response times exceeding 5 seconds are regarded as outliers—missing data.
variable in an optimal way (that is the various levels of the predictor variable are split into two mutually exclusive dichotomous groups in a way that accounts for the greatest amount of variance in the dependent variable). The procedure works down to less and less dependable findings on smaller and smaller subgroups. The variance analysis used by AID3 is a sequential one-way analysis of variance that is simple, robust, and easy to understand. AID3 is useful on large data sets that are not orthogonal and where the dependent variable is reasonably well-behaved (few extreme cases and no severe bimodalities). Predictor variables should be classifications or be capable of being made into classifications. The program operates sequentially, has a look ahead option, imposes a minimum number of assumptions on the data and does a great deal of searching. The algorithm does require that some parameters of the search be specified so that it is a reproducible result.

An AID3 analysis was performed on the data using a minimum group size of five. Response time was the dependent variable and the independent variables listed in Table 4 were used in the analysis. Figure 17 shows the tree structure that resulted from the AID3 analysis. Response time variability was split into nineteen different groups. The variable that is split is indicated on the figure. The levels of each variable and how these levels are grouped in each dichotomous split are also indicated at each split. In addition, the number of samples in each group and group mean and variance are included in the figure. Several observations can be made concerning the tree diagram.
Note: This analysis is based on data which had a coding error for one subject's data in one run. ISI's are also coded differently in later analyses. AID3 is no longer available at UTCC.

Figure 17. Aid 3 Analysis for Response Time.
The initial split is on variable X5 (order of run). This variable accounts for the greatest amount of variability in response time. The procedure groups levels of X5 into two dichotomous groups. Group 3 (levels 1, 2, 3, 4) and Group 2 (levels 5, 6). This result suggests that some learning takes place in the experiment between early and late runs; although this result may not be statistically significant in a standard analysis. This notion is supported also by the lower mean and less variable response time for Group 2 (levels 5, 6).

Another observation that must be made is that the prior signal interval variable (X3) does not enter into the analysis. It should be noted that the variable X3 was coded as X3 = 1 (ISIs > 60 seconds), or X3 = 2 (ISI = 5 seconds), or X3 = 3 (ISI = 60 seconds). The variables X1 (signal presentation time) and X2 (type of warning signal) come into the analysis early and can be considered to be dependable findings (splits farther down the tree are on smaller sample sizes and the variables account for a smaller proportion of the response measure variability). Group 2 splits X2 into two groups, Group 16 and Group 17. Group 16 contains levels (2, 3) and Group 17 contains level (1). Levels 2 and 3 indicate the presence of an auditory or visual prior signal whereas level 1 is the control condition—no prior signal.

Subject differences (X4) contribute less to response time variance than order of run, the time at which a signal was presented, and the type of prior signal. Another result worth noting concerns the variable X1 (time at which the signal was presented). X1 is split early in the tree diagram. Earlier signal presentation effects are grouped together into Group 4 (levels 1, 2, 3, 4, 5, 6) whereas the effect of
late signal presentation (level 7) is split into a separate group (Group 4). The mean and variance for Group 4 (early signals are \( Y = 2.27 \) and \( \text{VAR} = 1.30 \). The mean and variance for Group 5 (late signal) are \( Y = 3.13 \) and \( \text{VAR} = 1.84 \). The effect of the presentation of late signals on response time in this analysis is to increase the mean and variance of response time.

Response Time Distribution and Summary Statistics. The data was examined using standard descriptive analysis techniques. Figure 18 shows a stem and leaf plot, boxplot, normal probability plot, and summary statistics for response time. The stem and leaf plot is a histogram of the data except that more detailed information is shown. This particular computer-generated stem and leaf plot rounds values down to the next lower stem category if the exact category is not computed. For example, the response time 0.96 appears in the stem category "8". The stem and leaf plot shows some lack of symmetry; the distribution is skewed negative (skewed to left in a conventional histogram). The boxplot also shows this skewness towards lower values, i.e. the mean is greater than the median.

The normal probability plot indicates some departure from normality. If the data were perfectly normally distributed, the normal probability plot would be a straight line.

Since only qualitative judgments can be made about the distribution of the data from the stem and leaf plot and normal probability plot, a Kolmogorov-Smirnov (K-S) test was performed. The K-S test statistic is defined as \( D(n) = \text{maximum} \left[ F(x) - S(n) \right] \). The maximum
### Summary Statistics for Response Times

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>N</th>
<th>SUM WXTS</th>
<th>MEAN</th>
<th>STD DEV</th>
<th>VARIANCE</th>
<th>KURTOSIS</th>
<th>SKEWNESS</th>
<th>CV</th>
<th>STD EV</th>
<th>VARIANCE</th>
<th>LOWEST</th>
<th>HIGHEST</th>
</tr>
</thead>
<tbody>
<tr>
<td>RT</td>
<td>78</td>
<td>169.64</td>
<td>2.17</td>
<td>0.94</td>
<td>0.93</td>
<td>1.18</td>
<td>0.91</td>
<td>1.96</td>
<td>0.97</td>
<td>0.93</td>
<td>0.54</td>
<td>4.27</td>
</tr>
</tbody>
</table>

**Moments**

- Max: 99
- 90th Percentile: 90
- 75th Percentile: 2.59
- 50th Percentile: 1.99
- 25th Percentile: 1.48
- Median: 1.68
- Lower Quartile: 1.98
- Upper Quartile: 1.11
- Interquartile Range: 0.58
- Standard Deviation: 0.97
- Variance: 0.93
- Skewness: 1.18
- Kurtosis: 1.18

**Quantiles**

- 2.5th Percentile: 1.11
- 0.1th Percentile: 0.58
- 0.01th Percentile: 1.12
- 0.0001th Percentile: 4.07

**Box Plots, Normal Probability Plots**

Figure 18. Stem and Leaf Plots, Box Plots, Normal Probability Plots and Summary Statistics for Response Times.
difference between a sample function \( 5(n) \) and a theoretical curve \( F(x) \) is compared with a tabled value \( D_{\text{table}} \). If \( D(n) < D_{\text{table}} \), the hypothesis of normality is not rejected. \( D(n = 78) = 0.1214 \) was less than \( D_{\text{table}} = 0.1539 \) for a two-sided one sample K-S test at the 5% level of significance.

The distribution is normal for the entire data set. This analysis did not examine the distributions of the populations sampled under each treatment condition because of the limited amount of data. Application of ANOVA requires normality for each population sampled.

**Summary Statistics for Response Time by Independent Variable.**

Descriptive block graphs were developed for means and variances of response times for each level of each independent variable. These block graphs are shown in Figures 19 (means) and 20 (variances).

Run, time, and treatment information is summarized in Table 6. Sample sizes \( (n) \), means, and variances for response times are provided by time (315, 635, or 955 seconds), run (1, 2, 3, 4, 5, or 6), and treatment (\( T_1 = \) no prior signal or an ISI equal to 315 seconds; \( T_2 = \) prior auditory signal with an ISI of 5 seconds; \( T_3 = \) prior auditory signal with an ISI of 60 seconds; \( T_4 = \) prior visual artificial signal with an ISI of 5 seconds; and \( T_5 = \) prior visual artificial signal with an ISI of 60 seconds).

Response time averaged across subjects for signals at 315, 635, and 955 seconds were plotted by run. These plots are shown in Figure 21. The plots show graphically the same information presented in Table 6. Differences between runs and changes within runs at specific
Figure 19. Block Graph for Response Time Means for Each Level of Each Independent Variable.
Figure 20. Block Graph for Response Time Variance
Each Level of Each Independent Variable.
Table 6. Response Time Data Summary, Data Averaged Across Subjects, Runs, Times, and Treatments.

<table>
<thead>
<tr>
<th>Run</th>
<th>Time (seconds)</th>
<th>Summary of Run Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>315</td>
<td>635</td>
</tr>
<tr>
<td>1</td>
<td>T1</td>
<td>T1</td>
</tr>
<tr>
<td></td>
<td>N 4</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Mean 3.38</td>
<td>2.20</td>
</tr>
<tr>
<td></td>
<td>S.D. 1.45</td>
<td>1.18</td>
</tr>
<tr>
<td></td>
<td>N = 12</td>
<td>Mean = 256</td>
</tr>
<tr>
<td>2</td>
<td>T1</td>
<td>T1</td>
</tr>
<tr>
<td></td>
<td>N 3</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Mean 1.40</td>
<td>1.68</td>
</tr>
<tr>
<td></td>
<td>S.D. 0.64</td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td>N = 10</td>
<td>Mean = 2.25</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Summary of Runs 1 and 2 Time Statistics</td>
<td>N 7</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Mean 2.53</td>
<td>1.98</td>
</tr>
<tr>
<td></td>
<td>S.D. 1.52</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td>Mean = 2.42</td>
<td>S.D. = 1.26</td>
</tr>
<tr>
<td>3</td>
<td>T1</td>
<td>T4</td>
</tr>
<tr>
<td></td>
<td>N 4</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Mean 2.47</td>
<td>1.87</td>
</tr>
<tr>
<td></td>
<td>S.D. 0.41</td>
<td>0.62</td>
</tr>
<tr>
<td></td>
<td>N = 11</td>
<td>Mean = 2.26</td>
</tr>
</tbody>
</table>
Table 6. (continued).

<table>
<thead>
<tr>
<th>Run</th>
<th>Time (seconds)</th>
<th>Summary of</th>
<th>Run Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>315</td>
<td>635</td>
<td>955</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>0.81</td>
<td>0.25</td>
<td>2.02</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>0.83</td>
<td>0.33</td>
<td>1.53</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>1.00</td>
<td>0.30</td>
<td>0.48</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Run</td>
<td>Time (seconds)</td>
<td>Summary of Run Statistics</td>
<td></td>
</tr>
<tr>
<td>-----</td>
<td>----------------</td>
<td>---------------------------</td>
<td></td>
</tr>
<tr>
<td></td>
<td>315</td>
<td>635</td>
<td>955</td>
</tr>
<tr>
<td>1, 2, 3, 4,</td>
<td>N</td>
<td>21</td>
<td>22</td>
</tr>
<tr>
<td>5, 6, Time</td>
<td>Mean</td>
<td>2.26</td>
<td>1.80</td>
</tr>
<tr>
<td>Statistics</td>
<td>S.D.</td>
<td>1.08</td>
<td>0.68</td>
</tr>
</tbody>
</table>

**Treatment Summary:**

- T1 (no prior signal or ISI = 315 seconds) 34 2.56 1.20
- T2 (5 second auditory) 7 1.45 0.51
- T3 (60 seconds auditory) 7 1.89 0.54
- T4 (5 second visual) 8 1.88 0.68
- T5 (60 second visual) 7 1.88 0.64
Figure 21. Plot of Response Times by Run, Time, and Treatment (T1 = no Prior or 315 Seconds, T2 = 5 Second Auditory Warning Signal, T3 = 60 Second Auditory Warning Signal, T4 = 5 Second Visual Artificial Signal, T5 = 60 Second Visual Artificial Signal.)
times can be observed in these plots. Average response times are plotted for runs 1 and 2 combined and for all runs combined versus time. It should be noted that time trends within runs can only be determined from runs 1 and 2. Time effects, if there are any, are confounded with treatment effects in runs 3, 4, 5, and 6. Moreover, one of the problems with this research design can be observed from the plots as well as in Tables 5 and 6 - treatment 1 is overrepresented at times 315 and 955, i.e. T1 occurs more often at times 315 and 955. Other treatments are underrepresented at the various times. The overall design is unbalanced with regards to treatments. This shortcoming in the original experimental design imposes limitations on the analysis and forces the use of an ad hoc analysis strategy. Inferential analysis strategies and results are discussed in the following section.

Inferential Analysis of Response Time Data

The strategy developed to make inferences concerning response times consisted of the following stages:

1. **Run-time effects analysis**—run effects and time effects within runs have been assumed to be insignificant. Some empirical support exists for these assumptions although the evidence is not strong due inadequacies of the research design. The analysis strategies employed do not completely circumvent possible confounding problems.

2. **Treatment effects analysis**—treatments are T1 (no prior signal or ISI = 315 seconds), T2 (prior auditory signal with an ISI of 5 seconds), T3 (prior auditory signal with an ISI of 60 seconds), T4 (prior visual artificial signal with an ISI of 5 seconds), and T5 (prior visual artificial signal with an ISI of 60 seconds). The treatment analysis is based upon an assumption of insignificant run and time effects although evidence to support the validity of this assumption is limited.
(3) **Intersignal interval effects analysis**—like the previous treatment analysis, the ISI analysis is based upon the assumption of insignificant run and within-run time effects. Plots and a one-way ANOVA were used to assess ISI effects.

**Run-time Effect Analysis.** A general linear model procedure was used to analyze run effects and time trends for runs 1 and 2. Runs 1 and 2 are the only runs in which treatment effects are not confounded with time and run effects. The analysis assumed an $S \times A \times B$ repeated measures design where $S$ is a fixed subject effect and $A$ and $B$ are fixed time and run effects. The design is blocked on subjects and since subjects receive both runs 1 and 2 and signals at times 315, 635, and 955 seconds, runs, times, and subjects are all crossed. An $S \times A \times B$ is a noise reducing design since subject effects are partitioned out of the error. The model, $X_{abs} = M + a_a + b_b + s_s + ab_{abs} + as_{as} + bs_{bs} + abs_{abs} + e_{abs}$ confounds error with the highest order interaction. Subject effects are generally of little interest other than as a blocking factor to reduce error. Table 7 is a summary of the analysis. Although the model accounts for 77% of the dependent variables variation, the model is not significant since $F_{model} = 0.78$, $p > 0.6830$. Runs 1 and 2 do not differ and mean response times at 315, 635, and 955 seconds do not differ. The (time * run) interaction was not significant. Technically, the order of runs is fixed and there is a problem of induction—a lack of run effects cannot be generalized to runs 3, 4, 5, and 6.

Isolation of run and time effects in an overall analysis is really not possible for runs 3, 4, 5, and 6 for three reasons: (1)
Table 7. Analysis of Variance for Response Times for Runs 1 and 2 Using an S x A x B Design.

<table>
<thead>
<tr>
<th>Source of Variance</th>
<th>df</th>
<th>Sum of Squares</th>
<th>Mean Squares</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time (X1)</td>
<td>2</td>
<td>3.56</td>
<td>1.78</td>
<td>1.63</td>
<td>0.2722</td>
</tr>
<tr>
<td>Subject (X4)</td>
<td>3</td>
<td>2.32</td>
<td>0.77</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Run (X5)</td>
<td>1</td>
<td>2.09</td>
<td>2.09</td>
<td>1.08</td>
<td>0.3754</td>
</tr>
<tr>
<td>Time * Run</td>
<td>2</td>
<td>9.75</td>
<td>4.88</td>
<td>2.52</td>
<td>0.1961</td>
</tr>
<tr>
<td>Time * Subject</td>
<td>6</td>
<td>6.56</td>
<td>1.09</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Run * Subject</td>
<td>3</td>
<td>5.81</td>
<td>1.94</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Subject * Time * Run</td>
<td>4</td>
<td>7.75</td>
<td>1.94</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Model: $X_{abs} = M + a + b + S + ab + as + bs + abs + e_{abs}$

Model F-value $= (25.77/17) ÷ (7.75 (4) = 0.78$, $p > 0.6830$

Model $R^2 = (Model SS ÷ Total SS) = (25.7684 ÷ 33.5222) = 0.7687$
treatment effects within a run are confounded with time effects, (2) run effects are confounded with treatment and time effects because of a lack of balance in the design, and (3) sample sizes are small. There is no wholly satisfactory way out of this dilemma. The analysis can proceed on the basis of several simplifying assumptions, namely that run and time are not significant. The remainder of the analysis is based on the validity of these assumptions. How reasonable are the assumptions and what support can be generated for these assumptions? The run statistics shown in Table 6 indicate that run mean response times for runs 3, 4, 5, and 6 are nearly the same as for runs 1 and 2. If fact, a one way ANOVA across all runs confirms that the run means are equal, \( F = 0.63 \) compared with \( F_{0.95, 5, 60} = 2.37 \) despite time and treatment confounding. If the run mean times for runs 3, 4, 5, and 6 (runs that have confounded effects) are the same as for runs 1 and 2 where run effects are measured without confounding, it may be that confounded time-treatment effects within a run contribute little to each run mean. Another possibility is that the test lacks power to discriminate run effects due to small sample sizes. In fact, the power \((1-\beta)\) of the test is low (less than 0.30).

A one-way ANOVA was performed for response times at 315, 635, and 955 seconds across all runs and treatments. This analysis resulted in an \( F \)-value with borderline significance \( (F = 3.15 \) compared to \( F_{.95, 2, 60} = 3.15 \). Since, it is unclear whether this borderline significance is due to a time effect or a treatment effect, further examination is required. The overrepresentation of signals without proximate prior signals (treatment designation T1) at 315 and 955 seconds has been
previously mentioned as a design problem. An underrepresentation of T1 occurs at 635 seconds. Lower and less variable response times can be observed at 635 seconds and compared to 315 or 955 seconds--this tendency can be observed in Table 6 (\(\bar{X}_{635} = 1.80 \pm 0.68\)) versus \(\bar{X}_{315} = 2.26 \pm 1.08\) and \(\bar{X}_{955} = 2.58 \pm 1.21\)). Tukey's HSD test at the experimentwise error rate of 0.05 indicates that \(\bar{X}_{635}\) is different than \(\bar{X}_{955}\).

The previous result may be due to time or treatment effects or both. In order to get a better indication of the source of the difference, it was decided to perform a second one-way ANOVA on response times at 315, 635, and 955 seconds. The analysis was performed only for responses for signals without proximate prior signals (T1 designation in Table 6). This analysis assumes run effects are insignificant and permits an analysis without treatment confounding. The analysis resulted in an insignificant F-value of 1.33 (the critical value was \(F_{.95, 2, 30} = 3.32\)). Time effects are insignificant, but this test like the earlier test for runs, has low power (\(1-\beta < 0.30\)). The test cannot discriminate true differences when they occur.

The results from this analysis lend some support, albeit weak, to the notion that time and run effects are insignificant. If time and run effects do exist, the ability of the statistical tests to discriminate differences is low, i.e. large differences must occur to be detected given existing sample sizes. The null hypothesis will be falsely accepted a large proportion of the time.
Several things could have been done at the outset of the research to alleviate some of the forgoing analysis problems: (1) runs 1 and 2 could have been used as pilot runs to estimate and bound sample size requirements, (2) the number of subjects and runs should have been increased and treatments within runs counterbalanced, and (3) each run could be extended to 30 minutes to enable each treatment to be in a run.

**Treatment Effects.** Treatment effects were assessed by a one-way ANOVA. The response time data set was aggregated across runs and within runs. The F-value for the general linear model was $F = 2.70$ ($df_1 = 4$, $df_2 = 58$). This F-value was significant at the 0.0393 level. The model had an $R^2 = 0.1570$. A protected LSD multiple comparison test was used to compare treatment means. The mean for T1 (no prior signal or an ISI = 315 seconds) and T2 (prior auditory signal with an ISI of 5 seconds) were the only means significantly different at a comparisonwise alpha level of 0.05. (Simultaneous confidence interval multiple comparison tests such as the SNK test, Tukey HSD test, and Scheffe test which control experimentwise error rates at 0.05 did not show a significant effect.)

A summary of sample sizes, means, and variances by treatments are given in Table 8. The contrast $[(V/5) + (V/60)] - [(A/5) + (A/60)] = 0$ was tested with a t-test to determine if there was a modality effect. The test resulted in a t-value of $t = 0.9712$. This was compared to $|t| = 2.052$ for $\alpha = 0.025$ and df = 27. The contrast $[(V-5) + (V-60)] - [(A-5) + (A-60)]$ was found to increase sample size
Table 8. Summary of Treatment Statistics for Response Time.

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Treatment Designation</th>
<th>Sample Size</th>
<th>Sample Mean</th>
<th>Sample S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1</td>
<td>V-315</td>
<td>34</td>
<td>2.56</td>
<td>1.20</td>
</tr>
<tr>
<td>T2</td>
<td>A-5</td>
<td>7</td>
<td>1.45</td>
<td>0.51</td>
</tr>
<tr>
<td>T3</td>
<td>A-60</td>
<td>7</td>
<td>1.89</td>
<td>0.54</td>
</tr>
<tr>
<td>T4</td>
<td>V-5</td>
<td>8</td>
<td>1.88</td>
<td>0.68</td>
</tr>
<tr>
<td>T5</td>
<td>V-60</td>
<td>7</td>
<td>1.88</td>
<td>0.64</td>
</tr>
</tbody>
</table>
and power for a test of modality effects. The power of the test, however, remained low despite efforts ($1-\beta = 0.30$). The test also assumed time and run effects to be negligible. Although no difference in prior signal modality were found the test is not particularly sensitive to true differences. Furthermore, modality effects could not be tested at 315 seconds because the A-315 condition was not present in the experiment.

**Intersignal Interval Effects**

If modality effects are not significant, it is possible that intersignal intervals affect response times. Each signal is preceded by some time interval between it and the preceding signal. The time intervals between signals in this experiment were 5, 55, 60, 250, 255, 305, and 315 seconds. For ANOVA purposes, these ISIs were regarded as 5, 60, 250, and 315 seconds. These intersignal intervals (ISI) were derived from the variable $X_1$ (time at which a signal occurs) and regarded as a fixed effect.

A correlation analysis was performed for response time (RT) and ISI ($ISI = 5, 55, 60, 250, 255, 305, \text{ and } 315 \text{ seconds}$). The correlation coefficient for RT and ISI was $R = 0.348$, $p < 0.01$ ($N = 76$) when tested against $R = 0$, ($t = 3.193$, $t_{\text{critical}} (0.995, 80) = 2.638$). This result suggested that response times increased with ISI. Descriptive plots were made of RT means and variances versus ISI. These plots are shown in Figure 22. The RT mean versus ISI plot shows RT increasing with ISI. RT variance shows a slight decline in the range of ISIs of 5-250 seconds and then an increase for ISIs in the range
Figure 22. Means and Variances of Response Times Versus Intersignal Intervals.
305-315 seconds. As a result of these descriptive analyses, an ANOVA was performed for RT using ISI as an independent variable. Time and run effects were not found to be significant in earlier analyses and the data were aggregated across all times and runs for this analysis.

The ANOVA for RT and ISI showed a significant F value (F = 3.651 for $v_1 = 3$ and $v_2 = 72$, df, $\alpha = 0.0164$, $R^2 = 13.2\%$). A Tukey HSD test was used to test group means for significance. This test provided an experimentwise error rate of at least 0.05. The mean for the 5 second ISI group was significantly different than the mean of the 315 second ISI group (1.6993 versus 2.5190). A Bartlett-Box test for homogeneity of variance was performed which gave an F-value of 5.681 which was significant at the $\alpha = 0.001$ level. Although the variances were not homogeneous, ANOVA is robust for variances where one is up to 9 times larger than another in a fixed effects model (Box, 1954). A polynomial regression was fitted to the data and the highest order significant polynomial was linear with the equation $RT = 1.6895 + 0.0025 \times ISI$.

Figure 23 shows a plot of RT versus ISI with 95% confidence limits for the data. Table 9 summarizes intersignal interval-response time statistics. The power for the ANOVA was determined to be $1-\beta = 0.76$ using a procedure outlined by Neter and Wasserman (1974, p. 453) for fixed effect models.

**EEG Analysis Results**

Subjects were instrumented with an amplifier and electrodes to pick up the EEG. This was done to determine to what variables the brain signal was sensitive and to determine whether the changes in the
Figure 23. 95% Confidence Interval for Response Time Data.
Table 9. Summary of Intersignal-Response Time Statistics.

<table>
<thead>
<tr>
<th>Intersignal Interval (ISI) - Seconds</th>
<th>Sample Size N</th>
<th>Mean</th>
<th>Standard Deviation (SD)</th>
<th>Standard Error (SE)</th>
<th>95% C.I. for Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>14</td>
<td>1.6993</td>
<td>0.6440</td>
<td>0.1721</td>
<td>1.3275 - 2.0711</td>
</tr>
<tr>
<td>60</td>
<td>14</td>
<td>1.8857</td>
<td>0.5674</td>
<td>0.1517</td>
<td>1.5581 - 2.2133</td>
</tr>
<tr>
<td>225</td>
<td>8</td>
<td>2.0150</td>
<td>0.3828</td>
<td>0.1353</td>
<td>1.6950 - 2.3350</td>
</tr>
<tr>
<td>315</td>
<td>40</td>
<td>2.5190</td>
<td>1.1336</td>
<td>0.1792</td>
<td>2.1564 - 2.8815</td>
</tr>
<tr>
<td>Total</td>
<td>76</td>
<td>2.1983</td>
<td>0.9658</td>
<td>0.1108</td>
<td>1.9776 - 2.4190</td>
</tr>
</tbody>
</table>
brain signal could be related to subject alertness state. The raw seconds before a signal was presented, five seconds during the presentation of a signal, and five seconds after a signal was presented. This strategy was necessary because of memory limitations in the PDP-8 computer.

The EEG was analyzed according to Figure 24. Initially EEG files were created from the subjects master file for the purpose of further analysis. The main features of the analysis were:

1. Descriptive Analysis
   a. Univariate statistics
   b. Graphical analysis
   c. Correlational analysis of dependent measures
   d. AID3 Analysis

2. Inferential tests
   a. GLM and other inferential tests

Descriptive Analysis of the EEG. A univariate procedure from the Statistical Analysis System (SAS) was used to develop a descriptive analysis of the 500 data points for each recorded EEG segment. The descriptive analysis indicated that the brain signal is normally distributed for each segment according to a modified version of the Kolmogorov-Smirnov D-statistic developed by Stephens (1974) at a significance level of less than 0.01.

The means and variances for each 5 second EEG segment were extracted during descriptive analysis and were used as performance measures in subsequent analyses. Histograms, cumulative frequency
Preprocess raw data to extract brain signal

Descriptive Analysis of Data
- Extract $\bar{X}$ and $S^2$ as performance measures

Spectral Analysis of Data
- Extract power in $\alpha$, $\beta$, $\theta$, and $\Delta$ frequency ranges as performance measures

EGG Segment K-S Distribution Tests

Create common file of performance measures and code by treatments and subjects

Perform AID 3 analysis on performance measures

Perform GLM analysis on performance measures

Results

Figure 24. Analysis of Strategy for EEG.
distributions and summary statistics were developed for $Y_1$, $Y_2$, $\delta$, $\theta$, $\alpha$, and $\beta$. These EEG measures are defined in Table 10. Kolmogorov-Smirnov tests were used to test the normality of the distributions. The K-S test indicated normality for $\delta$, $\theta$, $\alpha$, and $\beta$ at a significance level of $<0.01$. $Y_1$ and $Y_2$ distributions were normal at the 0.10 level of significance. These tests aided in determining what inferential statistical tests were appropriate for the analysis.

**Spectral Analysis of EEG Segments.** A spectral analysis was performed for each of the EEG segments. Each brain wave segment consisted of a record five seconds long and was sampled at the maximum rate of the PDP-8 computer; every 0.01 seconds. Spectral analysis is the decomposition of a time series into frequency components. The spectral analysis procedure produced a periodogram that was smoothed by a moving average filter to produce estimates of spectral densities.

The procedure produced an output file of spectral densities for frequencies up to 50 cps. A Fortran program operated on the output file which summed the power in the delta ($\delta$), theta ($\theta$), alpha ($\alpha$), and beta ($\beta$) frequencies ranges of the spectrum. Each one of these frequency ranges was found to be associated with different alertness or information processing states from the literature survey. The power in the $\delta$, $\theta$, $\alpha$, and $\beta$ frequency ranges for each five second EEG segment was output as an EEG measure for further analysis.

**Independent Variables and EEG Measures.** A common file was created for analysis of brain wave data that contained performance
Table 10. EEG Measures.

<table>
<thead>
<tr>
<th>Performance Measure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y1</td>
<td>Mean of each brain wave segment</td>
</tr>
<tr>
<td>Y2</td>
<td>Standard deviation of each brain wave segment</td>
</tr>
<tr>
<td>δ</td>
<td>Power in 0.5-3.5 frequency range for each EEG segment (associated with sleep)</td>
</tr>
<tr>
<td>θ</td>
<td>Power in 4-7 cps frequency for each EEG segment (associated with drowsiness)</td>
</tr>
<tr>
<td>α</td>
<td>Power in 8-13 cps frequency range for each EEG segment (associated with relaxed state)</td>
</tr>
<tr>
<td>β</td>
<td>Power in 14-30 cps frequency range for each EEG segment (associated with mental activity)</td>
</tr>
</tbody>
</table>
measures derived from the descriptive analysis and the spectral analysis of the EEG. The independent variables shown in Table 11 were used in the EEG analysis. This table is the same as Table 4 used in the response time analysis except that two other variables have been added; X6, an EEG segment location indicator, and X7, a variable that indicates whether or not a subject detected a signal.

**Correlation Analysis of EEG Dependent Measures**

The results of a correlation analysis of dependent measures are presented in Table 12. The first number in each row of the table is the correlation coefficient for each row column measurement pair. The second number in each row column is the significance level of the correlation coefficient tested under the null hypothesis (i.e., \( R=0 \)).

The mean of an EEG segment (\( Y_1 \)) is independent of other performance measures. The standard deviation of an EEG segment (\( Y_2 \)) is correlated with \( \delta, \theta, \alpha, \) and \( \beta \) at the alpha = 0.0001 level. One might expect these significant correlations since the variance across these frequency ranges is contained in the total variability of the signal (\( Y_2 \)).

**AID3 Analysis of EEG.** An AID3 analysis was performed for EEG measures \( Y_1, Y_2, \delta, \theta, \alpha, \) and \( \beta \). The independent variables explained only 10.5% of the variation in \( Y_1 \) and most of this variability was attributed to subject differences and order effects. An ANOVA showed the independent variables to have no significant effect on \( Y_1 \). Because of these findings, \( Y_1 \) was dropped from further analysis.
Table 11. Independent Variables Used in the Analysis of the EEG.

<table>
<thead>
<tr>
<th>INDEPENDENT VARIABLE</th>
<th>VARIABLE LEVELS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signal Presentation Time ($X_1$)**</td>
<td>2 = 315 seconds</td>
</tr>
<tr>
<td></td>
<td>5 = 635 seconds</td>
</tr>
<tr>
<td></td>
<td>7 = 955 seconds</td>
</tr>
<tr>
<td>Type of prior Signal ($X_2$)</td>
<td>1 = No Auditory Warning Signal—No Visual Artificial Signal</td>
</tr>
<tr>
<td></td>
<td>2 = Auditory Warning Signal</td>
</tr>
<tr>
<td></td>
<td>3 = Visual Artificial Signal</td>
</tr>
<tr>
<td>Prior Signal Interval ($X_3$)</td>
<td>1 = 0 seconds</td>
</tr>
<tr>
<td></td>
<td>2 = 5 seconds</td>
</tr>
<tr>
<td></td>
<td>3 = 55 or 60 seconds</td>
</tr>
<tr>
<td>Subject ($X_4$)</td>
<td>1 = S1</td>
</tr>
<tr>
<td></td>
<td>2 = S2</td>
</tr>
<tr>
<td></td>
<td>3 = S3</td>
</tr>
<tr>
<td></td>
<td>4 = S4</td>
</tr>
<tr>
<td>Order of Run ($X_5$)</td>
<td>1 = Run 1 (Baseline 1-B1)</td>
</tr>
<tr>
<td></td>
<td>2 = Run 2 (Baseline 2-B2)</td>
</tr>
<tr>
<td></td>
<td>3 = Run 3 ($V_1$)*</td>
</tr>
<tr>
<td></td>
<td>4 = Run 4 ($A_1$)*</td>
</tr>
<tr>
<td></td>
<td>5 = Run 5 ($V_2$)*</td>
</tr>
<tr>
<td></td>
<td>6 = Run 6 ($A_2$)*</td>
</tr>
<tr>
<td>EEG Segment ($X_6$)**</td>
<td>1 = 5 second EEG segment before signal</td>
</tr>
<tr>
<td></td>
<td>2 = 5 second EEG segment during signal</td>
</tr>
<tr>
<td></td>
<td>3 = 5 second EEG segment after signal</td>
</tr>
<tr>
<td>Detection of Signal ($X_7$)***</td>
<td>1 = Detection of signal</td>
</tr>
<tr>
<td></td>
<td>2 = No detection of signal</td>
</tr>
</tbody>
</table>

*Run designations $V_1$, $A_1$, $V_2$, and $A_2$ are described in Figure 3.9

**Variables $X_1$ and $X_6$ were combined into one time variable in the inferential analysis with values 1, 2, ..., 9.

***$X_7$ is not really an IV but a performance-related variable.
Table 12. Correlation Coefficient (r) Matrix of EEG Measures.

<table>
<thead>
<tr>
<th>EEG Dependent Measure</th>
<th>Y2</th>
<th>δ</th>
<th>θ</th>
<th>α</th>
<th>β</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y2</td>
<td>-0.03828</td>
<td>-0.06272</td>
<td>0.00343</td>
<td>0.05542</td>
<td>-0.01070</td>
</tr>
<tr>
<td>0.6100</td>
<td>0.4029</td>
<td>0.9635</td>
<td>0.4599</td>
<td>0.9635</td>
<td>0.00343</td>
</tr>
<tr>
<td>Y2</td>
<td>0.67078</td>
<td>0.69345</td>
<td>0.42505</td>
<td>0.36974</td>
<td>0.0001</td>
</tr>
<tr>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
</tr>
<tr>
<td>δ</td>
<td>0.87643</td>
<td>-0.10321</td>
<td>-0.03564</td>
<td>0.6348</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.0001</td>
<td>0.1680</td>
<td>0.6348</td>
<td></td>
<td></td>
</tr>
<tr>
<td>θ</td>
<td>0.06730</td>
<td>-0.02135</td>
<td>0.02135</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.3694</td>
<td>0.02135</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>α</td>
<td>0.43217</td>
<td>0.0001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.0001</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: The first number in each row for each EEG measure is the correlation coefficient between row EEG and column EEG measures. The sample size for each entry is 180. The second number in each row is the significance level of the correlation under the null hypothesis Ho:R=0.
AID3 analyses were performed for all dependent EEG measures. All analyses specified a minimum group size of 5 and a default reducibility criterion of 0.8% for the parent group, i.e., 0.8% of the variability in the parent group had to be explained before a dichotomous split was made by the algorithm. Not all independent variables were represented in each AID3 analysis. Table 13 summarizes the independent variables that appeared in each AID3 analysis.

**AID3 Analysis of α.** The AID3 analysis for alpha (α) is shown in Figure 25. Approximately 36 percent of the variability in α is explained by the independent measures. The first split is on X7 (detection of signal). Group 2 contains all the EEG segments coded for detection and Group 3 contains all the observations coded for non-detection. The mean and variance of α for the detection group is lower than for the non-detection group. Alpha is generally associated with a relaxed state. A blocking of alpha occurs in an alerted state. The findings of lower mean and less variable α power for the detection group are consistent with the literature on alertness.

**Inferential Analysis of EEG Measures**

One problem that has constantly "dogged" the analysis is the lack of balance, of treatment effects. Treatment means are confounded with other effects-separation of effects is difficult or impossible. Because of the confounding problem special demands are place an analysis and interpretation.
Table 13. Summary of Predictor Variables That Appear in AID3 Analyses of EEG Measures.

<table>
<thead>
<tr>
<th>Predictor Variable</th>
<th>Dependent EEG Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Y2</td>
</tr>
<tr>
<td>Signal Presentation Time (X1)</td>
<td></td>
</tr>
<tr>
<td>Type of Prior Signal (X2)</td>
<td></td>
</tr>
<tr>
<td>Prior Signal Interval (X3)</td>
<td></td>
</tr>
<tr>
<td>Subject (X4)</td>
<td>✓</td>
</tr>
<tr>
<td>Order of Run (X5)</td>
<td>✓</td>
</tr>
<tr>
<td>EEG Segment (X6)</td>
<td>✓</td>
</tr>
<tr>
<td>Detection of Signal (X7)</td>
<td>✓</td>
</tr>
</tbody>
</table>
Figure 25. AID3 Analysis of Power in Alpha (8-13 cps) Frequency Range of EEG Segment.
If the design is severely unbalanced and "unconnected", data has to be analyzed piecemeal. As a result of the piecemeal analysis, statistical power suffers and may threaten the validity of results.

One-way analyses of variance were performed for EEG dependent measures. The results are reported in Table 14. A problem that exists in these results is confounding. For example, the EEG measure alpha shows significant effects in the one-way ANOVA for type of prior signal (X2), subject (X4), and detection of signal (X7) (detection of signal (X7) is not really an independent variable but more a concomitant indicator variable) but it is not clear whether all of these effects are dominant because of the lack of independence in effect means due to nonorthogonality. The significance of effects is not possible to determine without further analysis.

Table 14 indicates that strong subject effects are present in all measures. Run effects are present in Y2 and β. The type of prior signal (X2) appears to have an effect on α and β. X7 (detection of a signal) appears as a significance influence on Y2 and α. These effects will be discussed in detail below.

**Overall EEG Standard Deviation (Y2).** Y2 is the standard deviation of the raw EEG signal. It is a measure of the consistency of the amplitude of the EEG. Large values of Y2 could indicate the presence of low frequency and higher amplitude waves such as δ, θ, and α that are usually associated with low levels of alertness and drowsiness. Alpha is suppressed and β increased when the eyes are open and during cognitive tasks such as mental arithmetic. The results of the one-way
<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>Model</th>
<th>Y2</th>
<th>δ</th>
<th>θ</th>
<th>α</th>
<th>β</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time (X₁)*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>0.35</td>
<td>0.78</td>
<td>0.60</td>
<td>1.04</td>
<td>0.44</td>
</tr>
<tr>
<td></td>
<td>α</td>
<td>0.9466</td>
<td>0.6175</td>
<td>0.7788</td>
<td>0.4114</td>
<td>0.8939</td>
</tr>
<tr>
<td></td>
<td>R²</td>
<td>1.6%</td>
<td>3.5%</td>
<td>2.7%</td>
<td>4.6%</td>
<td>2.0%</td>
</tr>
<tr>
<td>Type of Prior Signal (X₂)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>2.86</td>
<td>2.89</td>
<td>1.20</td>
<td>3.41</td>
<td>3.08</td>
</tr>
<tr>
<td></td>
<td>α</td>
<td>0.0600</td>
<td>0.0581</td>
<td>0.3024</td>
<td>0.0354</td>
<td>0.0485</td>
</tr>
<tr>
<td></td>
<td>R²</td>
<td>3.1%</td>
<td>3.2%</td>
<td>1.3%</td>
<td>3.7%</td>
<td>3.4%</td>
</tr>
<tr>
<td>Prior Signal Interval (X₃)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>1.88</td>
<td>2.53</td>
<td>0.48</td>
<td>2.48</td>
<td>1.60</td>
</tr>
<tr>
<td></td>
<td>α</td>
<td>0.1564</td>
<td>0.0824</td>
<td>0.6165</td>
<td>0.0864</td>
<td>0.2044</td>
</tr>
<tr>
<td></td>
<td>R²</td>
<td>2.1%</td>
<td>2.8%</td>
<td>0.5%</td>
<td>2.7%</td>
<td>1.6%</td>
</tr>
<tr>
<td>Subject (X₄)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>44.16</td>
<td>7.24</td>
<td>12.65</td>
<td>4.64</td>
<td>27.97</td>
</tr>
<tr>
<td></td>
<td>α</td>
<td>0.0001</td>
<td>0.0002</td>
<td>0.0001</td>
<td>0.0039</td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td>R²</td>
<td>42.9%</td>
<td>11.0%</td>
<td>17.7%</td>
<td>7.3%</td>
<td>32.3%</td>
</tr>
<tr>
<td>Order of Run (X₅)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>2.41</td>
<td>1.21</td>
<td>1.55</td>
<td>2.07</td>
<td>3.88</td>
</tr>
<tr>
<td></td>
<td>α</td>
<td>0.0378</td>
<td>0.3072</td>
<td>0.1749</td>
<td>0.0704</td>
<td>0.0025</td>
</tr>
<tr>
<td></td>
<td>R²</td>
<td>6.5%</td>
<td>3.4%</td>
<td>4.3%</td>
<td>5.6%</td>
<td>10.0%</td>
</tr>
</tbody>
</table>
Table 14. (continued).

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>Model</th>
<th>Y2</th>
<th>$\delta$</th>
<th>$\theta$</th>
<th>$\alpha$</th>
<th>$\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Detection of Signal (X7)</td>
<td>F</td>
<td>7.42</td>
<td>3.61</td>
<td>1.59</td>
<td>36.99</td>
<td>2.73</td>
</tr>
<tr>
<td></td>
<td>$\alpha$</td>
<td>0.0071</td>
<td>0.0592</td>
<td>0.2087</td>
<td>0.0001</td>
<td>0.1005</td>
</tr>
<tr>
<td></td>
<td>$R^2$</td>
<td>4.0%</td>
<td>2.0%</td>
<td>0.9%</td>
<td>17.2%</td>
<td>1.5%</td>
</tr>
<tr>
<td>Treatment** (T)</td>
<td>F</td>
<td>1.20</td>
<td>1.73</td>
<td>1.11</td>
<td>2.01</td>
<td>0.95</td>
</tr>
<tr>
<td></td>
<td>$\alpha$</td>
<td>0.3109</td>
<td>0.1446</td>
<td>0.3537</td>
<td>0.0948</td>
<td>0.4362</td>
</tr>
<tr>
<td></td>
<td>$R^2$</td>
<td>2.7%</td>
<td>3.8%</td>
<td>2.5%</td>
<td>4.4%</td>
<td>2.1%</td>
</tr>
</tbody>
</table>

Note: The variables X1 and T above are recoded variables.
TIME (X1) = 1 (X1 = 315, X6 = 1), TIME = 2 (X1 = 315, X6 = 2), ... TIME = 9 (X1 = 955, X6 = 3).
TREATMENT (T) = 1 (X2 = 1, X3 = 1), T = 2 (X2 = 2, X3 = 2), T = 3 (X2 = 2, X3 = 3), T = 4 (X2 = 3, X3 = 2), T = 5 (X2 = 3, X3 = 3)
ANOVA suggest that Y2 is influenced by the detection of a signal. The mean of Y2 for detection is 56.7 (N = 168) whereas the mean of Y2 for non-detection is 67.2 (N = 12). Although the result is significant by an F-test, several observations should be made concerning the validity of this result. The twelve data points for non-detections represent four missed signals. Subject 3 missed three of the four signals and subject 4 missed one signal. The overall averages for Y2 for each subject were as follows: subject 1 (48.3, N = 36), subject 2 (46.6, N = 36), subject 3 (67.7, N = 54), and subject 4 (60.4, N = 54). The significant detection-non-detection effect in Y2 can be explained in terms of high average subject values present in Y2, i.e. the significant difference found in Y2 between detection and non-detection data was due in part to confounded subject differences.

**α-Wave Results.** Alpha data was further analyzed beyond the initial one-way ANOVA. The non-detection data was removed from the data so that further analysis only used detection data. An ANOVA was performed on the data that viewed the design as a subjects by treatments (S X A) repeated measures design. The observations were modeled as:

\[ X_{\text{ase}} = m + a_s + s_a + a_s + e_{(as)} \]

where the a-effect is considered a fixed effect and corresponds to type of prior signal and the S-effect is a random subject effect. When the data is analyzed in this way, the results of Table 15 are obtained. Subject effects dominate the results. Tukey's HSD test with an experimentwise error
Table 15. Subjects x Type of Prior Signal ANOVA for Alpha (α).

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>Type III and Type IV Sums of Squares</th>
<th>MS</th>
<th>F</th>
<th>PR&gt;5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of Prior Signal (X2)</td>
<td>2</td>
<td>1165.0</td>
<td>582.5</td>
<td>3.01</td>
<td>0.05</td>
</tr>
<tr>
<td>Subjects (X4)</td>
<td>3</td>
<td>3433.6</td>
<td>1144.5</td>
<td>5.92</td>
<td>0.0009</td>
</tr>
<tr>
<td>X2 * X4</td>
<td>6</td>
<td>1426.8</td>
<td>237.8</td>
<td>1.23</td>
<td>N.S</td>
</tr>
<tr>
<td>Error</td>
<td>156</td>
<td>30157.2</td>
<td>193.3</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Note: Model Values, F(11, 156) = 3.19, PR>F = 0.007, R^2 = 0.1834
Model: X_{ase} = M + A_a + S_s + A*S_{as} + E_{e(as)}, all effects fixed except E_{e(as)}. 

217
rate of 0.05 indicates that subjects 2 and 4 are significantly different in their mean EEG alpha levels. Type of prior signal (X2) produces a borderline effect with the PR > F = 0.05. Tukey's HSD test with an experimentwise error rate of 0.05 indicated that auditory warning signals resulted in higher average amplitude EEG alpha waves than visual artificial prior signals.

The comments concerning detection versus non-detection effects on Y2 also apply for alpha. Subjects 3 and 4 have higher amplitude alpha waves than subjects 1 and 2. Since subjects 3 and 4 are the only subjects to miss signals, it appears that non-detection of signals results in higher amplitude alpha waves; however, this effect is really confounded with subject effects.

**β-Wave Results.** The results obtained for β in the one-way ANOVA are confounded results due to nonorthogonality. Further analyses were performed to try to ascertain separate effects. The one-way ANOVA and Tukeys HSD multiple comparison test indicated that run 3 was significantly different from other runs. Run 3 effects were removed from subsequent analysis of variance tests. A two-way ANOVA based on subject x type of prior signal design assumptions was performed for β. The model assumed was:

\[ X_{ase} = m + a_s + s + e(\alpha, s) \]

The general linear models analysis resulted in a significant F-value (\( F = 8.39, V_1 = 11, V_2 = 132 \)) for the model with a significant level of PR > F = 0.0001. The model explained 41.1% of the variability in
However, the analysis indicated that nearly all of this variability was due to subject effects. No other effects were significant.

Eye Movement Analysis Results

Introduction. The display board which the subject monitored during the experiment was divided into three sections for calibration and analysis: section 1 (left meter and surrounding area), section 2 (middle section of the display), and section 3 (right meter and surrounding area). If the eyes traveled to the left meter (section 1), the time at which the eyes entered and fixated in section 1 as well as a display segment "1" code was recorded by the computer. If the eyes traveled to segment 2, the times at which the eyes made an excursion from display segment 1 and entered display segment 2 as well as correct display segment coding were sensed by the Biometrics Eye Monitor System and recorded by the computer. All eye movement recording was event based— whenever an event changed, a record was made by the computer of the event and the associated time of the change. A continuous record was made of all dwell times in each display segment as well as transition times between display segments. The entire eye movement analysis was based on these records.

Although the Biometrics Eye Monitor System was the primary eye movement instrumentation system, the activity of the right eye was picked up by a television camera with a telephoto lens mounted in the display board. The information from the right eye was recorded on video tape for the entire experimental run and used to corroborate data collected via the Biometrics Eye Monitor System.
system is more suitable for real time experimentation than is the video recording system because the data reduction task is simplified. However, the experimenter does not have any direct feedback of the Biometrics system during a run concerning how well the system holds calibration.

Eye Movement Analysis Strategy. The eye movement analysis strategy that produced the results reported in this section is shown in Figure 26. Eye movement data were extracted from the subject's master data file, preprocessed to get it in a usable form, and then used to calculate initial dwell time and activity measures.

A reliability analysis was performed on the data collected from the Biometrics Eye Monitoring System and the video recording system on a sampled basis. Five percent of the data collected from both systems for the same time period was used to determine system reliability.

A special analysis was performed for eye movement data collected on the video recording system. The purpose of this analysis was to examine eye movement behavior in relationship to signals that were missed on the display panel. These missed signals were obtained from the response time analysis.

Descriptive analyses were performed for the dwell times and transition times. Histograms, stem and leaf plots, box plots, normal probability plots, frequency tables, and descriptive statistics were determined for each subject for each experimental run.

The next eye movement analysis activity that was performed was that of filtering noise out of the data, developing a data aggregation
Figure 26. Eye Movement Analysis Strategy.
policy, and further developing performance measures. Examination of preprocessed data files revealed that considerable "noise" existed in the eye movement data. This noise was manifested at display segment boundaries where boundary flutter occurs. Boundary flutter consisted of transition times less that 0.10 seconds and rapid fluctuations back and forth across the display boundary, e.g., 1-2, 1-2 or 2-3, 2-3, etc. Sometimes dwell times were also registered in a display segment that were less than 0.10 seconds. The literature review suggested that the shortest time to define a fixation is 0.10 seconds (Lambert, Monty, Hall, 1974). Consequently, boundary flutter and dwell times of durations less than 0.10 seconds were filtered out of the data for subsequent eye movement analyses. However, these eye movement times of less than 0.10 seconds were recorded, tabulated, and used as a data quality check in the analysis.

A Fortran analysis program was written to aggregate data every thirty seconds and to calculate performance measures. Although any aggregate interval is somewhat arbitrary, the aggregation interval in this research was selected with two guidelines considered: (1) the aggregation interval should not be too narrow; otherwise little eye activity will occur in an interval and critical activity that might span several intervals would be missed and (2) the aggregation interval should not be so broad that any display-induced changes in eye movement behavior would be obscured because of averaging across time.

Variables Used in Eye Movement Analysis. The predictor variables used in the eye movement analysis are shown in Table 16. The basic
Table 16. Predictor Variables Used in Eye Movement Analysis.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Levels of Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time on the watch (X1)</td>
<td>Time on the watch ranging from 0-20 minutes (0-1200 seconds). Time was initially divided into 40, 30 second intervals for the eye movement analyses. For GLM analyses, time was further aggregated into five minutes (300 second) intervals.</td>
</tr>
</tbody>
</table>
| Signal Presence           | 1 = signal present in a 30 second interval  
|                           | 2 = signal not present during a 30 second interval.     |
| Subject (X4)              | 1 = subject 1                                           |
|                           | 2 = subject 2                                           |
|                           | 3 = subject 3                                           |
|                           | 4 = subject 4                                           |
| Order of Run (X5)         | 1 = baseline run 1 (B1)                                 |
|                           | 2 = baseline run 2 (B2)                                 |
|                           | 3 = run 3-visual artificial signals (V1)                 |
|                           | 4 = run 4-auditory warning signals (A2)                  |
|                           | 5 = run 5-Visual artificial signals (V2)                 |
|                           | 6 = run 6-auditory warning signals (A1)                  |
predictor variables used in the analysis were "time on the watch (X1)" and "type of prior signal (X2)."

Table 17 contains a list of three types of dependent variables used in the eye movement analysis: (1) raw measures, e.g., dwell times and transition times by display segment or type of transition, (2) derived performance measures (performance measures derived from raw measures but aggregated over 30-second intervals), and (3) transformed derived performance measures (percentage performance measures are transformed by an arc sine transformation).

Reliability Analysis Results. Eye movements of subjects were recorded by two different techniques during an experimental run. Eye activity was sensed by a Biometrics 5GHV/2 eye movement monitor. Also during an experimental run, a TV camera with a telephoto lens picked up the movement of the right eye.

Five percent (one minute from each run) of the data was sampled from both systems to determine the reliability. Data was examined from both systems from the beginning, middle, and end segment of an experimental run. The results of the reliability analysis are presented in Table 18. The reliability ranged from 79 percent to 100 percent with an average reliability across all subjects of 95 percent. Subject 3 exhibited the lowest percentage of correspondence (79 percent) between the two eye movement recording systems. The video taped showed this subject had her eyes closed during a large percentage of runs A1 and A2.
Table 17. Eye Movement Raw Measures and Performance Measures.

### Raw Measures

<table>
<thead>
<tr>
<th>Measure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DWELL</td>
<td>Dwell time spent in display segments 1, 2, or 3.</td>
</tr>
<tr>
<td>TRANS</td>
<td>Transition time. Time spent in travelling from one segment to another (1-2, 2-1, 1-3, 3-1, 2-3, 3-2)</td>
</tr>
</tbody>
</table>

### Derived Performance Measures (Aggregated over 30 second intervals)

<table>
<thead>
<tr>
<th>Measure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>E1</td>
<td># of events in display segment 1</td>
</tr>
<tr>
<td>E2</td>
<td># of events in display segment 3</td>
</tr>
<tr>
<td>E3</td>
<td># of events in display segments 1 and 3</td>
</tr>
<tr>
<td>E4</td>
<td>Total dwell time in display segment 1</td>
</tr>
<tr>
<td>E5</td>
<td>Total dwell time in display segment 3</td>
</tr>
<tr>
<td>E6</td>
<td>Total dwell time in display segments 1 and 3 (E6=E4+E5)</td>
</tr>
<tr>
<td>E7</td>
<td>Average dwell time in segment 1 (E7=E4/E1)</td>
</tr>
<tr>
<td>E8</td>
<td>Average dwell time in segment 3 (E8=E5/E1)</td>
</tr>
<tr>
<td>E9</td>
<td>Average dwell time in segments 1 and 3 (E9=E6/E1)</td>
</tr>
<tr>
<td>E10</td>
<td>Segment 1 dwell time variance</td>
</tr>
<tr>
<td>E11</td>
<td>Segment 3 dwell time variance</td>
</tr>
<tr>
<td>E12</td>
<td>Segments 1 and 3 dwell time variance</td>
</tr>
<tr>
<td>E13</td>
<td>% of total time spent in segment 1</td>
</tr>
<tr>
<td>E14</td>
<td>% of total time spent in 1-2, 2-1, 1-3, 3-1, 2-3, 3-2</td>
</tr>
<tr>
<td>E15</td>
<td>% of total dwell time spent in segment 3</td>
</tr>
<tr>
<td>E16</td>
<td>% of eye movement times &lt;0.1 seconds</td>
</tr>
<tr>
<td>E17</td>
<td>% of total time spent in segments 1 and 3.</td>
</tr>
</tbody>
</table>

### Transformed Performance Measures

<table>
<thead>
<tr>
<th>Measure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1</td>
<td>$\sqrt{\text{ARCSINE}(E13)}$</td>
</tr>
<tr>
<td>F2</td>
<td>$\sqrt{\text{ARCSINE}(E14)}$</td>
</tr>
<tr>
<td>F3</td>
<td>$\sqrt{\text{ARCSINE}(E15)}$</td>
</tr>
<tr>
<td>F4</td>
<td>$\sqrt{\text{ARCSINE}(E16)}$</td>
</tr>
<tr>
<td>F5</td>
<td>$\sqrt{\text{ARCSINE}(E17)}$</td>
</tr>
</tbody>
</table>
Table 18. Eye Movement Location Reliability Analysis—Television Recording System Versus Biometrics 5GHV/2 Eye Movement Monitor.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Run Type</th>
<th>Reliability (% Correspondence)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Baseline 1</td>
<td>88</td>
</tr>
<tr>
<td></td>
<td>Baseline 2</td>
<td>94</td>
</tr>
<tr>
<td></td>
<td>Visual (V1)</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>Visual (V2)</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>Auditory (A2)</td>
<td>88</td>
</tr>
<tr>
<td>2</td>
<td>Baseline 1</td>
<td>98</td>
</tr>
<tr>
<td></td>
<td>Baseline 2</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>Visual (V1)</td>
<td>95</td>
</tr>
<tr>
<td></td>
<td>Auditory (A1)</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>Visual (V2)</td>
<td>94</td>
</tr>
<tr>
<td></td>
<td>Auditory (A2)</td>
<td>100</td>
</tr>
<tr>
<td>3</td>
<td>Baseline 1</td>
<td>95</td>
</tr>
<tr>
<td></td>
<td>Baseline 2</td>
<td>96</td>
</tr>
<tr>
<td></td>
<td>Visual (V1)</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>Auditory (A1)</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>Visual (V2)</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>Auditory (A2)</td>
<td>79</td>
</tr>
<tr>
<td>4</td>
<td>Baseline 1</td>
<td>96</td>
</tr>
<tr>
<td></td>
<td>Baseline 2</td>
<td>95</td>
</tr>
<tr>
<td></td>
<td>Visual (V1)</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>Auditory (A1)</td>
<td>98</td>
</tr>
<tr>
<td></td>
<td>Visual (V2)</td>
<td>87</td>
</tr>
<tr>
<td></td>
<td>Auditory (A2)</td>
<td>84</td>
</tr>
</tbody>
</table>
Analysis of Missed Signals and Eye Movements. An analysis of the video tapes and responses to signals was performed to better understand missed signal behavior. Information about missed signals was obtained from the response time analysis. Subject 3 missed a visual signal at 955 seconds in run A1. This subject also missed a visual artificial signal during run V2 at 305 seconds and a visual signal at 955 seconds. In addition to these missed signals, subject 3 missed a visual signal during run A2 at 305 seconds. Subject 4 missed a visual signal during baseline run 2 at 315 seconds. These two subjects accounted for all missed signals during the experiment, a total of five missed signals. Figure 27 is a diagram that shows missed signals during a run as well as right eye behavior obtained from video tapes. Figure 27 shows a schedule of signals presented to a subject during a run in which signals were missed. Directly below each run signal schedule is a graph that shows both periods of long dwell times in a display segment (cross-hatched portions) and periods on the video tape when the right eye was closed for extended periods of time (solid shaded portions).

Signals presented for detection can be missed in three ways: (1) a person's eyes are closed during the time a signal is presented for detection, (2) a person is fixated too long in the wrong place, and (3) a person may be fixated on a meter on which the mean value has shifted (a signal) but not processing the information. This last category of explanation did not show up in the analysis of eye behavior from the video tapes. Figure 27 shows that when signals were
Figure 27. Eye Behavior Observed on Video Tape for Runs With Missed Signals.
missed, the concurrent eye behavior was that the subject was fixated on the wrong meter too long or that the eyes were closed during the appearance of a signal. A signal was presented for detection in this experiment for five seconds. If the eyes were closed or fixated in the wrong place for periods greater than five seconds during the presentation of a signal, the signal was missed. This shows clearly in Figure 27.

Prior to the eye being closed for extended periods (the dark shaded portion of the graph), subjects exhibited little or no scan activity. The eye would droop, be partially open, alternate between being partially opened and closed for brief durations of a few tenths of a second. The chin and head were fixed throughout the experimental runs. The observed behavior was similar to the visual behavior that precedes sleep and was consistent for the open spaces between dark shaded portions of the graphs for subject 3 during runs A1 and A2. The auditory warning signal did seem to alert subject 3 long enough to enable her to detect the visual signal presented for detection. No visual signal that was preceded by an auditory warning signal was missed.

A missed signal analysis was performed on the eye movement data recorded using the Biometrics Eye Monitor System. Five signals were missed across all runs. Two subjects accounted for the five missed signals and subject 3 accounted for four of these missed signals. Eye movement data was plotted for each missed signal. A total of 20 seconds of eye movement data was plotted around a missed signal (five
seconds before the onset of a signal, five seconds during the presentation of a signal and ten seconds after the signal is turned off). Figure 28 illustrates the plotting procedure for missed signals. The display segment (1, 2, or 3) in which the eyes were fixated versus a 20 second time interval around a missed signal was plotted. The zero point of the horizontal scale is the beginning of the five second interval before the onset of a signal. Figure 28 shows that subject 3 during run V2 had her eyes fixated on display segment 1 or 2 during the time that a signal arrived on the meter in display segment 3. This pattern is similar for subject 3, run V2 (artificial signal missed at 305 seconds) and run A1 (visual signal missed at 955 seconds). The eye movement patterns for subject 3, run A2 and subject 4, run B2 do not exhibit this pattern of being fixated on the wrong meter. The eye movement trace for subject 3 during run A2 indicated the subject was fixated on the correct meter but still missed the signal. The previous analysis of eye movements obtained from video tapes showed that subject 3's eyes were closed during the on-time of this signal (see Figure 27). When a subject closed his/her eyes, the Biometrics Eye Monitoring System registered that period of time as a dwell time in whatever segment the subject last had his/her eyes open. Subject 4, during run B2 (second baseline run) missed a signal at 315 seconds. His pattern is quite different from the other patterns. Subject 4's pattern is characterized by a great deal of darting back and forth across the display board. Forty-seven events were recorded during the on-time of the signal. Thirty-seven of these 47 events are
SIGNAL ON-TIME
METER IN DISPLAY 3 CHANGES — EYES ARE FIXED IN
DISPLAY SEGMENTS 1 OR 2. VISUAL ARTIFICIAL SIGNAL
IS MISSED.

Figure 28. Display Segment (1, 2, or 3) Activity Versus the
20 Second Interval Around a Missed Signal for
Subject 3-Run V2 (Horizontal Scale Zero Point
Occurs 5 Seconds Before the Onset of a Signal).

Note: A designation of "1" on the vertical axis indicates the
eyes are fixated at the extreme left of the display panel.
A "2" indicates fixation in the center of the panel and
a "3" indicates fixation at the extreme right of the panel.
of a duration less than 0.10 seconds. The literature review suggested that 0.10 seconds is the shortest time for which an eye fixation can be defined. If this is the shortest period for which information can be meaningfully processed then it is not surprising that subject 4 missed the signal. He only occasionally fixated long enough to process any meter movements. The frequency of the needle swing for each meter was 1-2 cycles per second. It takes a one-half cycle to determine whether a mean shift has occurred. Therefore, it takes approximately 0.5-1.0 seconds of meter fixation before the monitor can determine whether a shift in the mean value of the meter has occurred. It is possible that subject 4 was scanning the display too rapidly to obtain the information necessary to make a decision. The longest dwell time Subject 4 had on the left meter (the meter that changed) in display segment 1 was 0.18 seconds.

**Descriptive Analysis of Data.** The display was divided into a left, middle, and right section (designated 1, 2, and 3 respectively) for calibration and analysis purposes. Univariate statistics were calculated for eye dwell times for each display segment and for each boundary where an eye transition was made from one display segment to another. This type of analysis helped to identify ranges of dwell times in each display segment and transition times at display boundaries. Table 19 summarizes the means, variances, and number of events in a display segment or at display segment boundaries for each subject and each experimental run. Display sections are designated
Table 19. Means, Variances, and Number of Observations for Eye Movement Times in Each Eye Movement Dwell or Transition Region of the Display Panel.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Run</th>
<th>Parameter</th>
<th>1-1</th>
<th>1-2</th>
<th>2-1</th>
<th>2-2</th>
<th>2-3</th>
<th>3-2</th>
<th>3-3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>B1</td>
<td>Mean</td>
<td>1.65</td>
<td>0.28</td>
<td>0.07</td>
<td>0.77</td>
<td>0.06</td>
<td>0.28</td>
<td>1.43</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Variance</td>
<td>1.29</td>
<td>0.04</td>
<td>0.01</td>
<td>2.33</td>
<td>0.01</td>
<td>0.04</td>
<td>0.68</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N</td>
<td>284</td>
<td>288</td>
<td>289</td>
<td>148</td>
<td>307</td>
<td>307</td>
<td>289</td>
</tr>
<tr>
<td>1</td>
<td>B2</td>
<td>Mean</td>
<td>2.32</td>
<td>0.25</td>
<td>0.07</td>
<td>0.79</td>
<td>0.07</td>
<td>0.15</td>
<td>2.03</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Variance</td>
<td>2.34</td>
<td>0.04</td>
<td>0.01</td>
<td>1.69</td>
<td>0.01</td>
<td>0.03</td>
<td>1.28</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N</td>
<td>229</td>
<td>241</td>
<td>242</td>
<td>97</td>
<td>293</td>
<td>294</td>
<td>223</td>
</tr>
<tr>
<td>2</td>
<td>B1</td>
<td>Mean</td>
<td>1.76</td>
<td>0.39</td>
<td>0.06</td>
<td>0.06</td>
<td>0.06</td>
<td>0.30</td>
<td>1.53</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Variance</td>
<td>1.12</td>
<td>0.09</td>
<td>0.01</td>
<td>0.02</td>
<td>0.00</td>
<td>0.08</td>
<td>0.86</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N</td>
<td>288</td>
<td>305</td>
<td>305</td>
<td>128</td>
<td>303</td>
<td>303</td>
<td>288</td>
</tr>
<tr>
<td>2</td>
<td>B2</td>
<td>Mean</td>
<td>1.96</td>
<td>0.40</td>
<td>0.06</td>
<td>0.09</td>
<td>0.07</td>
<td>0.22</td>
<td>1.89</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Variance</td>
<td>0.94</td>
<td>0.08</td>
<td>0.00</td>
<td>0.07</td>
<td>0.00</td>
<td>0.08</td>
<td>1.09</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N</td>
<td>254</td>
<td>257</td>
<td>257</td>
<td>120</td>
<td>310</td>
<td>310</td>
<td>257</td>
</tr>
<tr>
<td>3</td>
<td>B1</td>
<td>Mean</td>
<td>3.02</td>
<td>0.24</td>
<td>0.03</td>
<td>0.06</td>
<td>0.04</td>
<td>0.12</td>
<td>2.13</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Variance</td>
<td>2.82</td>
<td>0.04</td>
<td>0.00</td>
<td>0.04</td>
<td>0.00</td>
<td>0.03</td>
<td>1.44</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N</td>
<td>205</td>
<td>204</td>
<td>205</td>
<td>89</td>
<td>349</td>
<td>350</td>
<td>218</td>
</tr>
<tr>
<td>3</td>
<td>B2</td>
<td>Mean</td>
<td>1.93</td>
<td>0.25</td>
<td>0.03</td>
<td>0.12</td>
<td>0.04</td>
<td>0.16</td>
<td>1.61</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Variance</td>
<td>1.29</td>
<td>0.06</td>
<td>0.00</td>
<td>0.14</td>
<td>0.00</td>
<td>0.04</td>
<td>0.97</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N</td>
<td>286</td>
<td>339</td>
<td>339</td>
<td>68</td>
<td>340</td>
<td>339</td>
<td>297</td>
</tr>
<tr>
<td>Subject</td>
<td>Run</td>
<td>Parameter</td>
<td>1-1</td>
<td>1-2</td>
<td>2-1</td>
<td>2-2</td>
<td>2-3</td>
<td>3-2</td>
<td>3-3</td>
</tr>
<tr>
<td>---------</td>
<td>-----</td>
<td>-----------</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>4</td>
<td>B1</td>
<td>Mean</td>
<td>0.86</td>
<td>0.23</td>
<td>0.06</td>
<td>0.15</td>
<td>0.07</td>
<td>0.20</td>
<td>0.91</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Variance</td>
<td>0.32</td>
<td>0.04</td>
<td>0.00</td>
<td>0.12</td>
<td>0.00</td>
<td>0.05</td>
<td>0.45</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N</td>
<td>492</td>
<td>575</td>
<td>575</td>
<td>178</td>
<td>553</td>
<td>553</td>
<td>478</td>
</tr>
<tr>
<td>4</td>
<td>B2</td>
<td>Mean</td>
<td>0.94</td>
<td>0.20</td>
<td>0.06</td>
<td>0.32</td>
<td>0.08</td>
<td>0.21</td>
<td>1.14</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Variance</td>
<td>0.53</td>
<td>0.04</td>
<td>0.01</td>
<td>0.43</td>
<td>0.01</td>
<td>0.05</td>
<td>0.83</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N</td>
<td>386</td>
<td>577</td>
<td>575</td>
<td>207</td>
<td>550</td>
<td>548</td>
<td>412</td>
</tr>
<tr>
<td>1</td>
<td>V1</td>
<td>Mean</td>
<td>1.56</td>
<td>0.35</td>
<td>0.07</td>
<td>0.18</td>
<td>0.05</td>
<td>0.44</td>
<td>1.41</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Variance</td>
<td>0.89</td>
<td>0.08</td>
<td>0.01</td>
<td>0.51</td>
<td>0.00</td>
<td>0.09</td>
<td>0.50</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N</td>
<td>298</td>
<td>323</td>
<td>323</td>
<td>168</td>
<td>302</td>
<td>301</td>
<td>300</td>
</tr>
<tr>
<td>1</td>
<td>V2</td>
<td>Mean</td>
<td>2.87</td>
<td>0.21</td>
<td>0.05</td>
<td>0.07</td>
<td>0.05</td>
<td>0.16</td>
<td>2.11</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Variance</td>
<td>1.70</td>
<td>0.03</td>
<td>0.00</td>
<td>0.03</td>
<td>0.00</td>
<td>0.03</td>
<td>1.28</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N</td>
<td>218</td>
<td>218</td>
<td>218</td>
<td>123</td>
<td>221</td>
<td>220</td>
<td>219</td>
</tr>
<tr>
<td>1</td>
<td>A1</td>
<td>Mean</td>
<td>2.76</td>
<td>0.24</td>
<td>0.05</td>
<td>0.43</td>
<td>0.06</td>
<td>0.17</td>
<td>1.92</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Variance</td>
<td>2.29</td>
<td>0.04</td>
<td>0.00</td>
<td>2.39</td>
<td>0.00</td>
<td>0.03</td>
<td>0.88</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N</td>
<td>221</td>
<td>221</td>
<td>222</td>
<td>143</td>
<td>223</td>
<td>223</td>
<td>214</td>
</tr>
<tr>
<td>2</td>
<td>V1</td>
<td>Mean</td>
<td>1.89</td>
<td>0.34</td>
<td>0.03</td>
<td>0.05</td>
<td>0.06</td>
<td>0.23</td>
<td>1.48</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Variance</td>
<td>0.62</td>
<td>0.05</td>
<td>0.00</td>
<td>0.01</td>
<td>0.00</td>
<td>0.07</td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N</td>
<td>284</td>
<td>292</td>
<td>293</td>
<td>160</td>
<td>450</td>
<td>450</td>
<td>284</td>
</tr>
<tr>
<td>Subject</td>
<td>Run</td>
<td>Parameter</td>
<td>1-1</td>
<td>1-2</td>
<td>2-1</td>
<td>2-2</td>
<td>2-3</td>
<td>3-2</td>
<td>3-3</td>
</tr>
<tr>
<td>---------</td>
<td>-----</td>
<td>-----------</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>2</td>
<td>A2</td>
<td>Mean</td>
<td>1.40</td>
<td>0.33</td>
<td>0.04</td>
<td>0.03</td>
<td>0.05</td>
<td>0.21</td>
<td>1.53</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Variance</td>
<td>0.39</td>
<td>0.06</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.06</td>
<td>0.45</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N</td>
<td>331</td>
<td>350</td>
<td>351</td>
<td>148</td>
<td>369</td>
<td>369</td>
<td>334</td>
</tr>
<tr>
<td>2</td>
<td>V2</td>
<td>Mean</td>
<td>1.39</td>
<td>0.35</td>
<td>0.03</td>
<td>0.02</td>
<td>0.05</td>
<td>0.26</td>
<td>1.55</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Variance</td>
<td>0.54</td>
<td>0.07</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.08</td>
<td>0.65</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N</td>
<td>327</td>
<td>337</td>
<td>338</td>
<td>143</td>
<td>337</td>
<td>337</td>
<td>327</td>
</tr>
<tr>
<td>2</td>
<td>A1</td>
<td>Mean</td>
<td>1.70</td>
<td>0.40</td>
<td>0.04</td>
<td>0.04</td>
<td>0.06</td>
<td>0.25</td>
<td>1.69</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Variance</td>
<td>0.72</td>
<td>0.09</td>
<td>0.00</td>
<td>0.01</td>
<td>0.00</td>
<td>0.09</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N</td>
<td>279</td>
<td>295</td>
<td>295</td>
<td>155</td>
<td>383</td>
<td>382</td>
<td>279</td>
</tr>
<tr>
<td>3</td>
<td>V1</td>
<td>Mean</td>
<td>1.22</td>
<td>0.16</td>
<td>0.03</td>
<td>0.14</td>
<td>0.05</td>
<td>0.17</td>
<td>1.39</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Variance</td>
<td>0.92</td>
<td>0.05</td>
<td>0.00</td>
<td>0.10</td>
<td>0.00</td>
<td>0.05</td>
<td>0.70</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N</td>
<td>373</td>
<td>779</td>
<td>778</td>
<td>96</td>
<td>376</td>
<td>375</td>
<td>362</td>
</tr>
<tr>
<td>3</td>
<td>A2</td>
<td>Mean</td>
<td>5.38</td>
<td>0.09</td>
<td>0.06</td>
<td>0.57</td>
<td>0.06</td>
<td>0.21</td>
<td>0.91</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Variance</td>
<td>230.26</td>
<td>0.03</td>
<td>0.01</td>
<td>0.81</td>
<td>0.00</td>
<td>0.06</td>
<td>0.76</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N</td>
<td>165</td>
<td>681</td>
<td>681</td>
<td>129</td>
<td>141</td>
<td>141</td>
<td>107</td>
</tr>
<tr>
<td>3</td>
<td>V2</td>
<td>Mean</td>
<td>1.95</td>
<td>0.21</td>
<td>0.11</td>
<td>0.86</td>
<td>0.06</td>
<td>0.18</td>
<td>0.81</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Variance</td>
<td>2.17</td>
<td>0.05</td>
<td>0.03</td>
<td>1.49</td>
<td>0.01</td>
<td>0.05</td>
<td>0.58</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N</td>
<td>305</td>
<td>414</td>
<td>415</td>
<td>286</td>
<td>290</td>
<td>290</td>
<td>191</td>
</tr>
</tbody>
</table>
### Table 19. (continued).

<table>
<thead>
<tr>
<th>Subject</th>
<th>Run</th>
<th>Parameter</th>
<th>1-1</th>
<th>1-2</th>
<th>2-1</th>
<th>2-2</th>
<th>2-3</th>
<th>3-2</th>
<th>3-3</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>A1</td>
<td>Mean</td>
<td>2.58</td>
<td>0.12</td>
<td>0.09</td>
<td>0.86</td>
<td>0.07</td>
<td>0.12</td>
<td>1.24</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Variance</td>
<td>16.08</td>
<td>0.04</td>
<td>0.02</td>
<td>1.50</td>
<td>0.01</td>
<td>0.03</td>
<td>4.28</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N</td>
<td>236</td>
<td>649</td>
<td>650</td>
<td>240</td>
<td>355</td>
<td>356</td>
<td>143</td>
</tr>
<tr>
<td>4</td>
<td>V1</td>
<td>Mean</td>
<td>1.08</td>
<td>0.27</td>
<td>0.04</td>
<td>0.09</td>
<td>0.04</td>
<td>0.19</td>
<td>1.04</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Variance</td>
<td>0.32</td>
<td>0.06</td>
<td>0.00</td>
<td>0.06</td>
<td>0.00</td>
<td>0.06</td>
<td>0.34</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N</td>
<td>416</td>
<td>492</td>
<td>492</td>
<td>133</td>
<td>555</td>
<td>554</td>
<td>437</td>
</tr>
<tr>
<td>4</td>
<td>A2</td>
<td>Mean</td>
<td>1.10</td>
<td>0.22</td>
<td>0.05</td>
<td>0.16</td>
<td>0.07</td>
<td>0.26</td>
<td>1.05</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Variance</td>
<td>0.37</td>
<td>0.06</td>
<td>0.00</td>
<td>0.08</td>
<td>0.01</td>
<td>0.06</td>
<td>0.65</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N</td>
<td>396</td>
<td>563</td>
<td>564</td>
<td>220</td>
<td>458</td>
<td>458</td>
<td>403</td>
</tr>
<tr>
<td>4**</td>
<td>V2</td>
<td>Mean</td>
<td>0.74</td>
<td>0.16</td>
<td>0.05</td>
<td>0.22</td>
<td>0.06</td>
<td>0.16</td>
<td>0.80</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Variance</td>
<td>0.38</td>
<td>0.03</td>
<td>0.01</td>
<td>0.14</td>
<td>0.01</td>
<td>0.04</td>
<td>0.35</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N</td>
<td>421</td>
<td>833</td>
<td>833</td>
<td>344</td>
<td>937</td>
<td>936</td>
<td>536</td>
</tr>
<tr>
<td>Subject</td>
<td>Run</td>
<td>Parameter</td>
<td>1-1</td>
<td>1-2</td>
<td>2-1</td>
<td>2-2</td>
<td>2-3</td>
<td>3-2</td>
<td>3-3</td>
</tr>
<tr>
<td>---------</td>
<td>------</td>
<td>-----------</td>
<td>-----</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>-----</td>
</tr>
<tr>
<td>4</td>
<td>A1</td>
<td>Mean</td>
<td>1.09</td>
<td>0.20</td>
<td>0.06</td>
<td>0.18</td>
<td>0.09</td>
<td>0.25</td>
<td>0.78</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Variance</td>
<td>0.69</td>
<td>0.05</td>
<td>0.01</td>
<td>0.08</td>
<td>0.01</td>
<td>0.06</td>
<td>0.65</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N</td>
<td>412</td>
<td>723</td>
<td>723</td>
<td>288</td>
<td>560</td>
<td>560</td>
<td>414</td>
</tr>
<tr>
<td>Display Section Mean time $\bar{X}$</td>
<td>1.88</td>
<td>.26</td>
<td>.05</td>
<td>.27</td>
<td>.06</td>
<td>.21</td>
<td>1.41</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standard Error of the Mean $\frac{S}{\bar{X}}$</td>
<td>1.00</td>
<td>.09</td>
<td>.02</td>
<td>.29</td>
<td>.01</td>
<td>.07</td>
<td>.43</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*The designation 1-1 means the subject had a dwell time in region 1 of the display panel. A designation of 1-2 means the subject made a transition from display region 1 to display region 2, etc.

**This subject during Run V2 made one transition from Display Segment 3 to Display Segment 1 that lasted 0.23 seconds. This was the only such occurrence of a transition not passing through Display Segment 2.
1-1, 1-2, 2-1, 2-2, 2-3, 3-2, and 3-3. For example, a 1-1 designation means the eyes entered segment 1 and fixated there for some amount of time. A 1-2 indicates the eyes required some amount of time to leave segment 1 and travel to segment 2. Designations for transitions 1-3 and 3-1 are absent from Table 19 because only one such transition occurred in the entire experiment. Mean dwell times and transition times are calculated across all subjects for each display section. These are shown at the bottom of Table 19. The dwell times for display segments 1 and 3 are not statistically different at the $\alpha \leq 0.05$ level. The average transition times for boundaries 1-2 and 3-2 are considerably larger than for 2-1 and 2-3. This finding can be explained in the following way: the eyes fixate and dwell in segment 1 and then a command is given to search and scan across other parts of the display. When the eyes get to boundary 1-2 the motion is relatively slow. However, by the time the eyes get to boundary 2-3 they have developed a maximum ballistic trajectory in order to arrive in display segment 3 where the relevant meter information is located. When the eyes begin in display section 3, a similar process occurs.

Subject 3 during run A2 showed a large average dwell time and variance in display segment 1. The video tape analysis showed the eyes to be closed a large proportion of the time for this subject and run. When the eyes travel to a location and then are closed, the Biometrics Eye Monitoring System records this as a dwell time if the eyes are reopened in this same display section.
Each distribution of dwell times and transition times was tested for normality using a modified Kolmogorov-Smirnov D-statistic. The null hypothesis is rejected for large values of D. PROC UNIVARIATE also reports the probability of a larger D, i.e., significance levels. The empirical distributions when tested against the normal distribution resulted in PROB ≤ .01 for 94 percent of the distributions. The remaining distributions were normal at a level of significance of ten percent. The distributions of dwell times and transition times were be regarded as normal in this experiment.

Correlation Analysis of Eye Movement Performance Measures

A correlation analysis of the performance measures was done in an effort to reduce the number of variables carried during the inferential analysis. Procedure CORR from SAS was used with the option RANK to determine all pairwise correlations of performance measures. The RANK option orders performance measures from the highest and most significant correlations to the lowest and least significant correlations. The procedure tests each correlation coefficient (RHO) against the hypothesis that RHO = 0.

The purpose of using correlation analysis to eliminate variables that were highly correlated was not particularly effective since nearly all performance measures were highly correlated with each other at a high level of significance (alpha ≤ .0001). This result comes about because of the way measures were constructed, e.g., E1 equals the number of eye movement events that take place in display segment one during a thirty second period. The measure E3 represents the
number of visual events (fixations) that take place in both display segments one and three during a thirty second interval. Performance measure E1 would be expected to be highly correlated with performance measure E3; indeed \( \rho = 0.90197 \) and \( \alpha < 0.0001 \). Another reason for most measures being significantly correlated with each other is the large sample size \( (N = 920) \) for each performance measure. The selection of performance measures to be carried in the analysis becomes a matter of judgment. The decision was made to carry all dependent eye movement measures to the inferential analysis.

**Correlation of the Mean and Variance of Eye Movement Measures with Time.** Eye movement measures were calculated for each subject across thirty second intervals. These performance measures were then aggregated across subjects and means and variances calculated for each eye movement measure. The means and variances for each measure were then correlated with time intervals. (The range of time is 1-40, 30 second intervals.) Table 20 shows the results obtained when the mean and variance of each eye movement performance measure was correlated with a time variable. Only significant relationships are reported.

**Inferential Statistical Tests for Eye Movement Measures.** General linear model (GLM) procedures were used to analyze the effects of time on the watch and signals on eye movement measures. Means and variances for each eye movement measure were calculated and plotted as a function of time. Runs A1 and A2 for Subject 3 were excluded from the analysis because of earlier results which showed this subject to have her eyes closed during large portions of these runs.
Table 20. Significant Correlations of Means and Variances of Eye Movement Performance Measures Aggregated Across Subjects Versus Time*.

<table>
<thead>
<tr>
<th>Eye Movement Measure Mean**</th>
<th>Correlation Coefficient</th>
<th>Significance Level</th>
<th>Eye Movement Measure Variance**</th>
<th>Correlation Coefficient</th>
<th>Significance Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>ME4</td>
<td>0.48125</td>
<td>0.0017</td>
<td>VE1</td>
<td>-0.45850</td>
<td>0.0029</td>
</tr>
<tr>
<td>ME6</td>
<td>0.32662</td>
<td>0.0397</td>
<td>VE2</td>
<td>-0.69336</td>
<td>0.0001</td>
</tr>
<tr>
<td>ME7</td>
<td>0.30892</td>
<td>0.0524</td>
<td>VE3</td>
<td>-0.64165</td>
<td>0.0001</td>
</tr>
<tr>
<td>ME13</td>
<td>0.48127</td>
<td>0.0017</td>
<td>VE5</td>
<td>-0.54050</td>
<td>0.0003</td>
</tr>
<tr>
<td>ME14</td>
<td>-0.43969</td>
<td>0.0045</td>
<td>VE6</td>
<td>-0.44869</td>
<td>0.0037</td>
</tr>
<tr>
<td>ME16</td>
<td>0.49870</td>
<td>0.0011</td>
<td>VE8</td>
<td>-0.59047</td>
<td>0.0001</td>
</tr>
<tr>
<td>ME17</td>
<td>0.32666</td>
<td>0.0397</td>
<td>VE9</td>
<td>-0.35547</td>
<td>0.0244</td>
</tr>
<tr>
<td>MF1</td>
<td>0.4112</td>
<td>0.0013</td>
<td>VE11</td>
<td>-0.38120</td>
<td>0.0152</td>
</tr>
<tr>
<td>MF2</td>
<td>-0.36501</td>
<td>0.0206</td>
<td>VE14</td>
<td>-0.49346</td>
<td>0.0012</td>
</tr>
<tr>
<td>MF4</td>
<td>0.54135</td>
<td>0.0003</td>
<td>VE15</td>
<td>-0.54045</td>
<td>0.0003</td>
</tr>
<tr>
<td>MF5</td>
<td>0.31100</td>
<td>0.0508</td>
<td>VE16</td>
<td>0.28643</td>
<td>0.0732</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>VE17</td>
<td>-0.44867</td>
<td>0.0037</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>VF2</td>
<td>-0.46334</td>
<td>0.0026</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>VF3</td>
<td>-0.61296</td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>VF5</td>
<td>-0.42584</td>
<td>0.0062</td>
</tr>
</tbody>
</table>

*Only correlations with significance levels < 10% are reported.

**The designations "M" and "V" in the performance measure indicate the mean and variance of the eye movement measure. A 1,200 second run is divided into 40, 30 second time intervals. The means and variances of each 30 second time interval across subjects is correlated with the time interval (N = 920).
Initially, a GLM main effects model of the form $E = \text{Eye movement measure} = \text{Time} + \text{Subject} + \text{Run}$ was used to analyze all the eye movement measures. Thirty-second time intervals were aggregated into four, five-minute intervals for the analysis. A summary of the signals and the thirty-second intervals used in this and subsequent analyses are shown in Table 21. Subject and Run effects were included in the model to reduce experimental error. If only Time was included in the model, Subject and Run effects would become a part of experimental error. Resultant F-tests might not show significance because of the increased noise. A summary of significant results is presented in Table 22.

Total dwell time in display segment 1 ($E_4$), total dwell time in display segment 3 ($E_5$), and percent of total time spent at display boundaries and in transition ($E_{14}$) showed significant time, subject, and run effects. The transformed variable $F_2 = \sqrt{\text{ARCSINE}} (E_{14})$ showed essentially the same result as the untransformed variable and is not reported here.

Dependent measures $E_4$, $E_5$, and $E_{14}$ were aggregated across five-minute segments of all runs except A1 and A2 for subject 3. The means and variances of these measures versus time are plotted and shown in Figures 29, 30, and 31. Figure 29 is a plot of mean eye dwell time in display segment 1 ($E_4$) versus each five-minute segment of the experimental runs. A significant increase occurs in the mean of $E_4$ in the range of 900-1200 seconds (15-20 minutes) of the runs. Figure 29 also shows that mean eye dwell time in display segment 3 ($E_5$) decreases in the range of 900-1200 seconds; thus $E_4$ and $E_5$ show opposite effects about the same time. This is not surprising since $E_4$ and $E_5$ have a
Table 21. Summary of Aggregation Intervals and Signals in Intervals for All Subjects.

<table>
<thead>
<tr>
<th>30 Second and 5 Minute Aggregation Interval</th>
<th>Range of Time (Seconds)</th>
<th>Experimental Run</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0 - 30</td>
<td>0 0 0 0 0 0</td>
</tr>
<tr>
<td>2</td>
<td>30 - 60</td>
<td>0 0 0 0 0 0</td>
</tr>
<tr>
<td>3</td>
<td>60 - 90</td>
<td>0 0 0 0 0 0</td>
</tr>
<tr>
<td>4</td>
<td>90 - 120</td>
<td>0 0 0 0 0 0</td>
</tr>
<tr>
<td>5</td>
<td>120 - 150</td>
<td>0 0 0 0 0 0</td>
</tr>
<tr>
<td>6</td>
<td>150 - 180</td>
<td>0 0 0 0 0 0</td>
</tr>
<tr>
<td>7</td>
<td>180 - 210</td>
<td>0 0 0 0 0 0</td>
</tr>
<tr>
<td>8</td>
<td>210 - 240</td>
<td>0 0 0 0 0 0</td>
</tr>
<tr>
<td>9</td>
<td>240 - 270</td>
<td>0 0 0 0 0 0</td>
</tr>
<tr>
<td>10</td>
<td>270 - 300</td>
<td>0 0 0 0 0 0</td>
</tr>
<tr>
<td>11</td>
<td>300 - 330</td>
<td>V V A-V V-V V</td>
</tr>
<tr>
<td>12</td>
<td>330 - 360</td>
<td>0 0 0 0 0 0</td>
</tr>
<tr>
<td>13</td>
<td>360 - 390</td>
<td>0 0 0 0 0 0</td>
</tr>
<tr>
<td>14</td>
<td>390 - 420</td>
<td>0 0 0 0 0 0</td>
</tr>
<tr>
<td>15</td>
<td>420 - 450</td>
<td>0 0 0 0 0 0</td>
</tr>
<tr>
<td>16</td>
<td>450 - 480</td>
<td>0 0 0 0 0 0</td>
</tr>
<tr>
<td>17</td>
<td>480 - 510</td>
<td>0 0 0 0 0 0</td>
</tr>
<tr>
<td>18</td>
<td>510 - 540</td>
<td>0 0 0 0 0 0</td>
</tr>
<tr>
<td>19</td>
<td>540 - 570</td>
<td>0 0 0 0 0 0</td>
</tr>
<tr>
<td>20</td>
<td>570 - 600</td>
<td>0 0 A V 0</td>
</tr>
<tr>
<td>21</td>
<td>600 - 630</td>
<td>0 V 0 0 A</td>
</tr>
<tr>
<td>22</td>
<td>630 - 660</td>
<td>V V V V V V</td>
</tr>
<tr>
<td>23</td>
<td>660 - 690</td>
<td>0 0 0 0 0 0</td>
</tr>
<tr>
<td>24</td>
<td>690 - 720</td>
<td>0 0 0 0 0 0</td>
</tr>
<tr>
<td>25</td>
<td>720 - 750</td>
<td>0 0 0 0 0 0</td>
</tr>
<tr>
<td>26</td>
<td>750 - 780</td>
<td>0 0 0 0 0 0</td>
</tr>
<tr>
<td>27</td>
<td>780 - 810</td>
<td>0 0 0 0 0 0</td>
</tr>
<tr>
<td>28</td>
<td>810 - 840</td>
<td>0 0 0 0 0 0</td>
</tr>
<tr>
<td>29</td>
<td>840 - 870</td>
<td>0 0 0 0 0 0</td>
</tr>
<tr>
<td>30</td>
<td>870 - 900</td>
<td>0 V 0 0 A</td>
</tr>
<tr>
<td>31</td>
<td>900 - 930</td>
<td>0 0 0 0 0 0</td>
</tr>
<tr>
<td>32</td>
<td>930 - 960</td>
<td>V V V V V V</td>
</tr>
<tr>
<td>33</td>
<td>960 - 990</td>
<td>0 0 0 0 0 0</td>
</tr>
</tbody>
</table>
Table 21. (continued).

<table>
<thead>
<tr>
<th>30 Second and 5 Minute Aggregation Interval</th>
<th>Range of Time (Seconds)</th>
<th>Experimental Run</th>
</tr>
</thead>
<tbody>
<tr>
<td>34</td>
<td>990 - 1020</td>
<td>B1 B2 V1 A1 V2 A2</td>
</tr>
<tr>
<td>35</td>
<td>1020 - 1050</td>
<td>0 0 0 0 0 0</td>
</tr>
<tr>
<td>36</td>
<td>1050 - 1080</td>
<td>0 0 0 0 0 0</td>
</tr>
<tr>
<td>37</td>
<td>1080 - 1110</td>
<td>0 0 0 0 0 0</td>
</tr>
<tr>
<td>38</td>
<td>1110 - 1140</td>
<td>0 0 0 0 0 0</td>
</tr>
<tr>
<td>39</td>
<td>1140 - 1170</td>
<td>0 0 0 0 0 0</td>
</tr>
<tr>
<td>40</td>
<td>1170 - 1200</td>
<td>0 0 0 0 0 0</td>
</tr>
</tbody>
</table>

0 = No signal is present in the 30-second interval.

A = Auditory warning signal.

V = Visual signal.

A-V = An audio alerting signal followed by a visual signal in the interval.

V-V = A visual signal followed by another visual signal in the interval.
Table 22. Analysis of Variance Summary Table for Eye Movement Measures.

<table>
<thead>
<tr>
<th>Eye Movement Measure</th>
<th>F-Value</th>
<th>Level of Significance for Model</th>
<th>$R^2$</th>
<th>Time*</th>
<th>Subject</th>
<th>Run</th>
</tr>
</thead>
<tbody>
<tr>
<td>E4 (Total dwell time in display segment 1)</td>
<td>20.26</td>
<td>0.0001</td>
<td>0.2121</td>
<td>0.0015</td>
<td>0.0001</td>
<td>0.0067</td>
</tr>
<tr>
<td>E5 (Total dwell time in display segment 3)</td>
<td>12.60</td>
<td>0.0001</td>
<td>0.1434</td>
<td>0.0207</td>
<td>0.0001</td>
<td>0.0001</td>
</tr>
<tr>
<td>E14 (% of total time spent in 1-2, 2-1, 1-3, 3-1, 2-3, 3-2)</td>
<td>17.72</td>
<td>0.0001</td>
<td>0.1906</td>
<td>0.0292</td>
<td>0.0001</td>
<td>0.0001</td>
</tr>
</tbody>
</table>

*Time was aggregated into 5 minute intervals.
Figure 29. Mean and Variances of Dwell Time in Display Segment 1 (E4) and Display Segment 3 (E5).
Figure 30. Mean of Total Time Spent in Display Segments 1-2, 2-1, 1-3, 3-1, 2-3, and 3-2 (E-14).
VARIANCE OF E14
AGGREGATED ACROSS RUNS, SUBJECTS
AND 5 MINUTE RUN TIME SEGMENTS
VERSUS 5 MINUTE RUN TIME SEGMENTS

Figure 31. Variance of Total Time Spent in Display Segments 1-2, 2-1, 1-3, 3-1, 2-3 and 3-2 (E14).
significant negative correlation \((\text{RHO} = -0.53897, p = 0.001, N = 920)\). The variances of E4 and E5 also are shown across five minute intervals in Figure 29. The variance of E4 is an inverted U-shaped curve with a peak in the range of 300-600 seconds. The variance of E5 is a maximum in the range of 0-600 seconds and then steadily declines over the remaining time.

The means and variances for E14 are shown in Figures 30 and 31. Both the mean and variance of E14 show a steady decline over five minute time segments. This can be interpreted to mean that eye activity declines over time and is less variable.

These results suggest that subjects eye dwell times in display segment 3 achieve a maximum in the range of 600-900 seconds. Dwell times in display segment 1 achieve a minimum during the same time. Eye transition times from one display segment to another decline over time suggesting less eye activity. The interpretation of the results of the analysis of E4, E5, and E14 is limited since signal effects are confounded with time effects. The changes that occur may not occur strictly as the result of time but also may be the result of an aggregation policy-signal schedule artifact. Table 21 indicates that no signals are presented during the first five minutes of runs. One would expect E4 and E5 to be similar in this time period unless an eye movement calibration problem existed or for some reason subjects were predisposed to look initially more in display segment 1 than 3. In fact, a significant difference does exist in the means and variances of E4 and E5 during the first five minutes of runs. The difference between the means of E4 and E5 during the first 5 minutes resulted in
a t-value of 2.139 (df = 418) which was significant at 0.05 in a two-sided test. An F-test for variance yielded a significant F-value of 1.08 (df₁ = df₂ = 209). No signals are present in the time interval 0-300 seconds. Eight signals are presented in the interval 300-600 seconds—six meter changes occur on the left meter. During the interval 600-900 seconds, eight signals are presented—six signals occur on the left meter. Six signals are presented in the interval 900-1200 seconds—five meter changes occur on the right meter. One would expect the means of E₄ and E₅ to be approximately equal in the interval 0-300 seconds. During the interval 300-900 seconds, one would expect the mean of E₄ to be greater than the mean of E₅. During the last interval one would expect the mean of E₅ to be greater than the mean of E₄ if the meter schedule conditions behavior. The mean of E₄ is greater than E₅ during the interval 0-300 seconds. This difference may result from an eye movement equipment calibration constant error or conditioning that occurred because a larger number of signals occurred on the left meter. Meter changes for practice trials (there were two such trials in a 5 minute practice period) always occurred on the left meter and this may have contributed to a conditioning effect in favor of the left meter (the mean of E₄ would be higher). The mean of E₄ decreases during the interval 900-1200 seconds which would support a conditioning hypothesis; however, E₅ is never greater than E₄. The significant results obtained in this analysis are likely due to time, signal, aggregation policy, and meter signal schedule. These
effects are confounded and it is not possible to entirely separate these effects from one another.

**Signal Effects and Eye Movements.** Further analysis was performed to determine whether signals affected eye movement measures. Each of the forty 30 second intervals in all twenty minute runs were examined and coded according to whether a signal was present or absent. An analysis of variance was performed using the following linear model:

\[
E(\text{Eye movement measures}) = \text{signal effects}
\]

This model was found to be significant at the alpha = 0.05 or better for eye movement measures E4 (total dwell time spent in display segment 1--left portion of the display), E5 (total dwell time spent in display segment 3--right portion of the display), E6 (E4 + E5), E11 (display segment 3 dwell time variance), and E14 (percent of total time spent in transition between display segments). E4, E5, and E14 were adopted as indicators of eye activity to be used in further analysis because they were found to be significant in the analysis of time effects previously reported in this section and because they were thought to be representative of eye dwell and transition activity. The F-values and significance levels observed for E4, E5, and E14 were 3.61 (alpha = 0.05), 4.56 (alpha = 0.03), and 13.51 (alpha = 0.0003). The amount of variation accounted for by signal effects in these measures was low; \( R^2 \) was only 1-2 percent but significant. The presence of a signal in a 30 second time interval increases the mean dwell time in both display segment 1 and 3, i.e., the mean of E4 and E5 increases if a signal is present. The analysis is not refined enough to distinguish between types of signals in the 30 second intervals--only whether a signal is present.
or absent. The mean of E14 (percent of total time spent in transition between display segments) decreases when a signal is present.

The results of the previous analyses are presented graphically in Figures 32, 33, and 34. The means for E4 for nonsignal and signal conditions are plotted versus five-minute intervals during a run. The general trends over time that were observed in Figure 29 for E4, E5, and E14 can be observed for both signal and nonsignal conditions in Figures 32, 33 and 34. In addition, the effects of the presence of a signal can be observed. The presence of a signal results in an upwards shift of the nonsignal plot for both E4 and E5. However, a downwards shift of the nonsignal curve is observed in the plot for E14 the presence of a signal results in reduced eye activity.

Heart Activity and Information Processing Analysis Results

Introduction. Heart activity was recorded by interfacing a modified Taber Telecardio Model 202-4 amplifier with the PDP-8 computer. The device had a threshold set point that was set by viewing the heart signal on Tektronix oscilloscope. The device was set to trigger at the peak of each R-wave and not on other parts of the waveform. The oscilloscope was then removed from the circuit for data collection to ensure subject safety. The heart monitoring system triggered a pulse when threshold was exceeded and this pulse and associated time were registered by the computer. All subsequent analysis of heart activity was based on the pulse times.

The research literature indicated that heart rate variability was associated with mental load. As the mental load increases the
Figure 32. Effects of Signals and No Signals on the Means of Dwell Times in Display Segment 1 (E4).
Figure 33. Effects of Signals and No Signals on the Means of Dwell Times in Display Segment 3 (E5).
Figure 34. Effects of Signals and No Signals on the Means of Dwell Times in Display Segment 1-2, 2-1, 1-3, 3-1, 2-3, and 3-2 (E14).
variability of the time between R-wave peaks decreases. As the load lessens, the variability of the heart rate increases. One would think that heart rate variability or the sinus arrhythmia score would decrease around the time that a signal is presented for detection. Another effect that would be expected is that as a S loses alertness in the experiment over time, the heart rate variability would increase over time. The heart signal is more a general response to information processing loads than a specific response. The heart signal would be expected to exhibit response lags in relation to changes in processing loads. Because of the lack of specific response of the heart signal, only general relationships were examined. The following hypotheses were examined: (1) heart rate variability gets smaller after a signal is detected, and (2) heart rate variability increases with time on the task.

Analysis Procedure. The analysis procedures that were used to study the research hypotheses are outlined in Figure 35. FORTRAN programs were written to create heart signal subfiles, detect artifacts and preprocess data for spectral analysis of the R-R interval. Preprocessing of R-R intervals is necessary for spectral analysis because the R-R intervals are not evenly spaced. Preprocessing transforms the train of R-R intervals into a regularly sampled signal—Lagrange interpolation is used to aid this process. The technique has been recommended by Sayers (1973), Mulder and Mulder-Hajonides Van Der Meulen (1973), and Mulder (1979).

Artifacts arise from two sources: (1) the heart itself, and (2) the recording equipment. The heart will at times skip beats or
Figure 35. EKG Analysis Strategy.
exhibit slow or fast rhythms. These conditions are not keyed to information processing tasks and make it difficult to use the heart signal for information processing research. Another source of artifact comes from the experiment and lead connections. If the leads are not secure and do not have a good contact, the recording of the heart signal may be affected. If the threshold is not set properly, a triggering of pulses may occur on parts of the waveform other than the R peaks. Heart signal recording can also be affected by movements on the part of the subject.

Two problems occurred in the recording of the digitized heart signal for Ss. One problem that occurred was the generation of R-R intervals that were small (F ≤ 0.2 seconds) and the other problem was the case where beats appeared to be skipped (R-R intervals > 1.5 seconds). The preprocessing program recorded the frequency of occurrence of these two types of intervals and included a strategy for handling these situations. The short intervals were kept to less than five percent of the total number of R-R intervals. The long intervals where estimation was necessary were kept to less than one percent of the total intervals and in most cases the number of these occurrences was zero. Files that did not meet these data quality requirements were rejected as not suitable for analysis. Less than one half of the files met the criterion for analysis (ten out of twenty-three files were considered suitable for analysis). The 10 subject-run combinations consist of two runs from subject 1, two runs from subject 3 and six runs from subject 4. Any analysis is limited by the scarcity of good data and subject 4's domination of the data.
Performance measures were developed as a part of the descriptive analysis. The following performance measures which are shown in Table 23 were developed and aggregated on the basis of ten-second intervals. This list of measures was later reduced to H1 and H2 for analysis purposes since the information contained in H3, H4, and H5 is contained also in H1 and H2. Later analyses also used an aggregation interval of thirty seconds.

A FORTRAN program was written and used with a CALCOMP plotter and the performance measures were plotted for each subject and each experimental run. Plots were also developed for performance measure means and variances aggregated across subjects for ten-second intervals. Another type of descriptive plot arose from the spectral analysis of R-R intervals. The R-R interval spectrum versus frequency was plotted. Finally, an inferential analysis was performed using non-parametric techniques to address the research questions.

Test of Distributional Assumptions for H1 and H2. H1 is the average R-R interval in a ten-second time segment for a single subject during a single run. H2 is the variance of the R-R intervals during a ten-second time interval for a single subject during a particular run. An histogram was plotted for H1 and H2 using the data from all subjects. These two histograms were tested for normality using a Kolmogorov-Smirnov statistic. Neither the distribution of H1 or H2 was normal at the five percent level of significance ($D = 0.079$, $D = 0.297$ respectively and the critical value of $D$ was $D(0.95, 1200) = 0.039$). This lack of normality was also reflected by the lack of linearity exhibited in normal probability plots for H1 and H2. This lack of

<table>
<thead>
<tr>
<th>Performance Measure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1</td>
<td>Average heart rate - Average of R-R Intervals</td>
</tr>
<tr>
<td>H2</td>
<td>Variance of R-R intervals</td>
</tr>
<tr>
<td>H3</td>
<td>Standard deviation of H2</td>
</tr>
<tr>
<td>H4</td>
<td>Range of R-R intervals</td>
</tr>
<tr>
<td>H5</td>
<td>Cumulative sum of R-R interval variability</td>
</tr>
</tbody>
</table>
normality in the distributions of HI and H2 suggested that nonparametric analyses should be used in the analyses of heart signal measures.

**Descriptive Plots of Means and Variances for HI and H2 Versus Time.** The means were calculated for HI and H2 (MH1 and MH2) for each of the 120 ten-second intervals during a run and plotted versus time. These descriptive plots are shown in Figures 36 and 37. Because these plots were not particularly informative, data was further aggregated into 5 minute intervals. Means and variances for HI and H2 were then calculated. These are shown in Table 24. The 10 subject-run combinations are two runs from subject 1, two runs from subject 3, and six runs from subject 4. The data is dominated by subject 4's EKG. Any analysis is limited by the scarcity of good data and by subject 4's domination of the data.

The mean of HI is stable over all time periods. The mean of H2 exhibits an inverted U-shaped function over five-minute time periods. The variance of HI is high initially and then plateaus at a constant value for the last 15 minutes of the experiment. The variance of H2, like the mean of H1, exhibits an inverted U-shaped relationship over time.

**Effects of Time and Signals on HI and H2.** Earlier it was reported that the distributions of the mean and variance of R-R intervals in each 10-second time interval were not normal. As a consequence of the nonnormality of the HI and H2 distributions, nonparametric tests were used to determine the effects of time and signals on HI and H2.
Figure 36. MHI=Mean of H1 (Mean of R-R Intervals) Aggregated Across 10 Second Intervals and Across Subject-Run Combinations Versus Time.
Figure 37. M_{H2}=Mean of H2 (Mean of Variances) Aggregated Across 10 Second Intervals and Across Subject-Run Combinations Versus Time.
### Table 24. Means and Variances of Heart Rate Performance Measures H1 and H2 Aggregated Across 5 Minute Periods of a Run.

<table>
<thead>
<tr>
<th>Run Time Interval (Minutes)</th>
<th>Heart Rate Measure</th>
<th>N</th>
<th>Mean</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-5</td>
<td>H1</td>
<td>100</td>
<td>0.84806</td>
<td>0.01156</td>
</tr>
<tr>
<td></td>
<td>H2</td>
<td>100</td>
<td>0.01502</td>
<td>0.00048</td>
</tr>
<tr>
<td>5-10</td>
<td>H1</td>
<td>100</td>
<td>0.84509</td>
<td>0.00917</td>
</tr>
<tr>
<td></td>
<td>H2</td>
<td>100</td>
<td>0.01346</td>
<td>0.00029</td>
</tr>
<tr>
<td>10-15</td>
<td>H1</td>
<td>100</td>
<td>0.84055</td>
<td>0.00969</td>
</tr>
<tr>
<td></td>
<td>H2</td>
<td>100</td>
<td>0.01239</td>
<td>0.00026</td>
</tr>
<tr>
<td>15-20</td>
<td>H1</td>
<td>100</td>
<td>0.85289</td>
<td>0.00943</td>
</tr>
<tr>
<td></td>
<td>H2</td>
<td>100</td>
<td>0.01723</td>
<td>0.00045</td>
</tr>
</tbody>
</table>
In order to determine if H1 and H2 changed significantly with time, each run was divided into five-minute time periods. Each five-minute time period consisted of ten, thirty-second intervals. The ten means and variances in each five-minute interval were compared using nonparametric van der Waerden tests. The van der Waerden test was not significant for H1 and H2, i.e., H1 and H2 did not change significantly with time at the five percent level of significance.

The combined effect of auditory and visual signals on H1 and H2 was determined by coding each 30-second interval in a run for the presence or absence of a signal. Two samples were formed; one sample consisted of all 30-second intervals that contained signals while the other sample consisted of all non-signal intervals. A van der Waerden test was used to test the null hypothesis of no difference between samples. The null hypothesis was rejected at the five-percent level of significance ($\chi^2 = 3.84$, df = 1, prob > $\chi^2 = 0.0499$) for H1 but not for H2. The descriptive statistics for H1 indicated that the mean R-R interval is longer for the signal condition than for the nonsignal condition (0.869 versus 0.843). Heart rate slows in the presence of a signal; however, the variance does not change.

**Run Tests for the Mean of H1 (MH1) and the Mean of H2 (MH2).**

Although H1 and H2 were not found to vary significantly across 5 minute time intervals using nonparametric tests, it is possible that changes occurring with time might be masked by a large aggregation interval. H1 and H2 were aggregated on the basis of a 10 second interval and a runs test used to examine trends in the time series MH1
and MH2, i.e. the time series shown in Figures 36 and 37 were tested for significance of runs. The null hypothesis, i.e., \( H_0 \): the pluses and minuses above and below the median occur in random order was rejected for both MH1 and MH2. MH1 and MH2 showed significantly fewer runs than would be expected. The Z-value for a two tailed test was -3.8503 which was significant at the 0.001 level for MH1. The Z-value for MH2 was -2.2002 which was significant at the 0.0278 level. 

Fewer runs than is expected suggests time trends in the data or a bunching of data due to a lack of independence from 10 second interval to 10 second interval. However, little can be said about changes that occur between a single R-R interval and the next because the time resolution of this analysis was only 10 seconds. This type of analysis does not reveal where dependencies exist; only that they do exist.

**Autocorrelation Function of the Time Series MH1 and MH2.** The previous run tests indicated that a lack of randomness existed in the aggregated mean and variance of H1 and H2 (MH1 and MH2), i.e., the mean and variance for 10 second intervals across subjects and runs. The autocorrelation function was calculated for 10 to 300 second lags for MH1 and MH2. Two sigma limits for the autocorrelation coefficients, Box-Pierce statistics and runs tests were performed for the two series. A summary of the results of the analysis is shown in Figures 38 and 39.

The series MH1 lies within ±2 sigma limits with the exception of lags 1, 17, 18, and 25. There are some periodicities in the data. The series can be regarded as stationary since it lies within two sigma limits with the exception of these few coefficients and since
Figure 38. Autocorrelation of Mean Heart Rate—MHI (H1 is Aggregated Every 30 Seconds).
Figure 39. Autocorrelation of the Variance of Heart Rate—MH2

(H2 is Aggregated Every 30 Seconds).
the Box-Pierce Q-statistic (Q=86.07) is less than $\chi^2 (120, 0.95) = 146$. This test indicated the autocorrelations were not significantly different from zero. The runs test for the autocorrelation coefficients yielded $Z = -2.96$ (significant at the 0.003 level) which indicates dependencies in the data.

The correlogram for MH2 indicates the series MH2 is stationary. All autocorrelations lie within ±2 sigma limits. The Box-Pierce Q-statistic is 25.54, considerably lower than the critical value Chi-sq (120, 0.95) = 146. The runs test yielded a test statistic of $Z = -1.84$ which was not significant at the 0.05 level in a two-sided test. These results do not confirm the results reported from the runs test in the previous section.

Spectral Analysis of R-R Intervals for Each Subject-Run Combination. A spectral analysis was performed on the R-R intervals for each subject. After data was suitably preprocessed, the data was then analyzed using PROC SPECTRA from SAS and an output file of the spectrum was created which was later plotted for each subject using a FORTRAN program and flat-bed plotter. PROC SPECTRA created a periodogram which was smoothed by a 15 point moving average. Figures 40 and 41 show examples of the spectral density plots for heart rate data. The plots are for subject 1 run V1 and subject 4 run A1. The plots are representative of the spectral density plots obtained with other subject-run combinations. Peaks in spectral density exist around 0.08 hertz and 0.20 hertz. The power in these peaks may differ from plot to plot and there may be slight shifts around these center frequencies but the general pattern is very consistent. The results
Figure 40. Spectral Density Plot of R-R Intervals for Subject 1--Run VI.
Figure 41. Spectral Density Plot of R-R Intervals for Subject 4--Run A1.
are remarkably similar to the results reported by Mulder (1979). Mulder reported the results of a frequency analysis of R-R intervals and identified three main energy bands. One band lies between 0.00 and 0.20 Hz and is related to body temperature and blood pressure regulation. A second area lies between 0.22 and 0.40 Hz which is associated with respiration activity. A third energy band lies between 0.42 and 0.60 Hz that is related to task activity. Mulder was using a paced task where a subject had to react to signals presented every two seconds. This task does not fit the usual definition of a sustained attention task. Mulder also reported that the first area contains a band from 0.08 to 0.12 Hz that reflected spontaneous fluctuations in blood pressure. The spectral density results reported in the present research show peaks around 0.08 to 0.20 Hz which are probably related to blood pressure oscillations and respiratory fluctuations.

Combined Analysis of Performance Measures

Introduction. The strategy for analysis of combined performance measures consisted of four major analyses: (1) a review of previous combined analysis strategies, (2) a simple linear correlation analysis of all heart rate and eye movement measures, (3) a cross-spectral analysis of eye movement and heart rate measures, and (4) a correlation of EEG measures with response time.

The combined analysis of dependent measures did not provide results that were as insightful as initially anticipated. This occurred for several reasons: (1) measures such as response time were
singular responses to a signal whereas other measures were continuous, (2) continuous measure aggregation intervals were large and information was lost, (3) EKG, EEG, and eye movements were difficult to reference to display signals as well as to each other, (4) the EKG data was limited--60% of the available data was from a single subject, and (5) the number of subjects was small.

**Previous Combined Analysis Strategies.** The analysis results presented in previous sections were results that were obtained from combined analysis considerations. For example, the analyses of the EEG was keyed to the response time analysis. Eye movement and heart rate analyses also considered the results obtained from the response time analysis.

The analysis results discussed here focus primarily on the relationships between dependent measures--particularly eye movement and heart rate measures since these measures were recorded during the entire experimental run.

**Correlation Analysis of Eye Movement and Heart Rate Measures.** Correlation coefficients were determined for the means and variances of eye movement performance measures with heart rate measures. The means and variances were determined for performance measures aggregated across subjects and across each one of 40 time intervals. An option was invoked in the analysis procedure which gave the five highest correlations for the mean and variance of eye movement performance measures with the means and variances of heart rate.
performance measures. The results are summarized in Table 25. Relationships are suggested between heart rate measures and eye movement measures. MH1 (mean across all subjects for each 30 second time interval) is correlated with VE7 (α<10%). MH2 (mean of heart rate variability across each 30 second time interval for all subjects) is significantly correlated with VF2 (variance of arcsine (E14) and VF5 (variance of the arcsine transformed percent of total time spent in display segments 1 and 3). VH1 is significantly correlated with eye movement measures VF5, VE6, and VE17 (VF5 is VE17 arcsine transformed). VH2 is significantly correlated with VE3, VE7, and VE9. Results are reported in Table 25 in descending order of significance.

Cross spectral Analysis of Eye Movement and Heart Rate Measures. In addition to linear correlation analyses, cross-spectral analyses were performed for the means and variances of the following performance measures: E4, E5, E9, E12, E14, F5 crossed with H1 and H2 and aggregated across subjects and runs. These aggregated measures are designated ME9, VE9, ME12, VE12, etc. M indicates the mean of the measure whereas V indicates the variance. Eye movement measures were crossed with heart rate measures to explore time relationships that might be present. The above measures were selected as being representative and sensitive measures of eye movement and heart behavior. Four kinds of plots were made for the analysis: (1) spectral density of eye movement measures versus frequency (S-01 vs. FREQ), (1) spectral density of heart rate measures versus frequency (S-02 vs. FREQ), (3) coherency squared versus frequency (K-01-02 vs FREQ), and (4)
Table 25. Correlation Coefficients and Levels of Significance for the Mean and Variance of Heart Rate Measures with the Mean and Variance of Eye Movement Measures (N=40).*

<table>
<thead>
<tr>
<th>Mean of Heart Rate Measure</th>
<th>Correlated Measure</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>VH1</td>
<td>MH2</td>
<td>VE7</td>
<td>VE16</td>
<td>VF5</td>
</tr>
<tr>
<td>MH1</td>
<td>0.81401</td>
<td>0.37159</td>
<td>-0.26483</td>
<td>0.23577</td>
<td>0.21354</td>
</tr>
<tr>
<td></td>
<td>0.0001</td>
<td>0.0182</td>
<td>0.0986</td>
<td>0.1430</td>
<td>0.1858</td>
</tr>
<tr>
<td></td>
<td>VH2</td>
<td>VH1</td>
<td>MH1</td>
<td>VF2</td>
<td>VF5</td>
</tr>
<tr>
<td>MH2</td>
<td>0.42759</td>
<td>0.37938</td>
<td>0.37159</td>
<td>0.31239</td>
<td>0.31106</td>
</tr>
<tr>
<td></td>
<td>0.0059</td>
<td>0.0158</td>
<td>0.0182</td>
<td>0.0497</td>
<td>0.0507</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variance of Heart Rate Measure</th>
<th>Correlated Measure</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MH1</td>
<td>VF5</td>
<td>MH2</td>
<td>VE6</td>
<td>VE17</td>
</tr>
<tr>
<td>VH1</td>
<td>0.81401</td>
<td>0.39207</td>
<td>0.37938</td>
<td>0.35728</td>
<td>0.35728</td>
</tr>
<tr>
<td></td>
<td>0.0001</td>
<td>0.0123</td>
<td>0.0158</td>
<td>0.0236</td>
<td>0.0236</td>
</tr>
<tr>
<td></td>
<td>VE9</td>
<td>MH2</td>
<td>VE7</td>
<td>VE3</td>
<td>ME9</td>
</tr>
<tr>
<td>VH2</td>
<td>0.51625</td>
<td>0.42759</td>
<td>0.39507</td>
<td>0.27353</td>
<td>0.23656</td>
</tr>
<tr>
<td></td>
<td>0.0007</td>
<td>0.0059</td>
<td>0.0116</td>
<td>0.0877</td>
<td>0.1417</td>
</tr>
</tbody>
</table>

*The designations "M" and "V" indicates the mean and variance across subjects for each performance measure. The numbers in the first row for each heart rate measure are the correlation coefficients and the numbers in the second row are the levels of significance for the coefficient tested against zero.
phase spectrum in radians versus frequency (PH-01-02 vs FREQ). The spectral density plots for the individual series (S-01 and S-02) show the distribution of variance over the frequencies between 0 to π radians. The coherency squared is analogous to a squared correlation coefficient and shows the degree of dependency that exists between two time series at a given frequency (K-01-02 ranges from 0 to 1). The phase spectrum of two signals indicates the lag and lead relationships between the two signals. For example, a phase spectrum value of 0 shows two signals to be "in step" with each other. A value of +2.00 indicates that the first signal leads the second by a positive two radians (about 115 degrees). A negative 2.00 indicates the second signal lags the first by about 115 degrees. Figure 42 is an example of the plotting procedure that was used for these four kinds of plots. Figure 42 shows individual spectral density plots (S-01 and S-02) for VE9 and VH2 respectively as well as coherency squared and phase spectrum plots. VE9 and VH2 were shown to have a significant linear correlation from previous analysis at a lag of zero.

The spectral density plot for VE9 (the variance of E9 aggregated across subjects) is shown as S-01 versus frequency. The variance declines with frequency although slight peaks of power exist at 0.8 radians and 2.3 radians.

The spectral density plot for VH2 is shown as S-02 versus frequency. A prominent peak exists at 0.80 radians. Another less prominent peak exists in the frequency range 1-60-2.40 radians.
Figure 42. Spectral Density Plots (S-01, S-02), Coherency Squared Plot (K-01-02) and Phase Spectrum Plot (PH-01-02) in Radians Versus Frequency in Radians for VE9 and VH2.
The coherency squared (K-01-02) plot for VE9 and VH2 is shown in the upper right hand corner of Figure 42. The peak coherency is approximately 0.92 at a frequency of 0.80 radians. The two signals are incoherent beyond a frequency of 1.0-1.2 radians.

The phase spectrum indicates that VE9, i.e., the variability of E9 is in phase with VH2, the variability of heart rate variability throughout a range of 0.0 to around 1.0-1.2 radians. When the frequency is around 1.8 radians, the eye movement measure lags the heart rate measure by 0.50 radians. Beyond a frequency of 1.8 radians, VE9 always leads VH2 by at least 0.5 radians. A lead of 0.5 radians in the frequency domain represents about 70 seconds in the time domain. The bandwidth resolution is broad due to the choice of the aggregation interval. The Nyquist frequency imposes a limit on the highest frequency about which one can get meaningful information. With an aggregation interval of 30 seconds, the Nyquist frequency is 1 cycle per 60 seconds. This situation is analogous to taking temperatures once per day at noon in a certain town. The observations tell us nothing about temperature variations within a day. In order to obtain more information about higher frequencies, the bandwidth resolution must be narrowed. A tradeoff exists between meaningful data aggregation levels and bandwidth resolution.

Correlation of EEG Measures With Response Time. Subjects were instrumented so the EEG signal was acquired 5 seconds before, 5 seconds during, and 5 seconds after the presentation of a visual signal for detection at 315, 635, and 955 seconds. This was done so as not to exceed the memory limitations of the PDP-8 computer. The
EEG signal was broken into the frequency components $\delta$, $\Theta$, $\alpha$, and $\beta$ for each 5 second segment using a spectral analysis routine. In addition, the mean ($Y_1$) and variance ($Y_2$) of the EEG signal was determined for each 5 second time segment. Response times were correlated with EEG measures for each 5 second time segment around the visual signals at 315, 635, and 955 seconds. Response times were not significantly correlated with EEG measures in the 5 second interval before a visual signal was presented for detection. However, alpha ($\alpha$) was positively correlated with response times during the 5 second interval when a visual signal was presented ($RHO = 0.42687$, significant at the 0.0014 level, $N = 53$). Delta ($\delta$), a low frequency EEG component was negatively correlated with response times in the 5 second interval after a visual signal was presented ($RHO = -0.29568$, significance was 0.0316 and $N = 53$). Caution should be used in the interpretation of these results since the data consisted of repeated observations on only four subjects. The $N=53$ observations cannot be viewed as fifty-three independent observations since only four subjects generated this data.
5. SUMMARY OF RESULTS, INTERPRETATIONS, AND RECOMMENDATIONS.

Introduction

The final chapter of this dissertation is organized into four sections; (1) summary of major findings of research, (2) review of vigilance theory, (3) interpretation of research results in terms of relevant vigilance theories, and (4) recommendations for future research suggested by present research.

Summary of Major Findings

The purpose of this study was to examine the effects of changes in display and task variables on human visual, physiological, and psychomotor behavior in a sustained attention task. More specifically, the study investigated the effects of artificial and warning signals on response time, EEG, eye movement dwell time and activity measures, and EKG in a sustained attention task.

The initial step in the research was a review of the research literature on sustained attention theories, task influences on performance, task duration, signal expectancy, signal duration, artificial signals, warning signals, signal intensity, and sustained attention task performance measures. Although the independent variables were "time on the watch," intersignal intervals, and artificial and warning signals, other influences had to be researched in the literature to design a research task, e.g., signal duration and expectancy for a signal.
The literature review provided information about the potential for a research study as well as information about research and experimental task design. As the result of the literature review and judgment, a research task was designed that was 20 minutes duration and investigated the effects of three conditions: (1) a control condition, (2) an artificial signal condition, and (3) an auditory warning signal condition. In addition, the research investigated the effects of a variable interval between artificial signals and warning signals and signals presented for detection. A control condition was used as well as 5 to 60 second intersignal interval conditions. Time on the watch was another independent variable that was investigated in the research.

In order to accomplish the research objectives, a real time experimental system was designed that used a PDP-8 minicomputer to present the experiment to four subjects each in six experimental runs as well as record their responses.

The analysis of data focused on five areas: response times, EEG, eye movements, EKG, and multivariate analysis.

Table 26 summarizes the major findings of the research. The research design was unbalanced and sample sizes were small and consequently it was necessary to qualify the results in Table 26 with appropriate comments about the validity of each result.

Vigilance Theory Summary

Early in the chapter, a summary of major findings were presented in Table 26. A more detailed discussion of the various vigilance
## Table 26. Summary of Major Findings.

<table>
<thead>
<tr>
<th>Dependent Variables</th>
<th>Observed Result</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Response Time</td>
<td>A one-way ANOVA resulted in a significant ISI effect on mean response time. Mean response time is lower with a short ISI higher with a long ISI, and insignificant for intermediate ISIs. Response time variance increased with ISI.</td>
<td>The ANOVA resulted in an F-value of $F = 3.651$ ($v_1 = 3$, $v_2 = 72$) which was significant at a level of $0.0164$; $R^2 = 13.2%$. The highest order polynomial fit to the data was linear and as follows: response time = 1.6895 + 0.0025 * ISI. The power of the F-test was approximately $(1 - \beta) = 0.76$; Tukey's HSD resulted in a significant difference between responses to signals with an ISI of 5 seconds and those with ISI = 315 seconds (the experiment wise error rate was at least 0.05). Response times for signals with ISIs of 5 seconds were significantly lower and less variable than response times with ISIs of 315 seconds. It is thought the ISI effect is unconfounded with run, time, and modality effects; however the evidence for no run, time, or modality effects is weak because of the low power of the statistical tests. More complex models such as $E(RT) = \mu + ISI + Subject$, etc. were tried; however, blocking on subjects was not effective—error degrees of freedom was reduced too much.</td>
</tr>
<tr>
<td>EEG Alpha Wave Mean power</td>
<td>Type of prior signal effect. Prior auditory warning signals with ISIs of 5 or 60 seconds resulted in higher amplitude EEGs in the alpha frequency range than prior artificial visual signals with ISIs of 5 or 60 seconds. A prior visual artificial signal with an ISI = 315 seconds did not produce a significant difference when compared with the other above conditions.</td>
<td>An $S \times T$ (subjects x treatment) design model was assumed for the analysis after non-detection of signal data was deleted. Treatment is type of prior signal ($X_2$). Model values were $F(11, 156) = 3.19$, Pr &gt; $F = 0.0007$, $R^2 = 18.3%$. The contribution to the model of the type of prior signal ($X_2$) had borderline significance (alpha = 0.05).</td>
</tr>
</tbody>
</table>
Table 26. (continued).

<table>
<thead>
<tr>
<th>Dependent Variables</th>
<th>Observed Result</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eye Movement Measures</td>
<td>The mean values of E4, E6, E13, E16, E17, F1, F4, and F5 are positively correlated with time.</td>
<td>These correlation coefficients and their significance levels are reported in Table 4.17. Significance levels when Rho is tested against zero are all less than 0.05. The measures E4, E6, E13, E17, F1, F4, and F5 are all dwell time measures and all correlations are positive, i.e. the results suggest that eye dwell times increase with time (N = 920). Signal effects are contained within these correlations.</td>
</tr>
<tr>
<td>Eye Movement Measures</td>
<td>The mean values for E14 and F2 are negatively correlated with time.</td>
<td>The significance levels for E14 and F2 when Rho is tested against zero are less than 0.05. The measures E14 and F2 are activity measures and negatively correlated with time, i.e. the results suggest that eye activity declines with time (N = 920). Signal effects are contained within these correlations.</td>
</tr>
<tr>
<td>Eye Movement Measures (Both dwell time and activity measures)</td>
<td>The variances for E1, E2, E3, E5, E6, E8, E9, E11, E14, E15, E17, F2, F3, and F5 are all significantly negatively correlated with time.</td>
<td>Measures of activity (E1, E2, E14, and F2) become less variable with time as do measures of dwell (E5, E6, E8, E9, E11, E15, E17, F3, and F5). Correlations are significantly different than zero at the 0.05 or better level. Signal effects are contained within these correlations.</td>
</tr>
<tr>
<td>Eye Movement Measures (Total dwell time spent in display segment 1)</td>
<td>Significant time effects in E4. (E4 is a dwell time measure.)</td>
<td>A main effects subject, time, and runs model was used to analyze the data. Significant subject, time, and run effects were found, $F_{model} = 20.26$, $\alpha = 0.001$, $R^2 = 21.2%$. Time, subject, and run effects were significant at the 0.001, 0.0001, and 0.0067 levels respectively. Figures 4.14 shows that the variance of E4 displays and inverted u-shaped curve. E4 has a peak variance in the</td>
</tr>
<tr>
<td>Dependent Variables</td>
<td>Observed Result</td>
<td>Comments</td>
</tr>
<tr>
<td>---------------------</td>
<td>----------------</td>
<td>----------</td>
</tr>
<tr>
<td>Eye Movement Measure</td>
<td>Significant time effects in E5. (E5 is a dwell time measure.)</td>
<td>A significant time effect was reported in Table 4.19 when E5 was modelled using a main effects subject, time, and runs model. Model values were $F = 12.60$, $\alpha = 0.0001$ and $R^2 = 14.3%$. All effects were significant. The mean of E5 reaches a minimum in the range of 600-900 seconds. The variance of E5 is a declining function with time beginning in the range of 600-900 seconds. These trends are shown in Figure 4.14. The presence of a signal significantly shifts the mean dwell time in display segment 3 upwards—this can be observed in Figure 4.18.</td>
</tr>
<tr>
<td>E5 (total dwell time in display segment 3.)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eye Movement Measure</td>
<td>Significant time effects in E14 (E14 is an eye activity measure.)</td>
<td>A significant time effect was reported in Table 4.19 for E14. The mean of E14 shows a decline with time—this trend can be seen in Figures 4.15. The mean activity during the first 5 minutes of a run is significantly higher than the last 5 minutes of a run. The variance of E14 follows a similar curve which is shown in Figure 4.16. The presence of a signal significantly shifts E14 downwards. This can be seen in Figure 4.19.</td>
</tr>
<tr>
<td>E14 (% of total time spent in transition across display segment boundaries)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EKG Measure H1 (Mean heart rate)</td>
<td>Mean heart rate H1 affected significantly by the presence of a signal—either auditory or visual or both.</td>
<td>A nonparametric van der Waerden test showed a significant signal effect ($\alpha = 0.0499$). The mean H1 is longer for the signal condition than for the nonsignal condition.</td>
</tr>
<tr>
<td>Dependent Variables</td>
<td>Observed Result</td>
<td>Comments</td>
</tr>
<tr>
<td>---------------------</td>
<td>----------------</td>
<td>----------</td>
</tr>
<tr>
<td>EKG and Eye Movement Measure Correlations</td>
<td>Significant correlations were found between EKG measures and eye movement measures.</td>
<td>The validity of this result is questionable due to the poor quality of EKG data; 60% of the data is from subject 4. However, the variance from individual differences is low with EKG measures.</td>
</tr>
<tr>
<td>EEG and Response Times</td>
<td>Significant correlation between EEG measure α and response times in the 5 second interval during which a signal was presented for detection.</td>
<td>These correlations are reported in Table 4.22. Significant correlations were found with the following pairs of variables: (MH2, VF2), (VH1, VF5), (VH1, VE6), (VH1, VE17), (VH2, VF9), (VH2, VE7). The common time interval upon which these correlations are based is broad; 30 seconds.</td>
</tr>
<tr>
<td>EEG and Response Times</td>
<td>Significant correlation between EEG measure 6 and response times in the 5 second interval prior to the time a visual signal was presented for detection.</td>
<td>Rho = 0.29686 for 6 versus response time; significant at the 0.0014 level, (N = 53). This result suggests that the amplitude of α waves increases when response times are longer. The literature reports that α rhythms are blocked or suppressed during alertness (amplitude decreases, frequency of EEG increases). The sample sizes are small in this analysis and prior signal modality and ISI effects are a part of this correlation coefficient.</td>
</tr>
<tr>
<td>EEG and Response Times</td>
<td>Significant correlation between EEG measure 6 and response times in the 5 second interval prior to the time a visual signal was presented for detection.</td>
<td>Rho = 0.42687 for α versus response time; significant at the 0.0014 level, (N = 53). This result suggests that the amplitude of α waves increases when response times are longer. The literature reports that α rhythms are blocked or suppressed during alertness (amplitude decreases, frequency of EEG increases). The sample sizes are small in this analysis and prior signal modality and ISI effects are a part of this correlation coefficient.</td>
</tr>
</tbody>
</table>

Rho = 0.29568 for 6 versus response times significant at the 0.0316 level. This result suggests that the low frequency component of EEG, 6, associated with low arousal decreases just prior to the presentation of a visual signal for detection. Response times decrease as 6 decreases. Again this correlation analysis mixes conditions, some visual signals are preceded by visual artificial signals with ISIs = 5, 60, or 315 seconds.
theories and an assessment of the theories was provided in the literature review in Chapter 2. This section provides an integration of theories. Table 27 provides a summary of various theories with comments concerning vigilance research results described by the theory. This section provides a framework by which to assess the results of the present research in terms of existing theories.

The previous discussion and development suggests several conclusions concerning vigilance theory. There is no single comprehensive vigilance theory that integrates the research findings. No one has presently accomplished this integrative task. It seems that a deeper understanding of vigilance behavior is limited by the state of the art. Warm's (1977) observation that vigilance research is in the third phase of its development is more than likely true. The third phase of evolution is described as a search for deeper understanding and exploration. Warm's (1977, 643-644) comments best summarize the existing situation:

... much remains to be done. Gaps in our knowledge exist and a complete psychophysical description of the relevant variables, as well as their relations to each other, is still to be written. A variety of theoretical models has been proposed to account for vigilance effects. Each has focused upon a somewhat different aspect of the problem, yet many can account for similar data. Further, it is difficult to establish a definitive test of one against the other and they all invite criticism on several grounds. It seems likely that future theorizing in this area will have to develop a synthesis of these different points of view.

It is worth noting that vigilance represents only one of several sub-categories in the general area of attention—an area which is of great interest to modern experimental psychology. In addition to vigilance, attention also encompasses problems of selection, concentration, search, and set. The degree to which vigilance has properties in common with these other sub-categories also remains to be determined.
<table>
<thead>
<tr>
<th>Vigilance Theory</th>
<th>Vigilance Decrement</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Classical conditioning/reactive inhibition (N.H. Mackworth, P.D. McCormack)</td>
<td>The unconditioned stimulus (Command of &quot;Now&quot;) and reinforcement (KOR - &quot;Yes that was correct&quot;) received during the practice session is not present during the experimental session. Inhibition to respond builds leading to extinction of the key pressing response.</td>
<td>Inhibition theory by itself is generally not accepted as an explanation for vigilance effects although the inhibition construct is a part of many current theories. Inhibition theory would predict that as the ISI increased inhibition would dissipate improving performance. Artificial signals should worsen performance because there are more unreinforced signals. Warning signals should be disinhibitory.</td>
</tr>
<tr>
<td>(2) Observing responses-operant conditioning (J.G. Holland)</td>
<td>Detections reinforce observing response. When observers are monitoring limited-hold signals on a variable interval schedule such as a Mackworth schedule and miss signals this has the effect of reducing observing responses. This further reduces detections.</td>
<td>Most of the work with observing responses used unlimited hold signals--little can be said about response times. Adding more signals to a signal schedule provides more reinforcement for observing responses if signals are detected. Broadbent (1971) relates this theory to expectancy--the less variable the signal schedule the more certain the observer is about when to observe (on a fixed interval schedule, the observer will not observe just after a signal).</td>
</tr>
<tr>
<td>(3) Elicited observing rate response theory (H.J. Jerison)</td>
<td>Jerison found a decrement with a high background event rate (30/minute) but not with a low background event rate (5/minute). He developed a theory to explain this result based on the following: (1) a decision model with two kinds of decisions--decisions about whether or not to observe the display and</td>
<td>Jerison found a decrement with high background event rates. Stroh (1971) has questioned whether Jerison's task can be considered a vigilance task due to high event rates. Event intervals were clearly defined unlike the usual vigilance task. Jerison's theory emphasizes intervening</td>
</tr>
</tbody>
</table>
Table 27. (continued).

<table>
<thead>
<tr>
<th>Vigilance Theory</th>
<th>Vigilance Decrement</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>(3) Continued</td>
<td>decisions about whether or not an event is a signal (vigilance according to Jerison, focuses on the first decision), (2) eliciting an observing response depends upon the relative costs and rewards of observing, (3) the costs of observing increases over time due to inhibition, and (4) observing quality may be perfect, blurred, or distracted; unrealistically high values of $\beta$ may be due to mixtures of types of observing over time (Jerison's model assumes changes in both $d'$ and $\beta$ although he only observed changes in $\beta$.)</td>
<td>variables which are unmeasurable and unobservable. Jerison's observing response idea is much like Broadbent's notion of filtering. It is possible that habituation is a factor in the vigilance decrement at high event rates. The notions of habituation and filtering are similar.</td>
</tr>
<tr>
<td>(4) Signal detection theory (SDT) (J.F. Mackworth, John A. Swets)</td>
<td>Monitoring tasks involve both discrimination and decision. $d'$-prime is a measure of discrimination or sensitivity and $\beta$ is a measure of decision bias. The vigilance decrement is generally associated with a criterion increment, i.e. a shift in the direction of more cautious behavior. (Loeb (1978) claims that decreases in FAs and increases in $\beta$ represent increases in learning and not criterion shifts.) This conclusion was derived from the observation in many studies that detections as well as false alarm decreased over a watch. $\beta$ increases as the signal probability decreases. The effects on $\beta$ of varying payoff matrices or providing KOR is equivocal. A few studies have shown the vigilance decrement to be associated with</td>
<td>SDT applied to vigilance has been questioned on several grounds. Craig (1977) claims that too often individual ROCs are averaged with the result that the shape of individual ROCs are masked. The averaged data may suggest an equal-variance SDT model. The researcher then calculates $d'$ and $\beta$ based on this model. Individual ROCs may suggest a different model. $d'$ and $\beta$ may be seriously inflated. Psychophysical experiments observers are expected to have several hundred observations before behavioral stability is reached. Vigilance research, particularly in operational settings may not permit this. This leads to a data reliability...</td>
</tr>
</tbody>
</table>
Table 27.

V ig ila n c e

(4 )

C o n tin u e d

T h e o ry

V ig ila n c e

(c o n tin u e d ).

d e c re m e n t

Comments

d e c r e a s e I n d ' as w e l l as in c r e a s e s I n p .
These
s e n s i t i v i t y s h i f t s a r e g e n e r a l l y fo u n d I n s t u d i e s
t h a t h a v e h ig h e v e n t r a t e s ( > 2 4 e v e n t s / m i n u t e )
I . e . w h e re t h e s e n s e r e c e p t o r has t o be p o in t e d
and h e ld op en s u c h as d i s c r e t e — f a s t o r c o n ­
t in u o u s t a s k s an d w i t h t a r g e t s t h a t r e q u i r e

p r o b le m when t h e

SOT m odel

is

a p p lie d

to

s u c c e s s iv e d i s c r i m i n a t i o n s s u c h as d e t e c t i n g t h e
In c r e a s e i n d e f l e c t i o n o f a m e t e r n e e d le ( a s
o p p o s e d t o t a r g e t s t h a t r e q u i r e a s im u lt a n e o u s
d is c r im in a t i o n , I . e . th e t a r g e t is s p e c if ie d
f u l l y w i t h i n a s t im u lu s e v e n t , t a r g e t an d non­
t a r g e t f e a t u r e s a r e p r e s e n t e d s im u lt a n e o u s l y ,
e .g . d e te c tin g a s p e c ifie d c o n fig u r a tio n in a
m ore c o m p le x p a t t e r n ) .

The u s e o f s u p r a t h r e s h o ld s i g n a l s i n
SDT g e n e r a t e lo w o r z e r o FAs— d ' c a n n o t
b e d e t e r m in e d .
Long and Waag ( 1 9 8 1 ) a ls o
m e n tio n e d t h a t in c r e a s e s i n p may r e f l e c t
le a r n in g r a t h e r th a n c r i t e r i o n s h i f t s .
SDT assum es t h a t d e c i s i o n g o a ls do n o t
c h a n g e fro m t r i a l t o t r i a l .
H o w e v e r, v a r i
a n c e i n t h e c r i t e r i o n may e x i s t w h ic h has
t h e e f f e c t o f i n c r e a s i n g s ig n a l and n o is e
v a r i a b i l i t y ; a p p a r e n t d e c r e a s e s i n s e n s i­

v ig ila n c e s e ttin g s .
C o r r e la tio n s a re o fte :
fo u n d i n v i g i l a n c e s e t t i n g s b e tw e e n d ' a nd
p.
SDT w o u ld s u g g e s t t h e s e m e a s u re s a r e
in d e p e n d e n t .
T h is p r o b a b ly r e f l e c t s
in a d e q u a t e a s s u m p tio n s a b o u t t h e m o d e l.

t i v i t y may r e s u l t ( D a v ie s and P a ra s u r a m a n ,
1 9 8 2 ).
E s t im a t i n g c r i t e r i o n v a r i a n c e i n a
SDT m odel i s d i f f i c u l t b e c a u s e a n o t h e r f r e
p a r a m e t e r i s in t r o d u c e d .
V ic k e r s e t a l
( 1 9 7 7 ) s t u d i e d a d a p t a t i o n t o d e c r e a s in g
s ig n a l p r o b a b i l i t y .
T h e y fo u n d t h a t Ds
and FAs in c r e a s e d w i t h t im e .
d ‘ d id n o t
ch a n g e b u t p d e c r e a s e d w i t h d e c r e a s e d
s ig n a l p r o b a b i l i t y .
These r e s u lt s a re
o p p o s it e t o t h e u s u a l SDT f i n d i n g s .
The
a u t h o r s s u g g e s te d t h a t t h e d i f f e r e n c e s
b e tw e e n s h o r t r u n p r o b a b i l i t y e s t i m a t e s
and lo n g te rm e s t i m a t e s i s m in im iz e d .
SD1
m o d e ls may n o t be a p p r o p r i a t e f o r s '> .e
v ig ila n c e s itu a t io n s .
SDT c a n be u s o i u l

ro
oo

uo


Table 27. (continued).

<table>
<thead>
<tr>
<th>Vigilance Theory</th>
<th>Vigilance Decrement</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>(4)  Continued</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(5) Probability matching (A. Craig)</td>
<td></td>
<td>The vigilance decrement is explained in terms of a model adopted from decision theory and probability learning. The model may describe behavior of naive, ill-informed, poorly trained observers. The decrement occurs where the signal rate is low, e.g. P(s) &lt; 0.02 and the observer expects and initially responds as if the signal rate is higher. The observer adjusts his response frequency downwards to match the signal rate as he gains more experience in the task. This results in the observed decreases in detections and false alarms.</td>
</tr>
<tr>
<td>(6) General activation/arousal theory-classical activation theory (W.E. Scott, E. Duffy)</td>
<td>Performance is related to activation in an inverted u-shaped unidimensional relationship. Both Duffy (1962, 1972) and Scott (1969) favor this view. More recently, Andreassi (1980) gives support for this view. Duffy viewed activation as the release of stored energy through metabolic activity in the tissues. Low activation is associated with a lack of alertness, lack of readiness to respond, and diminished activity. A high level of activation can harm performance due to behavioral disorganization. Perception is narrowed and the range of cue utilization is decreased.</td>
<td>The classical activation theory model is appealing because of its simplicity. However, the model does not fit some of the vigilance data. One instance where this seems to be true is in the situation where combined stressors are studied and a two-arousal mechanism is suggested. Evidence that activation does influence performance is shown by drug studies where performance is maintained by stimulants and lowered by depressants. Moreover, time of day effects are difficult to explain without arousal theory.</td>
</tr>
</tbody>
</table>

for the interpretation of vigilance results but care and caution need to be exercised in modeling.

A probability matching model may be appropriate to some signal detection situations where the observer is naive. There is some empirical support for such a model but the exact form of the model that best describes the data is unclear.
Table 27. (continued).

<table>
<thead>
<tr>
<th>Vigilance Theory</th>
<th>Vigilance Decrement</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>(6) Continued</td>
<td>Duffy asserts that activation is measurable and follows lawfulness in its variations. Duffy (1972, p. 598) states, &quot;Activation theory does not demand that all measures march in step alike an army but only that significant trends between most of the measures be concordant.&quot; Duffy disagrees with the position favored by Lacey of dissociation of autonomic, cortical, and behavioral systems. She discounts the dissociation position on the basis of inadequate measurement. Scott (1969) favors Duffy's view with the exception that the brain stem reticular formation plays more of a role in activation than accorded by Duffy. Activation level is influenced by drugs, hormones, chemical products of fatigue or stimuli past, present, or future. It is also affected by stimulus intensity, variation, complexity, uncertainty and meaningfulness. The vigilance decrement is due to a lack of variety in relevant and irrelevant stimuli. Because of the uniform environment, i.e. a lack of stimulus variation, habituation may arise in the reticular formation leading to a decrement in performance. A decrement may also arise from over-arousal. Welford (1976) has interpreted vigilance results in terms of classical activation theory and signal detection theory in terms of a fixed absolute criteria and shifts in noise and signal distributions right or left.</td>
<td></td>
</tr>
</tbody>
</table>
Vigilance Theory


Recent views of arousal are considerably more complex than the theory outlined above. Moreover, arousal theory suffers from a lack of agreement concerning definitions and a precise knowledge about the functioning of specific parts of the brain. Pibram and McGuiness (1975) evaluated nearly 200 studies and concluded there are three separate but interacting neural attentional systems: arousal, activation, and effort. Arousal is identified with the orientation reaction, i.e., phasic responses to input. The arousal circuits center in the amygdala. Activation represents a tonic response to input and is identified with a readiness to respond or vigilance. Activation control circuits seem to be located in the basal ganglia of the forebrain. A third control process regulates arousal and activation, an operation that demands effort. The control circuits for effort center in the hippocampus. Amygdala control circuits are facilitatory or inhibitory. An external stimulus is compared with an internal representation; a neuronal model in memory. The neuronal model is constructed and refined with repeated presentation of a stimulus. With each presentation, the error signal decreases and less orientation is produced. In a sense, there is an uncoupling of

Vigilance Decrement

There is not a unified state theory but instead a diverse body of often conflicting and controversial results. Research is evolving very rapidly in this area and it is likely that this area will contribute greatly in the future to a comprehensive theory of vigilance. Presently, an interpretation of research results in terms of this body of findings requires a review of many individual research studies instead of a summarizing and organized theory. Several reviews of this area are available, for example, Eysenck (1982) is one of the most organized and comprehensive reviews of arousal and attention in terms of cognitive psychology. Davies and Parasuraman (1982) provide a good summary of this area. Earlier, Mackworth (1969, 1970) and Broadbent (1971) provided a thorough review of this area. Control to modern activation theory is brain signal processing. In view of the fact that large areas of the brain have not been functionally mapped, progress will continue in activation theory to the extent this mapping is accomplished. Habituation has been offered as an explanation of the vigilance decrement in high event rate tasks such as Jerison's (Mackworth 1989)
Table 27. (continued).

<table>
<thead>
<tr>
<th>Vigilance Theory</th>
<th>Vigilance Decrement</th>
<th>Comments</th>
</tr>
</thead>
</table>
| (7) Continued                                                                     | stimulus and response with repeated presentations of a stimulus. Arousal asks the question, "What is it?" and activation asks the question, "What is to be done with it?" Pibram and McGuinness have identified serotonergic "Stop" mechanisms associated with arousal that act as a reequilibrium mechanism and a dopaminergic and norepinephrinergic catecholamine "Go" mechanism associated with readiness to respond and activation. A third mechanism in the hippocampus coordinates the arousal and activation mechanism. This control system leads to changes in the central representations, i.e. changes of state, set, or attitude. Living systems can alter their competence to process information. Increases in competence reflect increases in complexity and efficiency of the neuronal model to handle incoming information. Effort can be defined in terms of the attention "paid" to increase or maintain efficiency by reducing equivocation. Although, the authors do not mention the vigilance decrement per se, presumably the vigilance decrement would arise because of the effort required to bring about habituation. Situations where the match between model and input is good defined habituation. In such circumstances, the stimulus does not require arousal or activation and processing becomes automatic. Pibram and McGuinness where sensitivity decrements occur such as: Jerison's (Mackworth, 1969). Mackworth was suggested that d' exponentially decays with time. Mackworth also claims the evoked EEG habituates and this parallels performance decrements. There is no empirical support of these claims to date (Davies and Parasurman, 1982). However, vigilance decrements generally show up anywhere from 5-30 minutes if they show up. Habituation to signals is generally much faster--say around a minute or less (Grove and Thompson, 1970). Habituation theory must reconcile these time differences. A dual-arousal mechanism is supported by studies with two stressors where results on performance are larger or smaller than predicted. A single unitary theory has difficulties with interactions of stressors that are often observed, e.g. a unidimensional theory would predict that low doses of alcohol would decrease performance and incentives would improve performance. The results are quite different; what appears to happen is that incentives increase the adverse effects of alcohol rather than cancelling them. A single arousal mechanism is probably not sufficient to explain such findings. Other findings difficult
Table 27. (continued).

<table>
<thead>
<tr>
<th>Vigilance Theory</th>
<th>Vigilance Decrement</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>(7) Continued</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

suggest that the contingent negative variation (CNV) has been thought to reflect expectancy but probability reflects the central event of matching input to the neuronal model and readiness to respond. The full meaning of the CNV is not presently known. The amplitude of the CNV is correlated with heart rate deceleration. There are controversies in the literature with regards to heart rate during information processing. The orienting response produces a dual sympathetic-parasympathetic effect on heart rate. The cardiovascular response is sympathetically controlled and cholinergically mediated. Gross blood flow shifts occur during the initial phasic heart rate change. This will vary in magnitude with stimulus intensity due to parasympathetic inhibition. All attentional heart rate changes are under vagal regulation which act to restabilize the system. Heart response during vigilance follows this sequence: stimulus --> blood flow shift and respiration block --> deceleration until a response is made. Stroh (1971) refers to the role of the reticular formation in the general arousal response. Spinal nerves synapse with dorsal column nuclei whose axons synapse with the specific neurons (ventrobasal nuclei) of the thalamus. These neurons then activate well-defined functional areas of the cortex. This system to explain with a unidimensional theory of arousal involve the effects of noise, sleeplessness, alcohol, and drugs, time on the watch and these variables in combination. One can attempt to specify the specific effects of stressors on each arousal mechanism in a two mechanism theory or abandon this concept and develop separate theories of incentive, sleeplessness, etc.
Table 27. (continued).

<table>
<thead>
<tr>
<th>Vigilance Theory</th>
<th>Vigilance Decrement</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>(7) Continued</td>
<td>is a direct system that seems to be responsible for a phasic activation of specific regions of the cortex. These direct and specific sensory pathways do have connections with the reticular activating system (RAS) so that even though impulses take the direct route there is also a general activation effect that arises from stimulation of the RAS. An indirect pathway also exists that produces a tonic arousal effect in the cortex. These are afferent pathways that synapse with the RAS and then connect with nonspecific thalamic nuclei. This results in a general activation of the entire cortex and produces changes in alpha. The arousal response of the RAS blocks the cortical recruiting response of the nonspecific thalamic nuclei. The result is overactivation which may result in a loss of awareness and discrimination. Broadbent (1971) and Eysenck (1982) argue for the existence of two arousal systems. The two arousal systems are interactive; one is a compensatory, active, cognitive control system that monitors a passive arousal system and mobilizes extra effort and resources when necessary. Evidence exists for some form of two arousal systems. Considerable work has been done using the 5-choice serial reaction task, combined stressors and converging techniques. Arousal/</td>
<td></td>
</tr>
</tbody>
</table>
Vigilance Theory | Vigilance Decrement | Comments
--- | --- | ---
(7) Continued | activation theory is rapidly changing and it is not possible to identify a simple theory such as classical activation theory. Presently it is only possible to compare research results with a wide and diverse set of findings. Posner (1975) and Mackworth (1968) suggest that tasks with low event rates may lead to underarousal and increases in \( \beta \) whereas overarousal may lead to habituation and declines in \( \alpha \). | Filter theory can explain many of the results from Mackworth's clock tests. For example, the repeated presentation of a signal causes it to lose novelty thus decreasing performance. Broadbent (1971) predicts that response time and response time variability should increase with time on the task. Filter theory can explain effects such as observed in Jerison's research with fast versus slow event rates. Filter theory would predict worse performance with high event rates. Rest periods permit a signal to regain novelty which maintains performance. Similarly disruptions such as occur with telephone messages improve performance because novelty is regained. Filter theory also would predict that conspicuous signals receive priority; there is support for this notion. Filter theory would |
Table 27. (continued).

<table>
<thead>
<tr>
<th>Vigilance Theory</th>
<th>Vigilance Decrement</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>(8) Continued</td>
<td>In $d'$. High event rate tasks with transient signals occurring at unexpected times will decrease $d'$. Pigeon-holing links evidence states to category states and occurs further along in the information processing sequence. Increases in $\beta$ are thought to reflect changes in the rules that link evidence to response categories. Signal probability is thought to control $\beta$.</td>
<td>predict a performance decrement to be more likely with paced tasks than unpaced tasks. Results show decrements occur with both types of tasks. Filter theory would predict the vigilance decrement to be greater for tasks with multiple displays because of more timesharing. Jerison found less of a decrement with a three-clock task than a one-clock task. Filter theory is not unique in its predictions; habituation theory will predict equally well many of the research results.</td>
</tr>
<tr>
<td>(9) Expectancy theory (J. Deese, C.H. Baker)</td>
<td>Expectancy theory emphasizes the ability of an observer to form estimates of the occurrence of the next signal based on an averaging of previous information. The storage of information and ability to estimate time intervals is important to the theory. The observer's expectancy determines his performance level. Stroh (1971, p. 59) summarizes the expectancy position: (1) you are more likely to detect a signal if it occurs when you expect it to, (2) it is easier to determine when to expect a signal if the other signals have occurred at fairly regular and small intervals, and (3) when a signal occurs at a time when you don't expect it, you are more likely not to detect it. The vigilance intensity should help observers gain more knowledge about signals and thus form more accurate expectancies. The results do not support this notion. The human observer is not particularly accurate at time estimation which is necessary to forming expectancies. Jerison (1967) found that high event rates produce greater decrements than</td>
<td></td>
</tr>
</tbody>
</table>
Table 27. (continued).

<table>
<thead>
<tr>
<th>Vigilance Theory</th>
<th>Vigilance Decrement</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>(9) Continued</td>
<td>decrement occurs because of the &quot;vicious circle&quot; phenomenon. Observers do not estimate time intervals very well. Because of this signals to be detected occur at times when they don't expect them. As a result, signals are missed which alters the perceived signal structure which causes estimation and averaging to worsen even more. This cycle leads to the vigilance decrement.</td>
<td>low event rates. Signal frequency effects appear with only high event rates. Expectancy theory has some difficulty with these results—expectancies are not just based on a simple, temporal, averaging process. Expectancy theory explains particularly well the effects of experience with training task signal schedules but there are many effects not well explained by expectancy theory.</td>
</tr>
<tr>
<td>(10) Motivation theory (R. L. Smith, P. D. McCormack, F. Nachreiner)</td>
<td>Motivation is intrinsic to nearly all vigilance theories. Motivation is often called upon upon to account for unexplained variance or fluctuations in performance due to individual differences. Some researchers have made motivation central to the explanation of the vigilance decrement. Smith's (1966) theory is based on the interplay of monotony, boredom and motivation. He claims that most normal observers can detect 100% of signals for 1-2 hours; however, the vigilance decrement reflects individual differences, reward structures and willingness—not ability—to participate. Individuals can be classed as periodic participants or conscientious participants. Periodic participants do not observe at the limits of their capability whereas conscientious participants do. The vigilance decrement reflects the mix or</td>
<td>Like expectancy theory, motivation theory accounts for certain vigilance results. Studies such as cited by Nachreiner (1977) that manipulate pretask instructions are best explained by motivational theory. Other studies such as those dealing with KOR and incentives are explained in part by a motivation concept. Motivation theory like arousal theory and habituation theory can be considered as a theory emphasizing the state of the observer.</td>
</tr>
</tbody>
</table>


According to Smith, behavior can be influenced and modified by providing extrinsic motivation, e.g., KOR, rewards and punishment, kind of supervision, instructions, etc. Nachreiner (1977) has provided some support for this position. Subjects participated in tasks that they thought were related to job criteria; other subjects participated in what they were instructed was a research study. A vigilance occurred in the latter task but not the former; perceived operational relevance of a task has a significant influence on performance. McCormack performed a series of studies that examined the effects of ISI and KOR on RT. McCormack found that RT improved with KOR which McCormack interpreted as affecting motivation. KOR has a potent effect on performance—even false KOR improves performance but not as much as true KOR (Warm et al., 1974).

<table>
<thead>
<tr>
<th>Vigilance Theory</th>
<th>Vigilance Decrement</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>(10) Continued</td>
<td>periodic participators and conscientious participators in the experimental sample. According to Smith, behavior can be influenced and modified by providing extrinsic motivation, e.g., KOR, rewards and punishment, kind of supervision, instructions, etc. Nachreiner (1977) has provided some support for this position. Subjects participated in tasks that they thought were related to job criteria; other subjects participated in what they were instructed was a research study. A vigilance occurred in the latter task but not the former; perceived operational relevance of a task has a significant influence on performance. McCormack performed a series of studies that examined the effects of ISI and KOR on RT. McCormack found that RT improved with KOR which McCormack interpreted as affecting motivation. KOR has a potent effect on performance—even false KOR improves performance but not as much as true KOR (Warm et al., 1974).</td>
<td></td>
</tr>
</tbody>
</table>
Although the fourth phase of the evolution of vigilance research has not been achieved, namely the synthesis of findings, there are signs this phase is beginning to emerge. The work of Davies and Parasuraman (1982) represents a significant and noteworthy attempt at a synthesis of many vigilance findings. The approach taken by Davies and Parasuraman emphasizes an integration of research findings through the development of a vigilance task taxonomy, i.e., through a definition of the major vigilance task dimensions. The approach taken by Davies and Parasuraman is one that has been advocated and supported by this author throughout the present research (see Chapter 2).

The present research has had a broad-based focus. The dependent variables have included measures of response latency and eye, heart, and brain activity, i.e., behavioral, cortical and autonomis measures have been collected which cut across many theories. Although, most researchers in vigilance say that it is not possible at this time to test one theory against another it is perhaps useful to discuss the results of this research in terms of a few specific theories. The question is, "Which theory or theories are relevant to the present research?" If one evaluates the various theories of vigilance discussed in the research literature, one cannot help but notice the overlap of theories. For example, arousal theory is closely related to habituation theory. In fact, habituation theory is often referred to as inhibition-habituation-arousal theory. Habituation theory is also very much like filter theory but is described at a neurological level rather than a functional level. Expectancy theory assumes that an observer has a temporal model for future signals based on previous
experience with the signal schedule. Broadbent classified observing response theory as a subclass of expectancy theory. He asserted the observing response is similar to the concept of filtering except that filtering depends on the passage of time whereas observing responses depend on other variables. Most of the theories already described contain an element of expectancy theory. Broadbent's (1971) statement:

Broadbent (1958) divided the various theories of vigilance into four. As was clear at the time, these were not necessarily mutually exclusive, and it might well be that more than one of them was true . . . Without undue forcing, one can classify the work done on vigilance since 1958 into the same four groups . . . (pp. 20-21).

Broadbent's observation applies just as well in 1984 as it did in 1958 and 1971 with the exception that the original four theories should be augmented with more recent theories. An updated list is as follows:

1. Inhibition theory
2. Expectancy theory
3. Observing response theory/elicited observing rate re-sponse theory
4. Filter theory
5. Arousal theory
6. Habituation theory
7. Decision theory
8. Motivation theory

As was seen earlier, habituation theory is like filter theory and Jerison's observing response theory. Expectancy theory is like Holland's observing response theory, i.e., if a person knows when to observe a signal, performance should be better. One ends up with about
a handful of theories if the similarities of theories are considered: inhibition theory, expectancy theory, filter theory, arousal/habituation theory, decision theory, and motivation theory. Motivation theory is added to the list because it is in a class by itself. Other theories that have been previously discussed such as probability matching is a variation of expectancy theory. Strohs IAF theory is an amalgam of the above theories. Figure 43 classifies the various theories along a state-strategy dimension. The divisions between the various dimensions are blurred and will become more blurred as a synthesis emerges due to research.

Interpretation of Research Results in Terms of Theories

The research findings have been summarized and presented with comments about the validity of the results. The various theories have been described and discussed in terms of their validity. The present section discusses the present research findings in terms of their validity in relationship to existing theories.

Response Time and Intersignal Interval (ISI)

The intersignal interval (ISI) in vigilance research is completely described by specifying its distribution and association parameters. The mean and variance of the distribution are usually adequate to describe the ISI although some research studies report the actual sequence of signals used.

ISI is related to several other variables that are often treated as separate and independent in the literature when in fact they are
Vigilance Theories

State Theories
- Arousal theory
- Habituation theory
- Motivation theory

State-Strategy Theories
- Elicited observing rate response theory
- Signal detection theory
- Filter theory
- Inhibition theory

Strategy Theories
- Expectancy theory
- Observing response theory

Figure 43. Classification of Theories According to a State-Strategy Dimension
intimately related. Average signal interval defines signal rate (signal density and signal frequency are terms often used interchangeably with signal rate). The variance of the ISI defines signal regularity. If no variance exists a signal series is defined as completely regular. If the ISI has a large variance the signal series is irregular. Related to signal regularity is temporal uncertainty. The more variable the ISI the greater is the temporal uncertainty of a signal. (Spatial uncertainty defines a search task). The greater the signal frequency the less the temporal uncertainty of the signals; there is an inverse relationship between signal frequency and ISI. In some research tasks there are three distinct events that occur: signal events, non-signal events, and blank events. For example, Jerison and Pickett (1964) had two stimulus events; one a signal event and one a background event. A background event was the deflection of a light bar. A signal event was a slightly greater deflection of the light bar. The interval between events was a third type of event. In this case, it is appropriate to speak of stimulus density. Stimulus density is the frequency of occurrence of both signal events and background events. Signal probability is another term often referred to in the literature. The way in which signal probability is defined depends on whether the research task has two kinds of events or three kinds of events. Colquhoun (1961) used a task where a blank panel was presented; or a panel with six disks, one of which was more pale than the other five; or six disks, all the same hue. The latter two events, he termed "wanted signals" and "unwanted signals", respectively. In this
instance signal probability was defined in terms of conditional probabilities: Pr (wanted signal / a stimulus event occurs). A stimulus event is the occurrence of a wanted signal, unwanted signal or blank panel. When three events occur, it is appropriate to speak of inter-event intervals and intersignal intervals. These intervals need to be considered when defining signal or stimulus probabilities.

Signal duration and ISI also have a relationship. Signal duration is generally thought to be related to signal conspicuity; the longer a signal remains the higher its detection probability. For example, in tasks where only signals or background stimulation exists such as the present research, the duration of the stream of background events between signals defines the ISI. If the duration of signals are lengthened or shortened the ISI also changes; consequently signal duration is related to ISI.

The point of the previous discussion is to illustrate the inter-relationships that may exist between intersignal interval (ISI) and signal rate, stimulus density, uncertainty, signal probability, and signal duration. Moreover, it should be mentioned that studies such as Baker (1963a) where signal duration was studied using a clock test with a 2 hour task and a signal rate of 20 per hour are confounded with intersignal interval. Often times the research literature does not mention this kind of confounding that can occur with ISI and variations in signal and task characteristics, e.g. session duration. (The situation is much like that observed by Heisenberg--you cannot simultaneously measure the velocity and position of an atom--here you cannot vary
signal duration and keep session length constant without changing ISI. Signal rate would also be affected).

The effects of intersignal interval length has been reported from at least four theoretical viewpoints: classical conditioning (inhibition) theory, operant conditioning theory, expectancy theory, and Jerison's elicited observing rate response theory. Expectancy theory, however, has the greatest theoretical investment in ISI effects.

Deese (1955) wrote one of the earliest articles that compared reinforcement theory with expectancy theory. The paper has been an influential and theoretical benchmark in the development and growth of vigilance theory. Deese's paper is often cited and quoted. Deese was an exponent of expectancy theory and viewed expectancy theory as follows: "... expectancy should be low immediately after a signal, should increase as the mean intersignal interval is approached, and finally should become quite high as the intersignal interval grows beyond the mean (Deese 1955, p. 363)."

Deese's paper was certainly a contribution to the literature in the sense that it compared the current theories of the day but was obfuscating in several other ways because of a lack of precision in defining and discussing terms. Deese's paper can be faulted on several grounds:

1. Deese discussed reinforcement theory as a single theory. Reinforcement theory can be either classical conditioning (inhibition) theory or operant conditioning theory (observing response theory); the predictions from these theories are opposite. It is difficult to determine which theory he is referring to at times because of the interchangeability of theories on his part.
2. Deese presented the results of his research in the paper which showed the percent of signals to be an increasing function of signal rate. This he interpreted as support of expectancy theory. However, the results are just the opposite of what his earlier statement said expectancy theory should predict.

Many research studies, such as Deese's use percent detections as a criterion measure and do not report response latency. One would intuitively expect percent detections to be high when response latencies are low—in fact, Boulter and Adams (1963) state: "... the reasonable assumption [is] that [the] probability of detection and response latency are inversely related, so that short latency means high expectancy... (p. 208)." Jenkins (1958) studied response times and percent detections using a DC voltmeter needle deflection task and signal rates of 7.5, 30, 60, and 480 signals per hour. He found that percent detections declined with increases in response times for all signal rates.

What do the various theories predict with regards to response latency and ISI? The following summarizes the various positions:

1. Inhibition theory. Inhibition theory or classical conditioning theory predicts that performance will improve as the intersignal interval increases because inhibition to respond has more of an opportunity to dissipate. Mackworth (1950) did find that percent misses declined with ISI. Inhibition theory would predict that response times decrease with increasing ISI.

2. Operant conditioning or observing response theory. Operant theory takes the position that signal detections are reinforcing and the detections are reinforcing and the shorter the ISI, the better performance. This theory would predict detections would decrease with ISI and response latency would increase with increases in ISI.
3. **Expectancy theory (Deese).** Deese's version of expectancy theory assumes that observers are performing an averaging process based on ISIs previously experienced. Expectancy immediately after the detection of a signal (short ISI) would be low and increase up to the mean ISI and beyond. Presumably, Deese means that expectancy for a signal increases until another signal is presented for detection. However, one would not expect an observer's expectancy for a signal to keep growing indefinitely beyond the mean ISI to which he is accustomed. Moreover, if an observer has experienced a high signal rate, one would not think observer expectancy to be low immediately after detection of a signal. Expectancy theory is weak and fuzzy with regards to its explanation of how expectancies are formed and how observers modify their expectancies in the face of changing information. Deese's expectancy theory does not differ in its predictions from classical conditioning (inhibition) theory.

4. **Expectancy theory (Baker).** Baker was not satisfied with Deese's version of expectancy theory and modified it such that an observer's expectancy declined beyond the mean ISI. Baker viewed expectancies (and performance as measured by percent detections) as increasing as the mean ISI is approached and then flattening out, i.e. there is an interval of uncertainty, and then a gradual decline occurs. The expectancy curve is an inverted, somewhat u-shaped curve.

5. **Elicited observing response rate theory (Jerison).** Jerison (1967) reported a study that examined response latency and percent detections as a function of signal probability and event rate. He used the following conditions: 3/360, 15/360, 15/1800, and 75/1800 signals and events per hour. He found a straight function with no slope or a slightly increasing slope with detections at about 80% for the low event rate. At the condition 15/1800 signals and events per hour, performance was about 30% detections and increased to about 50% for the condition 75/1800, i.e. a significant event effect was found and a slight signal effect for high event rates. Jerison found response time averaged 1.8 seconds for the low-event rate (360/hour) and 1.0 seconds for the high event rate (1800/hour). He interpreted this effect to mean that good detection performance was associated with long latencies and poor performance with short latencies. According to Jerison, the observer had an inter-stimulus interval of about 10 seconds at the low event rate--he used the longer interstimulus interval to make a more leisurely response whereas for the high event rate events occurred on average of every 2.4 seconds.
In other words, with less time available, he took less time, i.e. he adopted his response to the time available. The monitor's performance appears to be some function of attentional costs per unit of the time available to observe and respond. Perhaps he adjusts the quality of observing and responding to fit the time available. Jerison's results could also be explained in terms of arousal, i.e. the observing response mechanisms is being overloaded which results in time stress.

The results of the present research and predictions of various theories are summarized in Figure 44 in relation to response time and intersignal or interstimulus interval. The graph suggests that elicited observing response rate theory (Jerison's theory), Deese's version of expectancy theory, or operant conditioning theory predict the results from the present research, i.e. response times increase with increasing ISIs. The remaining discussion will focus on the empirical evidence in this research and other research available to aid in discriminating between these the various theories.

Earlier in Chapter 5 when the various theories were reviewed, Jerison's theory was thought not be applicable because his research was based on clearly defined observation intervals and high background event rates. In the present research, signal events are clearly defined but events that occur in the ISI are characterized by a "sameness," particularly when the ISI is large. Since non-signal events are not clearly defined, it is not clear how subjects perceived event rates. There is a temptation to define event rate as it was defined when designing the research task, i.e. each five second interval was either a background event or a signal event. The event rate would then be 240 stimulus events/20 minutes or 720 stimulus events/hour. This event
Present research, Jerison's theory.

Baker's expectancy theory

Interval of uncertainty

Mean ISI

Inhibition theory, Deese's expectancy theory.

Figure 44. Summary of Research Results and Theoretical Predictions for Response Time as a Function of Intersignal Interval.
rate is in line with the event rates used by Jerison. However, one must exercise caution in interpretation because the background events were not clearly delineated as they were in Jerison's task.

The task used in this research had an element of search and in that sense required continuous monitoring. The continuous nature of the attentional demands would not make it so different from Jerison's high event rate task. It is not unusual that Jerison's theory predicts what expectancy theory predicts since Jerison incorporated expectancy into his theory (Jerison, 1967). Broadbent (1971) views observing response theory (both Jerison's and Holland's theory) as a part of expectancy theory. Presumably knowing when to observe provides more reinforcement (Holland's theory) and reduces the attentional cost of observing (Jerison's theory). It may not be possible to discriminate between the three theories because each includes an expectancy construct and they may all be basically the same (as Broadbent suggested).

One problem that exists with the theories discussed in regards to intersignal interval effects in that of missed signals and false alarms. Inhibition theory and operant conditioning theory predictions are based on the amount of elapsed time since the last detected signal. If a signal is missed, the ISI lengthens. If a signal is erroneously inserted the ISI shortens. The various theories do not address the effects of errors on perceived ISIs. The predictions of expectancy theory are based on mean ISIs rather than individual ISIs as are inhibition theory and operant conditioning theory. Little change in expectancy would be anticipated as the result of missed signals or false
alarms if the monitor has substantial experience with the task and perceived average ISI has stabilized. If expectancies are based on a small sample size, such as occurs at the beginning of a task, missed signal and false alarms may have a significant effect on expectancy. The present research has a low miss and false alarm rate and although some effect probably occurs, the exact nature of the effects are not known.

Another aspect of the present research should be discussed and that has to do with the experience that observers receive as the result of signal schedules with different degrees of uncertainty. Runs 1 and 2 have a constant ISI, i.e. no temporal uncertainty in the signal schedule. Although no temporal uncertainty exists with the signal schedules for runs 1 and 2, the observer may perceive temporal uncertainty in the signal schedule. Presumably, this uncertainty would be reflected in run means and variances. Expectancy theory would suggest that expectancies should be more easily formed with a regular signal series than an irregular series. Average run response times and variability should be less than for the more irregular runs, i.e. 3, 4, 5, or 6. Response time means and variances for long intervals for runs 1 and 2 were not significantly different than means and variances for runs 3, 4, 5, and 6. This prediction is not obvious from the data. Several explanations exist for this: (1) the observer does not have enough experience with the few signals (six) in runs 1 and 2 to form stable expectancies, (2) although temporal uncertainty may not exist in the signal schedule, the observer may perceive temporal uncertainty in
the signal schedule which makes it difficult to form precise estimates of the average ISI, (3) individual differences may "wash out" any kind of run effects, or (4) expectancy theory is not an appropriate explanation of the data.

An expectancy theory of vigilance is predicted on the assumption that a monitor's performance is based upon his ability to predict the occurrence of future signals based on past events. Closely related to expectancy theory is a body of research literature based on protensity. In general observers can estimate short intervals better than long intervals. Stroh (1971, pp. 77-78) reviewed the literature on time estimation and concluded: "... negative time errors tend to occur more frequently than do positive time errors, and that this tendency increases with an increase in the size of interval between the two stimuli." Negative time error means the tendency to overestimate time intervals and positive time error is the tendency to underestimate them. Observers tend to underestimate short ISIs and overestimate longer intervals although there is disagreement about the size interval where times error changes from positive to negative, i.e. where time error is zero (Stroh reports that somewhere between 1.5 and 12 seconds is the change over point.) One cannot conclude from the results of the present research that observers underestimate or overestimate ISIs; however, there is support for this notion in the data--response times do grow with ISIs.

Johnston, Howell, and Goldstein (1966) studied signal frequency (60 and 150 per 100 minute session), stimulus density (4, 8, 16, or 32
stimuli within an 8 x 8 matrix display on a CRT on each trial). The researchers found mean detection latency was a U-shaped function of intersignal interval in the range of 70 to 130 seconds and had a significant quadratic component. Boulter and Adams (1963) examined response latency as a function of ISI and temporal uncertainty. For the most variable ISIs, the authors found response latency to increase with ISI. Response latency for the medium uncertainty condition was high for short ISIs and low for long ISIs. Only one mean response latency point was available for the no uncertainty condition but it did coincide with the point defined by response latency and the mean ISI for the medium uncertainty condition (mean ISI was 220 seconds for both conditions). The results suggest that the medium uncertainty results follow predictions of inhibition theory whereas the high uncertainty condition results are very similar to results obtained in the present research (NU = 220 seconds; MU = 2 of 120 seconds, six of 220 sec. and four of 270 sec.; HU = 15, 15, 30, 30, 30, 60, 120, 120, 300, 420, 600, and 900 sec.). Dardano (1952) found results similar to those reported by Johnston et al., and Boulter and Adams. Dardano studies response latency as a function of interstimulus interval under low, medium, and high uncertainty. The low uncertainty condition used the following: 50, 55, 60, 65, and 70 seconds. The medium uncertainty condition had ISIs equal to 30, 45, 60, 75, and 90 seconds; the high uncertainty condition used ISIs equal to 10, 35, 60, 85, and 110 seconds. Dardano found that response latencies declined with increases in ISI for the low uncertainty condition. Response time appeared to be a u-shaped
curve for the intermediate uncertainty condition; Dardano found a significant non-linear trend. The high uncertainty condition showed response times to increase linearly with increasing ISIs. The curvilinear relationships found by Johnson et al. and Dardano follow the predictions of Baker (1963c) whereas the functions that show response latency increases with increases in ISI follow the predictions of Deese. The functions that show response time linearly decreasing with increasing ISIs or interstimulus intervals follow the predictions of inhibition theory.

It is not clear how subjects perceived signal rate, signal regularity, intersignal intervals and signal intensity in the present research since the number of signals and modality of signals varied with the particular run. The average ISI for runs 1 and 2 was 315 seconds (S.D. = 0 seconds). The average intersignal interval for runs 3 and 5 (artificial signal runs V1 and V2) was 187 seconds with a standard deviation of approximately 145 seconds. It is not clear how to define the ISI in runs with auditory warning signals but if it can be defined as the interval between signals regardless of modality, then the average ISI is also 187 seconds (S.D. = 145 seconds). There is no ISI equal to a mean ISI of 187 seconds in the present research. The ISI closest to the mean ISI of 187 seconds is an ISI = 250 seconds.

Boulter and Adams (1963) used a mean ISI = 220 seconds and the standard deviations of ISI for the high, medium, and low uncertainty conditions were 285, 52, and 0 respectively. Johnston, Howell, and Goldstein (1966) report mean ISIs of 40 sec. and 100 sec. normally
distributed but do not report the variance of the p.d.f. Dardano (1962) reported an average signal rate of 60/hour and ISI standard deviations of 7.9, 23.7, and 39.5 seconds for low, intermediate, and high uncertainty conditions. The present research ISI means and standard deviations were similar to those used by Boulter and Adams (1963) for their high uncertainty condition. Boulter and Adams found that response times increased with increasing intersignal intervals when temporal uncertainty was high. Dardano (1962) defined high uncertainty lower than Boulter and Adams but found response times increased linearly for the high uncertainty condition.

The following conclusions summarize the effects of intersignal interval on response times:

1. Response time was found to be a linearly increasing function of intersignal interval in the present research.

2. Increasing response time with ISI is consistent with predictions of expectancy theory and observing response theory. Observing response theories can be viewed as theories related to expectancy theory.

3. The results observed in this experiment are consistent with results found by Boulter and Adams (1963) and Dardano (1962) for performance (response time) and intersignal intervals while the signal schedule contains high uncertainty.

4. It appears that inhibition theory may explain response time--ISI relationships under low uncertainty, Baker's expectancy theory may predict response times under medium uncertainty, and Deese's expectancy theory may explain the response time--ISI relationships under high uncertainty.

5. More research in this area may be justified to establish more precise relationships between uncertainty, response times, and temporal uncertainty.
6. Arousal theory could also explain the finding that response times increased as ISIs increased. Although Jerison explained his background event rate results in terms of attentional costs, he allowed the possible explanation provided by arousal theory. The present research also does not dismiss arousal theory as a possible explanation of results. Jerison incorporated an aspect of arousal into his theory. Expectancy theory and Holland's observing response theory, however, assume the observer is in an aroused state and actively observing. Some support for an arousal component does exist in the present research in the eye movement and behavioral response time data. The missed signals in the research were not fixated in most cases. Missed signals were also accomplished by nodding and eye drooping behavior which was observed on video tapes of subjects' eye movements. Because of this observed behavior, it is believed that two subjects were in a low activation state during a few of the runs. This aspect will be further discussed.

7. The time courses of events needs to be considered with regards to the various theories. Specific time relationships are not shown in Figure 44; the figure is not drawn to a time scale. The figure only shows functional forms and not specific and detailed temporal relationships. Table 28 summarizes research and theories on the basis of ISI and response times. With the exception of Mackworth (1950), the present research used the widest range of ISIs. Jerison and Baker found differential response time effects with ISIs of less than 10 seconds. The present research did not find significant effects unless the full range of ISIs were considered. Another point should be mentioned and that is Jerison used event rates of approximately 0.5 events per second. Psychological refractory period probably does not impose a limit on Jerison's studies but temporal uncertainty and response complexity may increase the refractory period and put a limit on some high-event rate vigilance studies (it is questionable whether these studies should be called vigilance studies). A response time decrement occurs even when signals are temporally certain at high signal rates (2-3 signals every second). Expectancy does not seem to govern performance in this situation and Jerison's attentional costs or Broadbent's filtering or "pigeon holing" may be an explanation for performance decrements. The present research task differs from those used by other researchers listed in the table; it had elements both spatial and temporal uncertainty; the observers did not know which meter would change or when.
<table>
<thead>
<tr>
<th>Research or Theory</th>
<th>Signal Rate or Event Rate</th>
<th>Range of ISIs or Interstimulus Intervals</th>
<th>Average or Range Response Times.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Present research</td>
<td>3 or 5 signals per 20 minutes (9-15 signals/hour, 720 events/hour (?))</td>
<td>5-315 seconds (ISIs)</td>
<td>0.54-5.68 seconds RT = 2.29 seconds</td>
</tr>
<tr>
<td>2. Jerison (1967)</td>
<td>3-75 signals/hour, 360-1800 events/hour</td>
<td>2.4-10 seconds. (interstimulus intervals)</td>
<td>1.0-1.8 seconds</td>
</tr>
<tr>
<td>3. Baker (1963)</td>
<td>Ss received 20 signals, 10 seconds apart and a 21st signal at 2, 5, 20, 25, or 30 seconds</td>
<td>2, 5, 20, 25, 30 seconds (ISI Ss were most familiar with was 10 seconds)</td>
<td>0.40-0.64 seconds about 0.42 seconds</td>
</tr>
<tr>
<td>4. Deese (1955)</td>
<td>10, 20, 30, 40 targets per hour.</td>
<td>not reported</td>
<td>RT not used. % D reported</td>
</tr>
<tr>
<td>5. Mackworth (1950)</td>
<td>12 signals/30 minutes (24 signals/hour)</td>
<td>45-600 seconds (ISIs)</td>
<td>0.5-7.9 seconds RT about 1.2 seconds</td>
</tr>
<tr>
<td>6. Psychological refractory period-Creamer (1963)</td>
<td>A decrement in RT occurs for the second of two signals as a function of ISI. The decrement occurs out to an ISI of about 300 ms. Temporal uncertainty increases RT 50-100 MSEC.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 28. Comparisons of Different Research Results Based on ISI.
The task is the simplest of search tasks (previously, it has been mentioned that most search tasks don't exhibit a vigilance decrement, i.e. response time increment or percent detections decline over time.) Intersignal or interstimulus intervals and associated parameters may permit one to discriminate between theories; this is a direction for future research.

Interpretation of Physiological Measures

The use of physiological measures in vigilance research generally assume an arousal hypothesis although this is not always true. Occasionally, conditioning models are used in conjunction with a vigilance task. For example, Beatty, et al., (1974) used operant conditioning techniques to shape EEG 0 rhythms in a radar monitoring task. The use of such operant techniques falls in the category of biofeedback control of central processes. Some studies make use of evoked potentials such as CNV. The CNV generally occurs when a warning signal is present and these studies are classified as studies of expectancy.

Jerison (1977) developed a "truth table" to help sort out concepts with regards to attention. His table is reproduced and appears as Table 29. Jerison makes a plausible argument for separation of sustained attention and selective attention. Selective attention has been associated by Broadbent (1971) with vigilance. Broadbent (1958) initially attributed the vigilance decrement to failures of an absolute filter. Later, he modified this view of an absolute filter and incorporated into his attentional model the notion of a relative filter and pigeon holing to explain the vigilance decrement. Broadbent (1971) also likened his model to Jerison's observing response model. Clearly
Table 29. Truth Table to Diagnose and Distinguish Among Phenomena Related to Attention.

<table>
<thead>
<tr>
<th>Phenomenon</th>
<th>Decrement* Time Constant</th>
<th>Correlation With IQ</th>
<th>Effect of High Event Rate</th>
<th>CNV Present?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vigilance (Sustained attention)</td>
<td>5 minutes</td>
<td>No</td>
<td>Poor Performance</td>
<td>No</td>
</tr>
<tr>
<td>Selective Attention</td>
<td>1 hour or more</td>
<td>Yes</td>
<td>Good Performance (fore-period effect)</td>
<td>Yes</td>
</tr>
<tr>
<td>Arousal</td>
<td>indefinite</td>
<td>?No</td>
<td>High arousal</td>
<td>?</td>
</tr>
<tr>
<td>Habituation</td>
<td>1 second or less</td>
<td>?No</td>
<td>Rapid habituation</td>
<td>?</td>
</tr>
</tbody>
</table>

*Time constant to be interpreted as a kind of chronaxie for a decrement function—the time required for half of the eventual decrement to occur.

Note: Many entries in this table are provisional, especially with respect to contingent negative variation (CNV) or slow DC change in EEG baseline level. The point is to suggest a way to think about the categories used to describe vigilance (sustained attention) as opposed to the other phenomena listed. Note that arousal is shown as unitary, though many authors (e.g., Lacey, 1967) have clearly shown that several kinds of arousal (behavioral, automatic, EEG) may behave differently under diagnosis by this kind of "truth table."

Jerison does not equate selective attention processes with sustained attention processes. Jerison also makes definite distinctions between habituation, selective attention and sustained attention. The only theory that seems compatible with vigilance, according to Jerison, is arousal theory. If Jerison's "truth table" is accurate, sustained attention consists of possibly two phenomena: vigilance and arousal. Jerison's truth table could be expanded to be more discriminating of various phenomena. In addition, other theories could be added such as SDT or expectancy.

Previously, it has been mentioned that one either adopts a unitary concept of arousal (which has its difficulties) or is faced with a morass of often contradictory physiological findings that are not integrated into theory. The position taken here is that arousal theory is a necessary part of vigilance theory; otherwise many research findings would be difficult to explain. However, the exact nature and form of an arousal component needs to be worked out.

**EEG Measures and Arousal Theory**

Three significant EEG findings were observed in this research: (1) alpha power increased in the presence of an auditory warning signal but not in the presence of a visual artificial signal, (2) EEG alpha had a significant positive correlation with response time in the 5 second interval during which a signal was presented, and (3) EEG delta was positively correlated with response time in the 5 second interval prior to a signal.
The finding that EEG alpha power was greater for the auditory warning signal condition than it was for the visual signal condition was not expected. The auditory warning signal restores observing behavior when eye movement runs are examined. One would expect that a novel auditory signal would elicit the orientation reaction (Lynn, 1966) and faster and lower amplitude activity. Traditionally, an alerted state is associated with blocking or desynchronization of alpha rhythms. Recent evidence shows that cortical activation may not reflect a single state but may reflect several states depending upon task conditions and subject characteristics.

There is some question as to whether alpha should be recorded with the eyes opened or closed. Lippold (1970) has suggested that alpha activity is an artifact generated by tremor of the eyes. Shagrass (1972) also supports this assertion although he suggests that recording techniques and instrumentation can eliminate these artifacts. Herberg (1958), on the other side, could find no relationship between alpha measures and eye movement scores. Shaw, Foley, and Blowers (1970) declined with time on the task. Daniel (1967) found that a measure of alpha abundance increased with time on the watch which paralleled increases in response time and mean errors. Alpha seemed to index arousal but did not differentiate between errors and correct responses. Daniel had subjects monitor with eyes open. He interpreted his EEG results in terms of lowered arousal. O'Hanlon and Beatty (1977) claim:

alert individuals show a relatively high incidence of alpha activity with eyes closed and less as they become drowsy,
when alpha incidence and amplitude diminish as theta appears. On the other hand, alert individuals with eyes open show a preponderance of beta (14-30 Hz) activity while alert, followed by increasing alpha activity as they become drowsy (p. 191).

Gale (1977, p. 268) reviewed the EEG alpha literature and reported it to be confusing:

Alpha index decreases with time or increases with time, while theta activity, normally observed in drowsiness and associated with hypogogic imagery, is greater before good trials than bad trials. Pre-signal alpha is said to affect performance or have no effect. EEG activation in response to stimuli may take the form of an alpha blockade, reduction in amplitude, and desynchronization or, depending on resting level and time in task, take the form of alpha augmentation.

Despite the above quote, Gale is supportive of the EEG as a measure of activation. According to Gale, the EEG is as good as an index of arousal as heart rate, EDA, slow potentials, etc. He claims that tasks which contain either a large short-term memory component or response competition elements are likely to produce results in conflict with the literature. Such tasks call for lowered arousal if performance is to be successful whereas usually superior performance in vigilance tasks is associated with heightened arousal. Gale (1977) performed an experiment where he studied EEG as a function of signal rate and time on the watch. He found that theta and alpha in the range of 7.5-9.5 Hz (low alpha) increased with time while mean dominant frequency reduced. Gale interpreted this to mean lowered arousal. Stroh (1971) claims that open eyes with increases in alpha means decreased arousal.

Mackworth (1960) says the arousal response to visual stimuli usually results in the blocking of alpha rhythms; the opposite changes
may occur, especially if the subject is drowsy and the stimulus is auditory. The present research results of increased alpha with the presence of an auditory signal is consistent with Mackworth's contention of heightened arousal.

**Interpretation of Eye Movement Results**

Eye movement results confirmed initial hypotheses. The following results were found:

1. Significant positive correlations of dwell time with time.
2. Significant negative correlations of activity measures with time.
3. Significant negative correlations of dwell time and activity variability with time.
4. Significant mean dwell time and activity effects.
5. Significant decreased activity with time.
6. Signal effects; signals tend to increase average dwell time and decrease average activity.
7. Most missed signals were not fixated—observers were obviously the wrong meter too long.

The above results are consistent with an arousal hypothesis—dwell time tends to increase and activity decreases with time on the task. There is also a tendency for both dwell time and activity measure variability to decrease with time on the task. Signals have an alerting effect; this is evidenced by increased dwell time and decreased activity. Observers spend longer fixating on meters when a signal is present. This may represent a confirmatory action on the part of the monitor. Increased fixation time and activity measures are
complementary, i.e. one would expect changes in one to also be reflected in the other. Short brief dwell times should be accompanied by increased activity—less dwell time means more time spent in transition; more dwell time means less time spent in transition or scanning.

There are few models of eye movement behavior. Senders, et al., (1966) and Senders (1970, 1973) describe a visual sampling model that assumes the observer is a single channel processor of information. The observer samples a large number of information displays and the information is held in a queue until the monitor can further process the information. Senders (1979) outlines some of the dimensions of a mental workload model that considers switching time between visual inputs. Brown (1979) Howarth and Bloomfield (1971), Bloomfield (1975), Bloomfield (1973), Williams (1973, 1965, 1966) have developed models of visual search. Vigilance theories have been directed, in the main, at describing the performance decrement observed in many monitoring tasks. Previously, it has been mentioned that monitoring tasks with a search component are not as likely to show a performance decrement as tasks without spatial uncertainty. This conclusion, however, is debatable. Jerison and Wallis (1957) studied a three-clock monitoring task and found no decrement in performance although performance was considerably lower than for a one-clock task. Wiener (1964) however, reported that when signal density was controlled, performance is equivalent in multiple display. Monitoring tasks to that of single display monitoring tasks. The search models mentioned above e.g. Williams (1966) and Bloomfield (1973), often model the probability of
target detection as a function of time available for observation in the form of an exponential function—the probability of detection increases and these asymptotes towards 1. These models usually consider task variables such as information input rate, search time available, stimulus density, display size, target conspicuity but seldom consider changes of state of the observer over time. To this writer's knowledge there has been little integration of search and scan models with vigilance performance. Because there are few models or theories of eye movement behavior in relation to vigilance performance, one is faced with comparing research results with those obtained by other researchers—unfortunately there have not been many eye movement studies performed from a vigilance paradigm. For example, Mackie (1977) edited the published proceedings of a vigilance symposium where over thirty studies were published; only one involved eye movements.

Meyer and Liebl (1976) claim evidence for reduced arousal in an eye movement study. Pilots and weapons systems operators performed a 30 minute visual monitoring task. The researchers found that saccadic latencies and reaction times increased over time and that maximal angular velocities of saccadic eye movements decreased over time. Their results are consistent with the finding in the present research that eye activity decreases with time on the task.

Schroeder and Holland (1968) reported that eye movements, which they equated to observing responses, declined with decreases in signal rate. Eye movements and detections also declined with time on the task. The authors also reported that there was a greater tendency to
fixate a signal without reporting it the longer the time on the task. Although, the authors were interested in determining whether signals affected observing behavior, it cannot go unnoticed that they reported a vigilance decrement which was paralleled by declines in eye movement activity. No theoretical explanation for the decrement was offered. The declines in eye movement activity reported by Schroeder and Holland could be interpreted as due to decreased arousal.

The eye is always in motion. Even when the eye fixates on a target, the eye produces small movements which are necessary for perception. The present research recorded only gross activity, i.e. transitions from one display segment to another and dwell time in a display segment. Minature eye movements or fixation eye movements consist of tremor, slow drifts, rapid flicks or microsaccades, and irregular movements. These movements are considered to be involuntary as opposed to the large amplitude saccadic eye movements which are voluntary and move the eyes from point of fixation to point of fixation. The recording device used in the present research had a capability for recording minature eye movements but this capability was given up so that computer storage capacity was not exceeded. Gaarder (1966) discovered a unique fine eye movement pattern during a unique fine eye movement pattern during inattention. This oscillatory pattern was highly reliable during inattention and quite different from attentive fixation patterns. Future monitoring research should attempt to record these fine eye movements because they may be related to missed signals and arousal state. Increased frequency of this pattern over time may suggest declining arousal.
Recording a continuous record of eye movements via a limbus tracking monitor such as was used in the present research can provide an additional benefit. A continuous record would permit an analysis such as was performed by McDowell and Rockwell (1978) on the eye movements of five drivers. McDowell and Rockwell performed a spectral analysis of horizontal and vertical eye movement signals. They found shifts in power and frequency that were associated with different speeds and road geometry. Analysis similar to those done in this research for EKG, i.e. autocorrelation and spectral analysis and the cross-correlation, phase angle, and squared coherency analyses done with eye movements and heart rate should be performed. However, this should be done with a continuous eye movement signal rather than an averaged signal aggregated across a wide time interval.

More concern should also be given to the recording of blinks and the blink rate. Hall and Cusak (1972) reviewed the literature on voluntary eye movements and blinking. They suggested that blink rate may index arousal states such as boredom and excitement. Hall and Cusak's hypothesis is that blink rate is related to arousal by an inverted u-shaped curve such as is common to a unitary arousal theory. They also suggest the hypothesis that blink rate is low for processes requiring outwardly directed attention and increases for processes requiring inwardly directed attention. Kahneman (1973) also supports this idea. He says that saccadic rates are very high for thinking but inhibited during attentive listening.

Another area that needs to be considered in eye movement measurement is vigilance task is pupillometry. Kahneman believes that pupil
size is the best indicator of sympathetic activity. Proper instrumentation would allow one to record pupillary responses and relate changes to task and display variables. Simpson and Hale (1969) found increases in pupil size that paralleled the time at which decisions were made in a two alternative decision-making task. Bradshaw (1967) has investigated pupil size as a measure of arousal. Pupil dilation was found to be greater for difficult cognitive tasks than for easy tasks. Woodmansee (1966) has reviewed the problems of pupillographic experiments. Despite problems he does agree that pupil size decreases with decreasing arousal. The largest problem with the pupillary response is that of high variability. Goldwater (1972) has discussed the psychological significance of the pupillary response. Goldwater concludes that pupil diameter is an effective measure of autonomic activity albeit subject to a number of control considerations. Nunally, Knott, Duchnowski, and Parker (1967) performed an experiment that examined pupil size in response to semi-nude pictures, gunshot anticipation, auditory stimuli, novel stimuli and muscle strain. Significant increases in pupil size were found for all five arousing stimuli. Koff, Elman, and Wong (1975) have developed a bibliography of pupillary response literature. This bibliography addresses pupillary response in regards to vigilance as well as other types of research.

The present research found results similar to those of Mackworth, Kaplan, and Metlay (1964), i.e. most missed signals were not fixated (one out of five missed signals indicated the observer was looking
without seeing--all others were not fixated). A vigilance task with a one or two dial display was used to examine eye movements. Mackworth, et al. found that in contrast to the one dial monitoring task, most missed signals were not fixated at all. In addition, detection probability for the two dial case was approximately half that for the one dial case. Mackworth et al. found a significant dial effect but not a signal rate effect in an ANOVA for detection probabilities. Also no dial x rate effect was present. The presence of a dial effect can be taken as evidence for an arousal hypothesis, i.e. increased loading decreases performance.

The conclusions reached concerning eye movements are: (1) eye movement results are consistent with activation theory, (2) future research should aim at more precisely defining the visual response over time and in relation to stimulation by continuously recording voluntary eye movements and minature eye movements and by adding such measures as pupillary response and blink rate; (3) the line of research began here should be continued, i.e. cross-correlation analysis of eye movement measures with other measures--Lacey and others have reported low correlations of activation measures and this may be due to a different time course for activation measures--cross-correlation studies may help define the time relationships (lags and leads) between measures; and (4) future research should be aimed at developing predictive models of eye movement behavior--little has been done in this area.
EKG Measures

Mean heart rate was found to be greater for the signal condition than for the nonsignal condition. During the experiment, considerable difficulty was experienced with the heart monitoring device. Much of the data had to be excluded from the analysis. Sixty percent of the remaining data was from one subject. Although the data is not representative of the other subjects, between subject variability is low. Whether or not EKG, heart rate and heart rate variability are suitable physiological measures of activation states and are sensitive to task and environmental changes remains a controversy. A unified theory does not exist to predict heart signal changes. Duffy (1962) has suggested that heart rate and blood pressure tend to increase with increases in the intensity or significance of stimulation. The results of the present research tend to support Duffy's contention. However, the heart signals were averaged over 30 second periods and phasic responses of the heart tend to be smoothed and less apparent. Differences that show up tend to be tonic responses.

Lacey (1967, 1974) have advanced the sensory intake-rejection hypothesis; environmental rejection or attention to internal events results in phasic HR increases while environmental intake, i.e. attention to external events results in phasic HR decreases. The brain and heart make-up a feedback system. As heart rate and blood pressure changes, baroreceptors detect these changes and signal the brain; this signal results in either cortical activation or deactivation. The Laceys systematically examined the notion of global arousal such as put
forth by Duffy (1962, 1972). Their concept of directional fractionalization suggests that different fractions of the total somatic response may occur in opposite directions. Environmental intake, i.e. outward attention to the environment results in HR and BP, decreases, inhibition of eye movements, faster RTs and cortical deactivation, i.e. no alpha block. HR, BP, and cortical deactivation occur quickly - within the cardiac cycle.

Another position has been advanced concerning cardiac-somatic coupling is that of Obrist (1981). Whereas the Lacey's baroreceptor hypothesis relates heart and brain activity, Obrist favors a theory of heart rate changes based on metabolic demand. HR is a peripheral measure of activation rather than a cause of cortical and somatic activation. Obrist, et al. (1970) tested the Lacey idea of brain-cardiac coupling by using atropine to block the cardiac response. Obrist et al. reasoned that if faster RTs are caused by HR decreases, blocking the cardiac response should lead to longer RTs. Obrist et al. found support for their hypothesis. The Obrist cardiac-somatic hypothesis suggests that HR changes are parasympathetically mediated; HR changes in parallel with somatomotor activity much like it does with exercise. Obrist et al. used a classical conditioning paradigm. Obrist (1981) reports EMG, HR, and eye movement changes with the onset of an UCS. All of the measures follow an inverted U-shaped curve much like that suggested by Duffy.

Much of the work of psychophysiologyists over the last 10-15 years has centered around testing the hypotheses of Obrist and Lacey. Another
line of research uses heart rate variability as a measure associated with task demands. O'Hanlon (1972) found that HRV increased with driving time and decreased after realerting events. O'Hanlon (1977) found decreases in mean HR and increases in HRV with time.

The results of the present research are in line with a unidimensional theory of activation. A fine-grain analysis of cardiac measures was not performed. Consequently, phasic changes were not readily detectable. Caution must be exercised in interpretation, however, since the data is limited.

A promising line of investigation was suggested in the present research and should be continued and refined, i.e., frequency analysis of eye-movements and cardiac measures. Coherency plots, phase spectrum plots, and spectral density plots suggested lag-lead relationships between the two measurements. Cross-correlation analysis should also be performed in the time domain.

**EEG - Response Time Relationships**

Correlations between response time and EEG were found in the research that were not expected at the outset of the research. A positive correlation between alpha activity and response time in the 5 second interval during which a signal was presented was reported. A negative correlation was found in the 5 second interval prior to the presentation of a signal between delta activity and response time. It is not clear what these correlations mean.

Response time is a measure of central information processing and EEG is thought to be a measure of cortical activation. One might
expect some correlation between these two output measures. Although there has not been much research in this area there has been some. Lansing, Schwartz, and Lindsley (1959) correlated RT with alpha and found that RTs were significantly faster when alpha blocking occurred before a visual stimulus. The alpha block was produced by an auditory warning signal.

The results from studies correlating EEG and RT are mixed. Woodruff (1975) found small correlations between brain wave period and RT. Subjects that produced fast brain wave activity produced faster RTs than Ss who exhibited slow activity. This study examined 6, 8, 10, and 12 Hz frequencies. Thompson and Botwinick (1966) could find little support for an EEG-RT correlation. Most of the recent work in this area has been motivated by the hypothesis of Lacey.

Although correlations in the present research were significant, sample sizes were small. Increases in response time with alpha power would suggest decreasing arousal. This finding is consistent with the predictions of traditional activation theory but not with Lacey's predictions. The negative correlation of delta activity with response in the 5 second interval prior to a signal suggests that low frequency activity increases with decreases in response time—the shorter the response latencies, the more activity in the delta frequency range (sleep related EEG activity). Such a finding is more in line with Lacey's environmental intake hypothesis. The findings for alpha and delta are opposed to each other. The results are also confounded by the effects of prior signals. The credibility of this result is questionable and the results might have arisen by chance.
Heuristic Model for Present Research

A heuristic model for the results of the present research is shown in Figure 45. The model is a flow model of information processing. Central to the model are arousal, expectancies, time estimation, and a simple search process. The model assumes that expectancies and arousal influence each other and in turn influence performance. For example, if the observer has a high expectancy for a signal it will have an arousing effect. Conversely, high arousal will also affect the way expectancies are formed although the exact nature of these interactions are presently unknown. The model relies upon an observer's ability to estimate ISIs. The model is quite specific to the present research. However, it does suggest ways that future research could be directed to test implied relationships. Moreover, it is easy to see where and how the model should be modified to include task and display modifications.

Recommendations for Future Research

Recommendations for future research take two forms: (1) recommendation for research based on a review of the research literature, and (2) recommendations arising out of the present research.

Recommendations for research based on a review of the research literature. The recommendations derived from the literature are presented in synoptic form with its source where possible.

1. Vigilance research studies should employ multifactor studies with adequate controls over all relevant variables (Stroh, 1971).

2. Stroh (1971) warns that when studying arousal, one should be careful to compare only those signals that are immediately adjacent in time.
Figure 45. Heuristic Model for Present Research.
3. The problem of stressors in combination needs to receive more attention in research (Stroh, 1971; Mackie, 1977; Broadbent, 1971).

4. The validity of laboratory vigilance research results needs to be considered in terms of operational settings (Kibler, 1965; Nachreiner, 1977; Jerison & Pickett, 1963; Elliot, 1960; Craig & Colquhoun, 1975; Mackie, 1977). The proper view of vigilance may be as Kibler suggested--vigilance experiments represent a small part of the continuum of monitoring tasks to be found in field situations; vigilance research results may have low generalizability for the total spectrum of these jobs.

5. Physiology and psychology need to be linked in sustained attention research. However, it is not clear at this time which physiological measures are sensitive indicators of observer states. Kahneman (1973) advocates the use of pupil size changes and skin conductance as good measures. Stroh (1971) claims that EEG alpha incidence and skin-conductance measures have been the most consistent in showing arousal changes. Kak (1981) summarized the features of different physiological measures. According to Kak, EKG, heart rate, blood pressure, body fluids, and EEG score the highest when compared against the following list of criteria: concept validity, empirical validity, freedom from artifacts, standardization, reliability, representativeness, lack of invasiveness, and practicality and cost. Stern, Ray, and Davis (1980); Wickens (1979); Soyers (1975), Ursin and Ursin (1979), and Hamilton, Mulder, Strasser, and Ursin (1979) provide recent reviews and assessments of the various physiological measures.

6. There is a need to begin building an integrated theory of vigilance performance by building upon a task taxonomy (Davies and Parasuraman, 1982).

7. Parasuraman and Davies (1976) found that response latencies varied depending upon the response. Response latencies were found ordered as follows: $FA > M > D \equiv CR$. More research in this area should provide a better understanding of this phenomenon. Other measures, e.g. physiological measures should be studied concomitantly.

8. Signal detection theory has extended vigilance theory and provided the derived measures of performance $d'$ and $\beta$. Craig (1977, 1978) and Jerison (1967) as well as others, however, have questioned the applicability of SDT
to vigilance. More work needs to be done to relate such performance measures to the various independent variables but also to other dependent measures such as response latencies and physiological measures. Caution should be exercised in the application of SDT to vigilance. Individual ROC curves should be examined and distributional forms verified to the extent possible. Egan (1975) provided ROC analyses, based on normal, negative exponential, Rayleigh, Chi-square, Bernoulli, and Poisson p.d.f.s. Ideal observer analysis usually assumes equal variance, normal p.d.f.s. and maximization of expected value as a decision goal. Other decision goals need to be examined in terms of idealized observers in order to provide measuring sticks for actual behavior (Willeges, 1976).

9. Jerison's elicited observing response hypothesis is based on the idea that observing has an attentional cost associated with it. An observer's utility for observing changes over time in the Jerison and Pickett (1963) model. Existing measurement technology does not provide for measuring changing utilities in a decision matrix. Although, elicited observing response theory is appealing in its notion of attentional costs, the theory is largely untestable because of its measurement problems. The technology needs to be expanded in this area if possible.

10. Poulton (1975, 1977) has discussed the problems of experimental design in vigilance research. He claims that many vigilance experiments are within-subject designs which confound effects because of asymmetric transfer, i.e. uncontrollable carry-over effects exist in the design that could be eliminated by use of between subjects' designs. More consideration needs to be given to this problem than appears to be given in the literature (Expectancy research, by its nature, involves the study of carryover effects. In these instances care should be taken not to eliminate these effects).

11. Mackie (1977) mentions the need for vigilance theory to consider the role of individual differences and motivation in performance. A review of the literature seems to suggest that most of vigilance research has emphasized task and display variables and to a lesser extent subject variables.

12. More and more, it appears the vigilance literature is reporting two kinds of research: (a) low event rate
studies and (b) high event rate studies. Researchers will have to agree on a definition of vigilance. Presently, the agreement is not universal. For example, some researchers claim that driving is a vigilance task while other disagree. Low event rate and high event rate studies may require different explanations of performance, e.g. hyperarousal and hyparousal states.

13. The point has been made by several researchers e.g. Loeb and Alluisi (1977) and Warm (1977) that no single theory is rich enough to account for all vigilance research results. Considerable overlap exists with the different models and theories. Definitive tests are not available in most instances to test one theory against another. Warm (1977) claims that future theorizing will need to integrate various reviews. If the different theories cannot be distinguished by empirical tests, then either the state of technology is not sufficiently developed or researchers are speaking of distinctions without any real differences. If the latter is true then theory integration should proceed ahead and begin to take shape.

Recommendations Arising Out of the Present Research

Future research should incorporate the previous recommendations where possible. In addition, several recommendations arise from the present research:

1. Control should be tightened by counterbalancing, randomization and careful selection of the research design. This will eliminate rival hypotheses, i.e. threats to validity.

2. Sample sizes should be increased to achieve reasonable levels of power. This will also have the effect of reducing threats to validity.

3. Visual artificial signal intensity should be cross-modality matched with auditory signal intensity.

4. Better instrumentation techniques should be used to reduce the loss of data due to equipment malfunction.

5. A balanced experimental design should be used. This will eliminate the problems of confounding due to non-orthogonality. A balanced design is more efficient in its estimation of effects than an unbalanced design.
6. Physiological signals should be recorded continuously if possible. This would permit a more refined finer-grain analysis. Proper attention should be given to sampling rates.

7. Unobstrusive and nonreactive eye movement recording techniques should be used.

8. On-line analysis methods should be used to reduce the post-experimental computational load where possible.

9. Subjects should be instructed to respond to auditory signals as well as visual signals. Responses to auditory signals provide a potentially valuable source of information that was not utilized in the present research.

10. Future research should examine relationships between ISI and temporal uncertainty. The research literature suggested a family of curves was necessary to predict response times under low, medium, and high temporal uncertainty. The curves look like those that would be predicted by classical conditioning, Deese's expectancy hypothesis and Baker's expectancy hypothesis. Manipulation of ISI and temporal uncertainty over a wide range of ISI's should clarify these relationships.

11. Expectancies have been shown to affect performance. Arousal states have also been shown to affect performance. Little work has been done, however, to clarify the relationship between expectancy and arousal. For example, how does time error change under different activation states, say sleep deprivation versus a highly aroused state.

12. More work needs to be done to establish the validity, reliability, and sensitivity of various physiological measures. Early research has criticized the use of physiological measures because of low correlations. Low correlations are not surprising because of curvilinear relationships and different response latencies among measures. Time series analysis as well as frequency analysis should provide insights into the nature and role of activation in vigilance.

13. More recently developed dependent measures such as the pupillary response, adrenalin gland output of corticosteroids and catecholamines, head temperatures, evoked potentials of the brain, and electrodermal measures should be investigated. Their relationship to task and display changes and to each other should be studied.
14. More use should be made of subject debriefing. The present research did not take advantage of this information—no formal debriefing occurred.

15. A systematic mapping of arousal in relation to various stressors singly and in combination needs to be done in order to better understand the nature of the arousal mechanisms. For example, does arousal depend on one, two, or three mechanisms? Sternberg's converging operation techniques, e.g. additive factor techniques may be useful.

16. Little work has been performed to determine the effects of various types of signals and other types of signals. Operational settings typically use mixes of signals with few design guidelines to indicate the mutual effects of various mixes of signals on performance. Binford and Loeb (1963) did some early work that studied readily detected auditory signals in the presence of obscure visual signals but little has been done since that time. Binford and Loeb found that latency of responses decreases with increasing frequency of occurrence of the auditory signal and increases of latency with time on the task. It may be fruitful to augment behavioral measures with physiological measures in this kind of research. Physiological measures in the present research suggested a sensitivity to signals. The present work should be extended to include other signal modalities such as tactile and vibratory modalities.

**Future Research**

The notion of expectation plays an especially important part in human factors research and design. The design of systems should consider S-R compatibility, population stereotypes and language norms. Another way of saying this is that designers should take advantage of the "expected" relationships that people have. Display and control arrangements and codings that are not anticipated worsen performance. The notion of expectancy is pervasive to human factors. Expectancy also has a significant role in affecting performance in monitoring tasks.
Expectancy has links with arousal. Expectancy assumes an alert state, but what happens to expectancies when activation wanes or when people are overaroused? How many times have people responded to what they thought was a stimulus when it was not really present? The stimulus-response sequence is much like it would be if the stimulus was really present. Some pattern in the environment is matched with some encoded set of features in memory. The degree of feature extraction and matching is probably influenced by the task loading and the activation state of the monitor. The model in memory may be partially formed and incomplete. Feature extraction may be incomplete either because of low arousal or excessive environmental demands and time pressures. The outcome of the matching process may be "rough" in which case expectancy bridges the gap between external events and stored models of previous similar situations. False alarms and misses that occur in vigilance tasks are an outcome of this incomplete feature matching process. False alarms and misses represent uncertainty about the feature-matching process. More time is spent searching memory for patterns in this ambiguous situation than where the match has good correspondence. Evidence exists for this assertion. Davies and Parasuraman (1982) report longer response times for errors than for correct responses. They report latency relationships as follows: FAL > ML > DL ≡ CRL. A false alarm latency is longer because the feature extraction process has proceeded further along than for a missed signal although both responses represent situations where considerable uncertainty exists. The situations are as if feature extraction proceeds
along as far as it can and then a response depends upon whether or not a threshold based on a single feature or a stored set of features is exceeded. Expectancy may be viewed as the mechanism that bridges the gap or completes the matching process.

A model of vigilance based upon expectancy and activation does not explain all research results but it can account for many. Future vigilance research can benefit from a more thorough understanding of expectancy and activation in the feature matching process.

A vehicle or paradigm for better understanding the role of expectancy and activation in vigilance would consider manipulating ISI distributions and distribution characteristics, temporal and spatial uncertainty and stimulus event rates. Instrumentation should include behavioral and physiological measures. Earlier research in this area has focused on behavioral measures to the exclusion of physiological measures.

Initial steps in the study of expectancies and arousal should concentrate on the variables that seem to be most directly related to both concepts; namely temporal, spatial, and conspicuity variables. Once these variables are integrated into a model, environmental, subject variables and other task variables can be brought into the research to determine how they affect the model.

Figure 46 outlines some of the display and task variables directly related to activation and expectancy. No single experiment can address the complexity of all these variables. The outline represents a program of research to extend the present research and to build and clarify relationships and theory.
Figure 46. Temporal and Spatial Display and Task Variables in Vigilance Research.
One of the first experiments should aim at clarifying the relationship of ISI to expectancy and activation. Initial experiments should be simple and straightforward and not complicated by the presence of too many other variables. ISI characteristics can be manipulated on a single display. Other variables should be fixed or eliminated. The research extends the work of Boulter and Adams (1963); Baker (1963c); Faulkner (1962); Deese (1955); Jenkins (1953); Dardano (1962); Johnston; Howell, and Goldstein (1966); Bevan; Hardesty; and Avant (1965); Adams and Boulter (1964); Baddeley and Colquhoun (1969); Smith, Warm, and Alluisi (1966); and McCormack (1960, 1962), Warm, Epps and Ferguson (1974) and Warm (1977). Warm (1977) suggests that response time is an increasing linear function of stimulus uncertainty. Moreover, Hick's Law applies to vigilance tasks as well as choice reaction time tasks. However, studies by Boulter and Adams (1963), Bevan, Hardesty, and Avant (1965), Johnston, Howell, and Goldstein (1966), and Dardano (1962), and Jenkins (1953) suggest different relationships between response time, ISI and signal uncertainty.

Certain limits are imposed in mapping the effects of ISI and signal regularity (variability of ISIs) on performance. For example, if the mean ISI is 10 seconds, the standard deviation of the ISIs cannot be larger than 5.7 seconds; otherwise some of the ISIs in the p.d.f. would have negative values which is not possible. Table 30 shows mean ISI and the limits on regularity that are imposed by the non-negativity restriction. The range of mean ISIs (10-640 seconds) includes the range that was used in the present research. Because of
Table 30. Mean Intersignal Interval (ISI) and Limits on Regularity (Standard Deviation of ISI) for ISIs With a Uniform P.D.F.*

<table>
<thead>
<tr>
<th>Mean ISI (seconds)</th>
<th>Average Signal Rate Per Minute</th>
<th>Average Signal Rate Per Hour</th>
<th>Maximum Possible Standard Deviation of ISI (seconds)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>6.0</td>
<td>360</td>
<td>5.7</td>
</tr>
<tr>
<td>20</td>
<td>3.0</td>
<td>180</td>
<td>11.5</td>
</tr>
<tr>
<td>40</td>
<td>1.5</td>
<td>60</td>
<td>23.0</td>
</tr>
<tr>
<td>80</td>
<td>0.75</td>
<td>45</td>
<td>46.1</td>
</tr>
<tr>
<td>160</td>
<td>0.375</td>
<td>22.5</td>
<td>92.3</td>
</tr>
<tr>
<td>320</td>
<td>0.1875</td>
<td>11.25</td>
<td>184.8</td>
</tr>
<tr>
<td>640</td>
<td>0.09375</td>
<td>5.625</td>
<td>369.5</td>
</tr>
</tbody>
</table>

*Distribution characteristics - uniform p.d.f.:

\[
\begin{align*}
    f(x) &= \frac{1}{A - B} \quad \{A \leq x \leq B \} \\
    E(x) &= \frac{A + B}{2} \quad \text{o otherwise} \\
    \text{VAR}(x) &= \frac{(B - A)^2}{12}
\end{align*}
\]
the limits on regularity (temporal uncertainty), the range of conditions do not lend themselves to a straightforward balanced experimental design. Dardano (1962) avoided this problem by using a mean ISI of 60 seconds and ISI variances of 62.5, 562.5, and 1562.5 for low, medium, and high uncertainty conditions. The research suggested here is more extensive in its conditions. A summary of conditions to be studied are shown in Table 31.

Design

Three factorial experiments can be conducted as shown in Table 31. This arrangement would provide for efficient experimentation and analysis and allow one to determine interactions. Experiment 1 is a 4 x 7 factorial design, experiment 2 is a 2 x 3 factorial design and experiment 3 is a 2 x 5 factorial design. The remaining experiment with the conditions (ISI = 20, VAR (ISI) = 133), (ISI = 40, VAR (ISI) = 40), (ISI = 40), VAR (ISI) = 533), (ISI = 160, VAR (ISI) = 8533), and (ISI = 640, VAR (ISI) = 136533). Each S will be randomly assigned to conditions and experience only one treatment combination.

Subjects: The subjects in the experiment would be college students with 20/20 vision and no hearing defects. Three Ss would be used in each cell of the factorial experiments and ten for each condition of the single factor experiment for a total of 182 subjects.

Apparatus: A display such as Vickers et al (1977) used would be appropriate to this research. Vickers used two vertical lines 40 mm apart with their upper ends terminated by a third, horizontal line (similar to the symbol \( \pi \)). This symbol is displayed for 0.2 seconds in the middle of a CRT which is observed by the monitor from a distance of 1.6 mm. The shorter leg is always 140 mm. whereas the longer leg is 141 mm. for the signal condition (neutral events have both legs equal). The differences of 1 mm. in line length
Table 31. Summary of Proposed Research Conditions—Mean Intersignal Interval and Variance of ISI.

<table>
<thead>
<tr>
<th>Variance ISI</th>
<th>Mean ISI (seconds)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td></td>
</tr>
<tr>
<td>.40</td>
<td></td>
</tr>
<tr>
<td>80</td>
<td></td>
</tr>
<tr>
<td>160</td>
<td></td>
</tr>
<tr>
<td>320</td>
<td></td>
</tr>
<tr>
<td>640</td>
<td></td>
</tr>
</tbody>
</table>

Factorial Experiment 1

Factorial Experiment 2

Factorial Experiment 3
subtends a visual angle of 0.04 degrees. The side which is longer should be randomized. Event rate should be held constant for all conditions. Treatment combinations (mean ISI and variance of ISI) are as shown in Table 31.

The CRT would be driven by a PDP 11/34 computer. Subject responses to the display would be recorded by the computer. Subjects would be instrumented for response latency (by type of response), EKG, EEG, electrodermal activity, pupillary response and blink rate. In addition, urine levels of corticosteroids and catecholamines should be taken before and after each session. The EKG, EDA, EEG, and pupillary and blink responses would be recorded around a signal continuously for all measures.

Procedure: Subjects will be trained on a one half hour session with conditions the same as they will encounter in their experimental sessions (except they will not be instrumented for physiological measures). The task duration for each treatment combination will be one hour. Subject debriefings will be conducted after each experiment.

Another, perhaps more satisfying design approach, is to randomly sample ISI levels and regard this independent variable as a random effect. This approach would be more efficient than a complete mapping of ISI effects; moreover, the results would generalize to other ISI levels. After ISI effects are determined, other variables such as appear in Figure 46 could be researched to further develop the expectancy activation constructs.
LIST OF REFERENCES
LIST OF REFERENCES


Davidson, Michael L. "Univariate versus multivariate tests in repeated-measures experiments." *Psychological Bulletin*, 1972, 77(6), 446-452.


Firth, P. "Psychological factors influencing the relationship between cardiac arrhythmia and mental load." *Ergonomics,* 1973, 16(1), 5-16.


Fisk, A. D. and Schneider, W. Controlled and automatic processing during tasks requiring sustained attention: a new approach to vigilance. Human Attention Research Laboratory, Psychology Department, University of Illinois, Champaign, IL, Report No. 8006, 1980.


James, Mary Ellen; Deborah Pierce; Deborah Warren and Rhonda Macchello. Bibliography and classification of the literature of pupillary response. Waltham, Massachusetts: G & W Applied Science Laboratory, 1975.


Kalsbeek, J. W. H. "Do you believe in sinus arrhythmia?" *Ergonomics*, 1973, 16(11), 99-104.


Lambert, R. H., R. A. Monty, and R. J. Hall. "High-speed data processing and unobtrusive monitoring of eye movements." Behavior Research Methods and Instrumentation, 1974, 6(6), 527-528.


MacNeilage, Peter F. EEG amplitude changes during different cognitive processes involving similar stimuli and responses. *Psychophysiology*, 1966, 2(4), 280-286.


Namboodiri, N. Krishnan. "Experimental designs in which each subject is used repeatedly." *Psychological Bulletin,* 1972, Vol. 77, No. 1, 54-64.


ONSER, Summary of Driver Vigilance Research Conducted by ONSER's Psychophysiological Laboratory. Organisme National de Securite Routiere, Division de Psychologie de la Conduite, Laboratoire de Psychophysique, Autodrome de Linas - Montlhery, 91 Montlhery, France.


Papakostopolous, D. and Bernard Fenelon. Spatial distribution of the contingent negative variation (CNV) and the relationship between CNV and reaction time. Psychophysiology, 1975, 12(1), 74-78.


Poulton, E. C. "Unwanted range effects from using within-subject experimental designs." Psychological Bulletin, 1975, 80(2), 113-121.


Roeckelein, Jon E. "Time perception: a selective review." JSAS Catalog of Selected Documents in Psychology, 1972, 2:103. (Ms. No. 208)


Wilder, J. The law of initial values in neurology and psychiatry. Journal of Nervous and Mental Disease, 1957, 125, 73-86.


Williams, L. G. Target conspicuity and visual search. Human Factors, 1966, 8, 80-92.

Woodruff, Diana S. "Relationship among EEG alpha frequency, reaction time, and age: a biofeedback study." Psychophysiology. 1975, 12(6), 673-680.

Yerkes, R. M. and J. D. Dodson. The relation of strength of stimulus to rapidity of habit-formation. *J. of Comp. and Neurol. Psychol.*, 1908, 18, 459-482.