1985

The Effect of Differing Goal Strategies on Subjective and Physiological Indices of Workload Across Time

Kevin C. Uliano
University of Central Florida

THE EFFECT OF DIFFERING GOAL STRATEGIES ON SUBJECTIVE AND PHYSIOLOGICAL INDICES OF WORKLOAD ACROSS TIME

BY

KEVIN CHARLES ULIANO
B.A., University of Central Florida, 1983

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science
in Industrial/Organizational Psychology
in the Graduate Studies Program
of the College of Arts and Sciences
University of Central Florida
Orlando, Florida

Spring Term
1985

ABSTRACT

The purpose of this study was to investigate the effects of differing goal strategies on subjective and physiological indices of workload across time. The sample consisted of 16 males and 24 females from undergraduate psychology classes at the University of Central Florida. Subjects were assigned to four goal conditions: time/accuracy, time, accuracy, and no goal, and asked to perform a computer-based decision making task comparing visual and semantic information. A trial consisted of a 15-minute baseline and three 5-minute task periods. Dependent variables included electromyopotential (EMG) measured in microvolts and a paper and pencil workload scale utilizing a Likert-type format and measuring three dimensions: general psychological stress (GPS) load, mental effort load, and time load. Results indicated that assigned goal strategy had no effect on the workload indices. Analyses of variance and trend analyses, however, revealed that EMG and mental effort load both increased from baseline to task period 1 then decreased across time. This relationship was just the reverse for GPS load . In addition, time load decreased across time in a significant linear fashion. Zero-order correlational analyses were also performed using all dependent variables. EMG and time load were inversely related during task periods 1 and 2 whereas mental effort and GPS load were related only during task period 1. Results are discussed with reference to future research methodology in the area of workload assessment.

ACKNOWLEDGEMENTS

This was by no means an individual effort. First, I would like to thank my parents, George and Marie, for so many thoughtful and loving gestures; too many to mention; too many to forget. Second, to Terese, the love of my lifetime, for her support and sound critiques at various stages of this undertaking. My thesis committee members were also a limitless source of encouragement and were always available to lend their expertise -- for this I am truly grateful. Finally, a special thanks to Sara Bennett and Mark Arnold -- two extremely competent and insightful research assistants.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Roll (1981) has observed that the term stress has been borrowed from the natural sciences. It is a term used to refer to the "elastic limit" of an entity or substance. This basic definition of stress has not been shared by colleagues in the social sciences especially psychologists.

Problems with the Definition of Stress

The term stress is not employed with a great deal of consistency in the psychological literature. Alluisi (1982) made an inquiry into the on-line computer data base of the American Psychological Association known as PsychInfo. He found that the literature dealing with stress tends to be organized into areas emphasizing physiological, psychological, social, and many other categories. This trend is empirically sound yet lacking in utility because, more often than not, operational definitions get lost in the research shuffle. Similarly, Lester (1979) conducted a data base search dealing with psychological stress from projects funded by the Department of Defense. He also cited evidence to support the notion that psychological stress has been employed to cover a multitude of variables. Specifically, Lester stated that the 356 reports produced a total of 647 categorizations which could be grouped into 12 general topic clusters ranging from task performance to coping strategies.

Hence, research into the nature of stress has most certainly evolved into a variety of "specific" disciplines. Ironically, this progression has taken place almost 30 years after the now famous general definition of stress given to us by Hans Selye. He viewed stress as "the non-specific [italics added] response of the body to any demand made upon it" (Selye, 1956, p. 27).

Stress research now finds itself floating in conceptual confusion. Each writer must define and re-define his or her term(s) anew, and one must carefully check each article to make sure one understands the author's vocabulary. This is especially tedious when stress research transcends discliplines as diverse as clinical and applied psychology, anthropology, sociology, psychosomatic medicine, and others. Hogan and Hogan (1982) observed that review articles (see for example Averill, 1973; Pervin, 1968) tend to be highly specific to various discliplines, tracing one or another specialized facet of stress research. From a different perspective, Lazarus, Deese, and Oster (1952) talked about the pervasiveness of individual differences in stress reactions. This is something McGrath (1970) called the "cognitive appraisal theme" -- the idea that stress reactions are a function of an individual's perceptions, expectations, experiences, moods, and personal appraisals of the stressors themselves.

How, then, does a researcher gain insight into measures that are so individual and specific? One possible answer is to explore the concept of subjective self-report data. Soutendam, writing in Wilkens (1982), explains: "Emerging from all this is that, potentially, it may well be that the easily administered and very economical paper and pencil research methods do provide reasonably good indicators that parallel the physiological indicators" (p. 78). He also states that psychological (i.e., subjective) factors may make their presence known before the physiological (i.e., objective) factors do.

Human Factors and the Concept of Workload

The field of Human Factors Psychology, which typically concerns itself with the enhancement of performance through the design and arrangement of training devices to fit human capabilities (Anastasi, 1979), has adopted the construct of stress and renamed it "workload." This approach serves to narrow the scope of inquiry to stressors that can only be linked to the device(s) or task(s) in question. Actually, stress is supposed to indicate the effects of workload upon man (Rasmussen, 1979). In any case, two parallels are immediately obvious when one compares workload research with stress research.

First, there is no agreed-upon definition of workload. Some definitions emphasize the physical components while others emphasize the psychological components. For example, Chiles (1982) uses the term "level of operator workload" to refer to a hypothetical construct that is determined by or related to the total task demands placed on the operator by the system of which he or she is an integral part. Eggemeier (1984) views workload as that portion of an operator's limited processing capacity which is actually required to perform a particular task or system function. The term "system" used in these definitions refers to some man-machine configuration and therein resides the intention. That is, by using words like "operator" and "system," it is desirable to assure that system demands do not exceed the information processing capabilities of the human operator. In these definitions the system is the sole contributor of the phenomenon called workload.

Moray (1979) goes one step further in stating that subjective mental load (SML) is the only real meaning of mental load. These approaches are strikingly straightfoward in their view of workload. Johannsen et al. (1979) express this succinctly when they write: "...if the person feels loaded and effortful, he is loaded and effortful...." Senders (1970), on the other hand, asserted that unless there is "time stress" in a task there exists no SML. Other writers (e.g., Leplat,1978) maintained that mental workload should be linked to personality variables and to such social variables as social pressure and expectations. Unfortunately, Moray (1982) has noted that despite the widespread use of subjective techniques, very little has been published since 1968 concerning this approach. This trend, however, is beginning to reverse itself (cf. Alluisi, DeGroot, and Alluisi, 1984).

The second parallel that workload research shares with stress research is that there is no agreed-upon method to measure workload (Wierwille & Williges, 1978; Moray, 1979; Ogden, Levine, & Eisner, 1979; Williges & Wierwille, 1979). The consensus is that workload is a multidimensional concept composed of behavioral, performance, physiological, and subjective components (Johannsen et al., 1979). This multidimensionality of workload provides a convenient taxonomy or classification scheme in which to view workload (Moray, 1979; Eggemeier, 1984; Kramer, Wickens, & Donchin, 1983).

The first category houses those studies dealing with physical and physiological parameters. Definitive work in this area rests with Gunnar Borg and his colleagues at the Institute of Applied Psychology located in Stockholm, Sweeden (Borg, 1978a, 1978b; Borg, Bratfisch, & Dornic, 1971a, 1971b). Their interest in this area was spurred by the observation that persons engaged in strenuous manual labor reported evaluations that differed from those of their own personal physicians in regard to the person's working capacity. These subjective experiences of physical performance and working capacity led to the Ratings of Perceived Exertion (RPE) scale (Borg, 1970). The RPE is a category rating scale from 6 to 20; this is said to match heart rate variation from 60 to 200 beats/min. Every second number is anchored with verbal expressions such as "fairly light," "hard," "very,very hard," etc. Validation studies (e.g., Borg and Noble, 1974) using the RPE scale have reported correlation coefficients as high as .85 between ratings and heart rate. Reliability studies (e.g., Stamford, 1976) similarly report coefficients ranging from .76 to .90. Also, Wardle (1978), using Borg's RPE scale, concluded that people perceive an extremely close relationship between the actual strenuousness of their work output and their bodily state. Heart rate and heart rate variability are the most researched areas when looking for promising physiological correlates of workload.

Wierwille (1979), however, has looked at other viable measures such as pupil dilation, body fluid analysis, evoked cortical potential, and electromyography (EMG). He concluded that more research is needed to provide convincing evidence of viability. The basic assumption which governs this facet of workload research is as workload changes, involuntary changes take place in the physiological processes of the human body. The key in using these measures (or any other workload measures for that matter) is to determine the extent of "sensitivity" of the instrument. A sensitive workload estimation technique is defined as one that can discriminate between different workload levels (Casali and Wierwille, 1983). Ostensibly, if it is possible to produce different levels of a construct then one may actually be measuring different degrees of the the same concept. That concept is defined as workload. In the Casali and Wierwille study, 16 potential techniques for estimating workload were investigated. The authors concluded that two subjective opinion measures were sensitive to changes in workload. In contrast, pupil diameter was only one of five physiological measures that proved sensitive to changes in workload. Wierwille and Connor (1983), in a similar study, evaluated 20 workload measures. Again, two subjective rating scales demonstrated significant load effects; the only physiological measure that showed significance was mean pulse rate.

Another criterion of concern is intrusiveness. This occurs when task performance is degraded by the introduction of the assessment technique (Eggemeier, 1984). The degree of intrusion associated with subjective and physiological workload measurement techniques have been reported to be minimal (Eggemeier, 1984; Rahim &

Wierwille, 1982). Intuitively, subjective assessment techniques would typically present no significant intrusion problem since rating scales and other report procedures are usually completed subsequent of task performance. At the same time, however, Rehmann, Stein, and Rosenberg (1983) effectively argue that this procedure relies too much on the operator's memory. They propose collecting subjective workload data during task performance thus closing the time gap between experiencing the work situation and attempting to report on it.

Thus, the question now arises as to whether a mentally demanding task invokes the same kind of responses as a physically demanding one. The second category of the classification scheme deals strictly with subjective opinion procedures that have been alluded to previously. To recap, in the Casali and Wierwille (1983) and Wierwille and Connor (1983) studies, the authors found that subjective opinion measures differentiated between different workload levels.

The historical foundation of subjective workload assessment can be traced back in the human factors literature to the Cooper (C) and Cooper-Harper (CH) scales (Cooper, 1957; Cooper & Harper, 1969). These scales were initially developed to measure the handling characteristics of aircraft and as such focus their attention on the "machine" side of the system. The C and CH scales, in any case, are well established and validated subjective instruments (McDonnell, 1969, Moray, 1982). They are useful instruments when one wants to talk about the "flyability" of aircraft. Unfortunately, the success of the C and CH scales has not generalized to other applications. Some modified forms of the original scales have appeared (North and Graffunder, 1979; Casali & Wierwille, 1983) yet the focus still remains on specific aircraft characteristics or on the psychomotor aspect of the task. We must only assume that if a pilot or operator states that an aircraft is difficult to fly, this is then similar to the assertion that the task is producing workload.

There has recently been a movement toward workload generalization instruments. To illustrate, Wierwille and Casali (1983) presented a validated rating scale for global mental workload measurement. It is a modified version of the Cooper-Harper Scale (called MCH) with all the references to specific pilot/aircraft characteristics being changed to accommodate new wording such as task accomplishability, errors, difficulty, performance, and mental workload. The authors make the assumption that as systems become more complex, there is a tendency to use human operators less frequently as active control system elements and more frequently for their other abilities such as perception, communication, and problem-solving. In any case, the argument must be made that the MCH scale, like the original Cooper-Harper Scale, uses an awkward decision tree format that may necessitate an extended training and instruction period for naive users.

At the United States Air Force Aerospace Medical Research Laboratories (AMRL) there is also a concerted effort to understand workload. AMRL has developed a subjective workload scale called the Subjective Workload Assessment Technique (SWAT) (Reid, Shingledecker, & Eggemeier, 1981; Reid, Shingledecker, Nygren, & Eggemeier, 1981). Using this technique, workload is defined as being composed of three dimensions. The first is time load. This refers to how much time is available for an operator to perform a task. If applicable, this also may include time between individual task presentations (i.e., task pacing). The second dimension is mental effort load. This refers to the amount of attentional capacity or effort required without regard to the amount of time available or to task pacing. The last dimension is psychological stress load. This has been referred to as anything that makes that task more difficult by producing anxiety, frustration, fatigue, etc. In the SWAT process, each subject provides an ordering from 1 to 27 representing his/her opinion of the workload associated with combining descriptors for three levels on each of the three dimensions. The dimensions are then combined through a mathematical process known as conjoint measurement and scaling which attempts to produce scales from 0 to 100 that have interval level properties. SWAT has been applied to more general tasks such as display monitoring (Notestine, 1983) and verbal short term memory (Eggemeier, Crabtree, Zingg, Reid & Shingledecker, 1982) yet its strength remains system evaluation. Eggleton and Quinn (1984) even discuss a projective workload assessment procedure (PRO-SWAT) that seeks to evaluate workload inplications of technology options before they exist in hardware form.

The last classification scheme of workload is the idea of a performance-based technique (e.g., Gartner & Murphy, 1976; Williges & Wierwille, 1979). This technique focuses on some measure of operator behavior or activity as the basis of a workload index. There are two generally agreed-upon performance techniques. First, is the primary task method (Rolfe, 1976; Gartner & Murphy, 1976; and others). This particular technique, seeking to provide an estimate of workload, examines some aspect of the operator's capability to perform a task. Basically, deviations from criterion task performance would indicate a primary task measure of workload. Gopher and Braune (1984) add that there is little justification in developing a workload measure that is not related in some way to the actual behavior of subjects. The second type is the secondary task method (Knowles, 1963; Odgen, Levine, &

Eisner, 1979) or spare mental capacity (Williges & Wierwille, 1979). This technique seeks to determine "...how much additional work the operator can undertake while still performing the primary task to meet some system criterion" (Ogden et al., 1979, p. 529). The secondary task method has been criticized for being ineffective and impractical (Pew, 1979; Fisk, Derrick, & Schneider, 1983). Fisk et al. (1983) have claimed that haphazard combinations of two tasks, one or both of which may not be realistic or practical, may lead to misleading results.

Based on the literature thus far reviewed, there clearly exists a need to measure workload reactions to a single cognitive decision making task that is not part of any implicit system configuration. Past research has utilized independent variables that were created by the manipulation of the system. The pervasive design has been to use a moving-base flight simulator and vary the difficulty of the simulated air-to-ground communication requirements (Casali & Wierwille, 1983) or attempt to produce levels of psychomotor workload by manipulating wind-gust disturbance level and pitch stability (Wierwille & Connor, 1983). The more "mundane" tasks have been overlooked.

Similarly, there exists a clear need for a generally applicable, easy to administer, and easy to understand subjective measurement technique that serves strictly to operationalize the term workload. Subjective workload can then be defined in terms of the characteristics of the instrument that seeks to measure it. These characteristics include asking the subject questions about the degree of workload imposed by a cognitive task and its assigned strategy for performance. It could be that "...mental load both depends upon the goals aimed at and the strategies used, and can also influence them" (Hacker, Plath, Richter, & Zimmer, 1978, p. 187). We know that goals serve to direct attention and action (Locke & Bryan, 1969). We also know that goals are immediate regulators of human action (Locke, Shaw, Saari, & Latham, 1981), but can they be considered regulators of workload? The hypothesis is that the more difficult goal would be achieved by expending greater effort and attention than would be expended to achieve a less difficult goal. To this author's knowledge, however, no attempt has been made to examine the effects of differing goal strategies on workload. Workload results can in turn be compared to a physiological parameter of workload operationally defined as EMG frontalis.

Research Objectives

The first question being investigated in this study was: Are there differences in subjective and physiological indices that are a function of an assigned task performance strategy or goal?

The second question asked: Is there a relationship between subjective and physiological workload data? The Hacker, Plath, Richter, and Zimmer (1978) study comprises the woefully scant research that has attempted to demonstrate such a relationship. This study reported a correlation coefficient of .50 between physiological measures such as heart rate and critical flicker fusion and ratings of "mental impairment" and "emotional state." This study unfortunately does not provide an adequate method section in which to judge the results. Hicks and Wierwille (1979), on the other hand, have reported insignificant intercorrelations between five workload measuring techniques. These techniques included heart rate, subjective ratings, and primary and secondary task performance. As is evident, the results are mixed.

A final question involved the use of a repeated measure design as discussed by Rehmann, Stein, and Rosenberg (1983). It is postulated that if EMG and subjective variables are to be considered valid measures of workload then there should be a practice or habituation effect reflected by a decrement in these variables as the time-on-task progresses.

METHOD

Subjects

Forty male and female students from the University of Central Florida (24 women and 16 men ranging in age from 17 to 35 years of age) served as subjects for this experiment. They were volunteers recruited from psychology classes and working toward an extra credit laboratory assignment. Subjects who completed the experiment received bonus points toward the course grade. Each subject was required to read and sign an informed consent form (see Appendix A) before beginning participation in the study. Subjects denied any physical impairments or health problems. Subjects were randomly assigned to four experimental conditions, n=10, N=40, reflecting four differing goal strategies.

Experimental Design

The primary purpose of this study was to gather data on the effects of differing goal strategies on workload. Specifically, what is the effect on dependent measures (such as subjective paper and pencil dimensions and an objective physiological reading) when subjects acquire or learn a decision making task under differing goal conditions (i.e., time/accuracy strategy, time strategy, accuracy strategy, and no assigned strategy). To this end, a Three-Way Analysis of Variance (ANOVA) with repeated measures was used (Hays, 1981). In this mixed design, the two between-subjects factors (i.e., time and accuracy) each contained two levels thus creating the matrix of four groups just described. The within-subjects factor contained six levels of the time-on-task variable with select dependent measures being collected at six 5-minute intervals during the trial. Separate ANOVAs were calculated; one for each dependent measure. Post hoc analytical and/or pairwise comparisons across treatment variables were also calculated in the presence of a significant main effect. In addition, a secondary objective was identified: to determine the relationship between dependent measures. Zero-order correlations were calculated for the purpose of addressing this secondary objective.

Apparatus

Programmed instructions and stimuli were presented by an Apple II Plus microcomputer system which consisted of a 48K Apple II Plus, two disk drives, and a nine inch (measured diagonally) green phosphor video monitor. The MBERT program (Uliano & Carey, 1984) was used to standardize the presentations of instructions in an effort to minimize the possibility of experimenter contamination. The significant aspect of computerized interaction with subjects is that any unknown contaminating variable(s) will be consistent across subjects and groups. A commercially available computer software package (Conduit, Laboratory in Cognition and Perception) was modified and employed to deliver the stimuli (task).

Electromyograms

Electromyograms (EMG) were recorded from the frontalis (forehead) muscle group by an Autogen 1700 Electromyograph (Autogen Systems, Inc.) using the 100-200 Hz bandpass. The EMG meter was concealed from the subject by taping a 3 X 5 index card over it. Readings from the muscle site were integrated by an Autogen 5100 Digital Wave Form Analyzer (Autogen Systems, Inc.). The function of the 5100 is to compute the cumulative average value of a constantly changing physiological parameter for a preselected period of time, in this case mean EMG readings across six 5-minute segments of the trial.

Mediational (Cognitive) Task

Like the MBERT program described previously, the task chosen was microcomputer-based and delivered on the computer system described. It required the subject to process visual and semantic information to arrive at a decision (Trotter, 1980; see Appendix B for task instructions). More specifically, the subject was presented with two stimuli; a "*" (star) and a "+" (plus). A statement about the relative positioning of the "star" and "plus" followed (see Appendix C for a sample frame). A typical statement might have been as follows: "The plus is not below the star." Therefore, for each trial the statement was randomly assigned one of two stimuli (i.e., "plus" or "star"), one of two verbs (i.e., "is" or "is not"), and one of two prepositions (i.e., "above" or "below"). The subject was required to respond either "true" or "false" to each presentation by pressing the appropriate key on the computer keyboard. The subject was allowed up to five practice trials to acquire familiarity with the task.

Workload Measurement Scale

The Workload Measurement Scale (WMS) developed specifically for this study is a paper-and-pencil rating form which views the construct of workload as a combination of three dimensions: time load, mental effort load, and general psychological stress load (Eggemeier, Crabtree, & LaPointe, 1983). Each item comprising the WPS was given unit weighting toward the total dimension score. Items 1 and 2 (see Appendix D) measured mental effort load. Items 3 through 8 measured general psychological stress load.

The WMS for the task periods (see Appendix E) used the same items in a slightly different order and also included items 1 and 2 which measured time load. The rating form is similar to the Likert (1932) approach of developing attitude scales. Each of the items was rated on a seven-point scale with behavioral anchors placed at the "1," "4," and "7" locations.

Procedure

Subjects were randomly assigned (n=10) to four groups of a Three-Way ANOVA with repeated measures (N=40). The within-subjects (repeated) variable was time-on-task. The dependent variables under consideration and the logistics associated with their measurement dictated the number of levels for the within-subjects variable. More specifically, EMG was recorded across all six 5-minute intervals. The mental effort load and general psychological stress load dimensions of the WMS were recorded at four intervals, baseline (total of 15 minutes) and each one of the three task periods. The time load dimension of the WMS, since it applied only to the task, was recorded on each of the three 5-minute task periods. The between-subjects variables were general task performance strategies. The two levels of the two between-subjects variables (i.e., time and accuracy) created four groups: time/accuracy strategy; time strategy; accuracy strategy; and no assigned strategy. The dependent measures under observation were (1) EMG data and (2) WMS ratings on the three dimensions described previously.

Subjects, upon arrival at the laboratory, were first required to read and sign an informed consent form. Next, they were instructed by the experimenter to sit in a straight-back chair at approximately one meter's distance from the computer console. The programmed instructions (MBERT) informed subjects of the basic nature of the experiment and provided general computer operation procedures.

Electromyographic data were collected by silver-silver chloride electrodes attached to the skin with adhesive collars. Prior to attachment, the electrode sites were cleaned with alcohol-moistened cotton balls. Redux Paste (Hewlett Packard Medical Electronics) was used as the conductive medium. Specific locations for electrode placement on the frontalis followed the procedure outlined by Lippold (1967). Briefly, this entails surface electrode placement approximately one inch above the eyebrow and 1.5 inches on either side of the midline. A ground electrode was placed on the center of the forehead. The electrodes were attached in such a manner as to not interfere with the individual's peripheral vision or overall comfort. With completion of these preparations, the experimenter exited the room to monitor the Data Integrator (i.e., Autogen 5100) in the adjacent office.

The total time required to complete the experiment was 30 minutes and was broken up as follows: First, a 15-minute baseline period allowed for the collection of data prior to the introduction of the independent variable. This data collection included both baseline EMG readings as well as baseline WMS ratings. For the purpose of the latter, subjects were instructed to consider the word "task," used in the construction of the WMS scale, as the baseline period as well as the task per se. In addition, the items concerning "time load" (i.e., questions 1-2, Appendix E) were omitted from the baseline WMS instrument since they did not apply to the baseline period per se.

After baseline, the experimenter re-entered the room and administered the WMS instrument. The procedure for completing the rating form was explained in MBERT and was verbally supplemented at this point by the experimenter on an as-needed basis. Prior to the beginning of the task intervention period, the experimenter read, verbatim, a prepared statement outlining the goal strategy for the task (see Appendix F for Strategy by Group). Next, the final fifteen minute task period began. This period was partitioned into three 5-minute sub-periods and constituted the repeated measure. At the conclusion of each sub-period, the experimenter entered the room and instructed the subject to temporarily stop what he/she was doing; at that time the WMS was administered. Also, the experimenter reminded the subject of the task strategy. This procedure was repeated for each of the three sub-periods. Removal of the electrodes followed; finally, the subjects were thanked for their participation. Subject name and social security number (or class I.D. number) were recorded to insure the awarding of extra credit bonus points.

RESULTS

The reliability of the three dimensions used in the WMS was determined using coefficient alpha (Cronbach, 1951). Coefficients of .72, .68, and .57 were attained for time load, mental effort load, and general psychological stress load respectively.

Goal strategy (i.e, time/accuracy strategy, time strategy, accuracy strategy, and no assigned strategy) had no effect on any one of the four dependent variables (please refer to Tables 2 through 5) used. There also appeared no first or second order interactions. Moreover, the hypothesis that the more demanding goal strategy (i.e., time/accuracy strategy) would produce greater workload was rejected. The repeated measure or time-on-task variable, however, produced significant results across all four dependent measures (see Table 1 for means and standard deviations).

Table 2 shows that time-on-task had an effect on the general psychological stress (GPS) load dimension of the WMS ($F = 2.74$, $df = 3{,}108$, $p < .05$). Figure 1 graphically portrays this relationship. Post hoc comparisons across time-on-task revealed a decrease in GPS from baseline to the first task period ($F = 6.81$, $df = 1{,}36$), $p < .02$). In

TABLE 1

MEAN AND STANDARD DEVIATIONS OF DEPENDENT VARIABLES ACROSS TIME-ON-TASK (REPEATED MEASURE)

	(B A S E L I N E)			(T A S K P E R I O D S)		
D.V.	5'	10'	15'	20'	25'	30'
EMG	$\bar{X}$=2.96 SD=1.18	$\bar{X}$=2.91 SD=1.31	$\bar{X}$=2.87 SD=1.57	$\bar{X}$=4.70 SD=2.22	$\bar{X}$=4.57 SD=2.12	$\bar{X}$=4.38 SD=1.83
TIME LOAD	$\bar{X}$=N/A SD=N/A	$\bar{X}$=N/A SD=N/A	$\bar{X}$=N/A SD=N/A	$\bar{X}$=3.25 SD=1.21	$\bar{X}$=2.95 SD=1.18	$\bar{X}$=2.84 SD=1.27
MENTAL EFFORT LOAD	$\bar{X}$=N/A SD=N/A	$\bar{X}$=N/A SD=N/A	$\bar{X}$=1.45 SD=0.73	$\bar{X}$=4.35 SD=1.29	$\bar{X}$=4.15 SD=1.27	$\bar{X}$=3.98 SD=1.44
GPS LOAD	$\bar{X}$=N/A SD=N/A	$\bar{X}$=N/A SD=N/A	$\bar{X}$=3.10 SD=0.91	$\bar{X}$=2.77 SD=0.80	$\bar{X}$=2.91 SD=0.82	$\bar{X}$=3.05 SD=0.93

N/A= not applicable

TABLE 2

SUMMARY ANALYSIS OF VARIANCE WITH REPEATED MEASURES FOR THE EFFECT OF ASSIGNED GOAL STRATEGY ON THE GENERAL PSYCHOLOGICAL STRESS DIMENSION OF THE WORKLOAD MEASUREMENT SCALE

Source of Variation	Sum of Squares	df	Mean Squares	F
Between Subjects				
Time Strategy (A)	.45	1	.45	< 1
Accuracy Strategy (B)	.02	1	.02	< 1
A X B	3.25	1	3.25	1.46
Error-Between	79.93	36	2.22	
Within Subjects				
Time-on-Task (R)	2.74	3	.91	3.17 *
A X R	.37	3	.12	< 1
B X R	1.90	3	.64	2.20
A X B X R	.22	3	.07	< 1
Error-Within	31.16	108	.29	
Total	120.04	159		

* $p < .05$

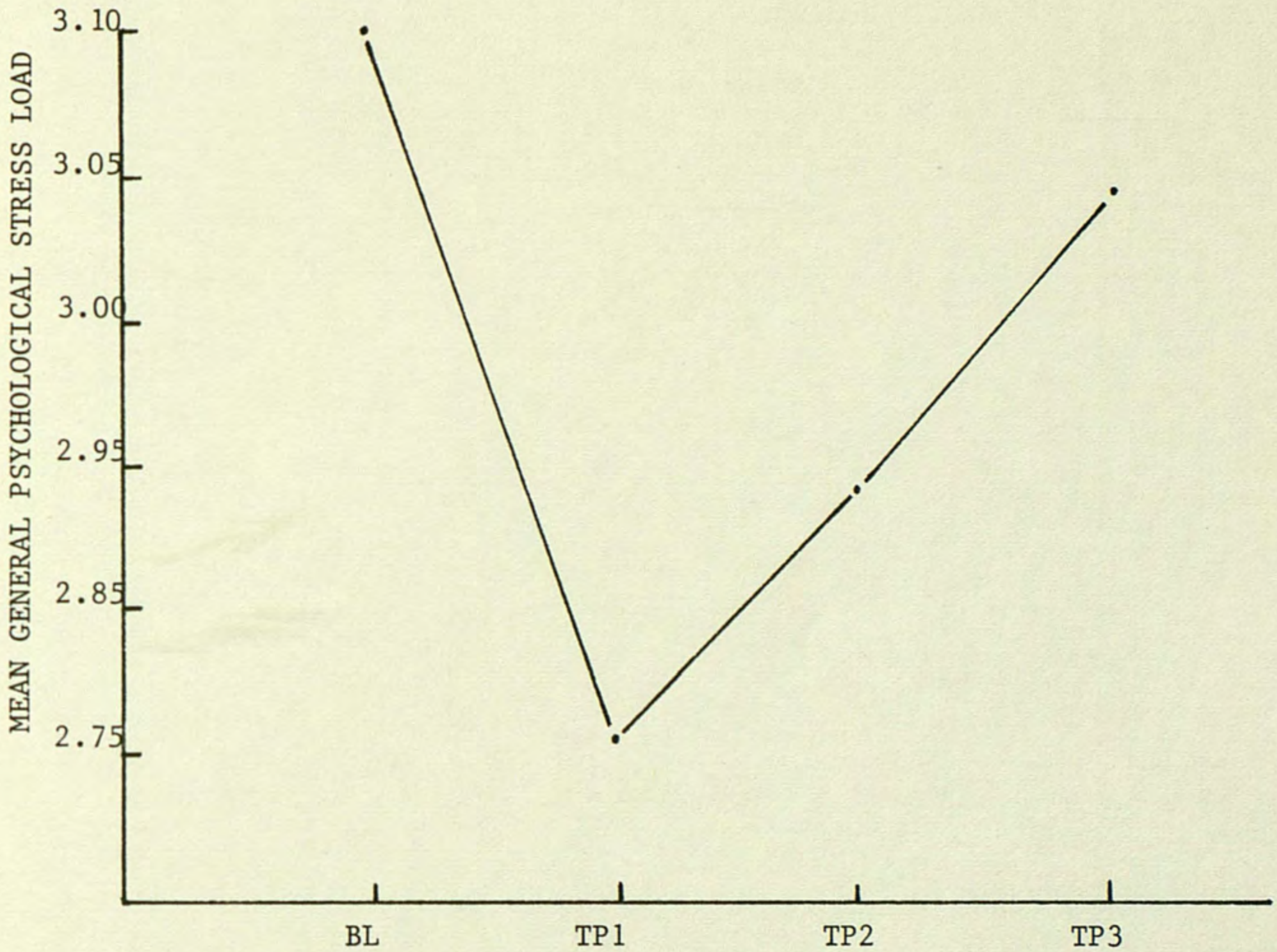

Figure 1. Mean General Psychological Stress (GPS) Load Dimension From the WMS Across Baseline and the Three Task Periods (four levels)

addition, analysis of trend indicated that as the subject worked at the task GPS load increased in a linear fashion ($F = 5.66$, $df = 1,36$, $p < .05$). Interestingly enough, the final GPS load rating was not significantly different from the previous baseline rating.

Table 3 reveals that time-on-task also had an effect on the time load dimension of the WMS ($F = 4.20$, $df = 2,72$), $p < .01$). Figure 2 and post hoc trend analysis indicated a linear decrease ($F = 6.89$, $df = 1,36$, $p < .02$) in time load as the subject worked on the cognitive task.

Table 4, similarly, indicates that time-on-task had an effect on the mental effort load dimension of the WMS ($F = 99.92$, $df = 3,108$, $p < .001$). Moreover, the mean task rating of this dimension was higher than baseline ($F = 160.31$, $df = 1,36$, $p < .001$), and Figure 3 shows an apparently linear decrease in mental effort load as the task progressed ($F = 3.96$, $df = 1,36$, $p = .05$).

Finally, Table 5 shows that time-on-task also had an effect on the physiological variable (EMG) under study ($F = 27.89$, $df = 5,180$, $p < .001$). Post hoc comparisons revealed that there were no differences between the three baseline readings (see Figure 4); however, the mean EMG reading during the task was greater than mean EMG during baseline ($F = 42.37$, $df = 1,36$, $p < .001$). In addition, EMG decreased in a linear fashion during the task periods ($F = 3.65$, $df = 1,36$, $p = .06$).

TABLE 3

SUMMARY ANALYSIS OF VARIANCE WITH REPEATED MEASURES FOR THE EFFECT OF ASSIGNED GOAL STRATEGY ON THE TIME LOAD DIMENSION OF THE WORKLOAD MEASUREMENT SCALE

Source of Variation	Sum of Squares	df	Mean Squares	F
Between Subjects				
Time Strategy (A)	10.50	1	10.50	2.84
Accuracy Strategy (B)	1.10	1	1.10	< 1
A X B	2.00	1	2.00	< 1
Error-Between	133.31	36	3.70	
Within Subjects				
Time-on-Task (R)	4.20	2	2.10	5.69 *
A X R	.12	2	.06	< 1
B X R	.02	2	.01	< 1
A X B X R	.22	3	.07	< 1
Error-Within	26.57	72	.37	
Total	178.08	119		

* $p < .01$

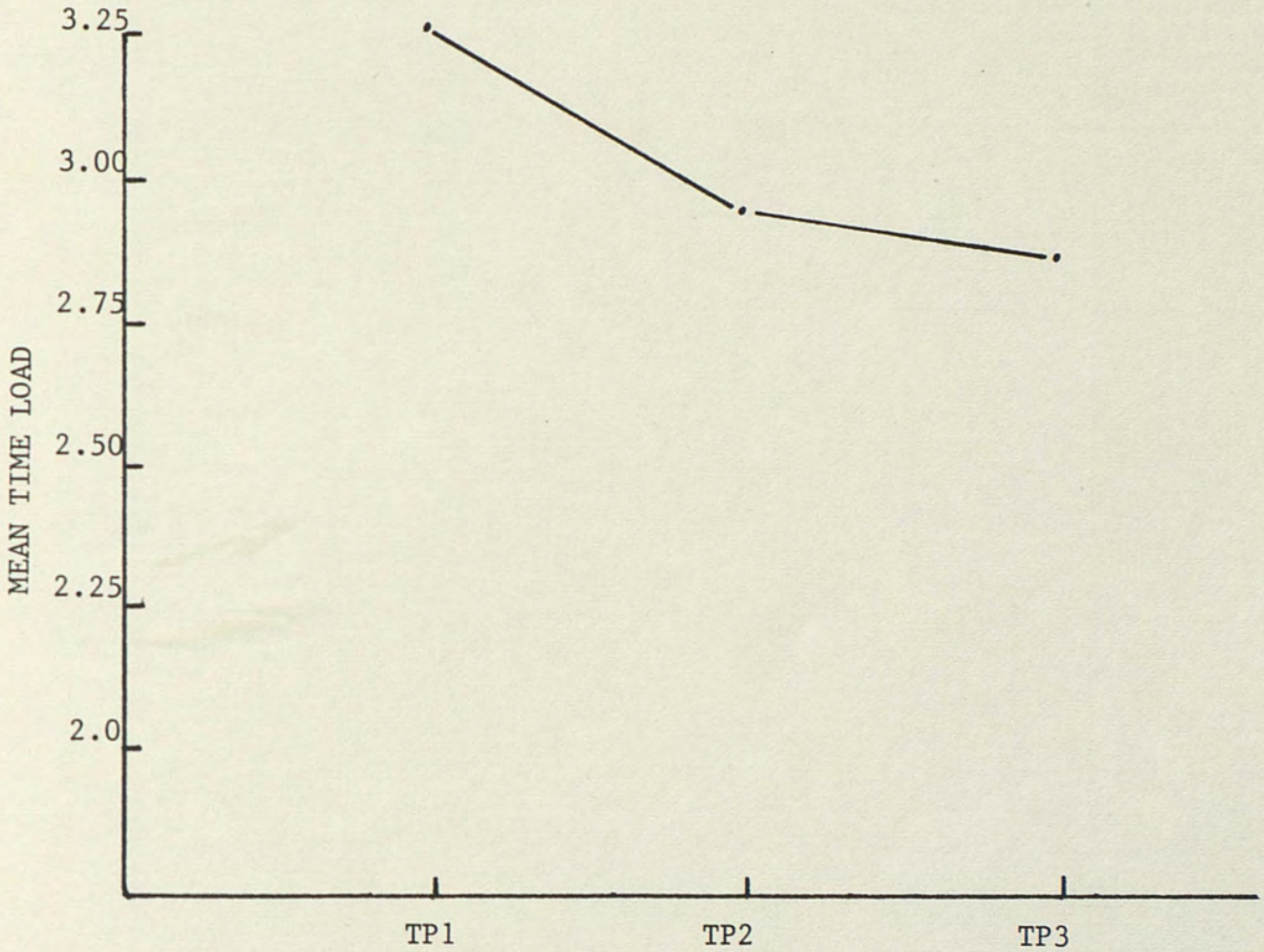

Figure 2. Mean Time Load Dimension From the WMS Across the Three Task Periods (three levels)

TABLE 4

SUMMARY ANALYSIS OF VARIANCE WITH REPEATED MERASURES FOR THE EFFECT OF ASSIGNED GOAL STRATEGY ON THE MENTAL EFFORT LOAD DIMENSION OF THE WORKLOAD MEASUREMENT SCALE

Source of Variation	Sum of Squares	df	Mean Squares	F
Between Subjects				
Time Strategy (A)	.31	1	.31	< 1
Accuracy Strategy (B)	2.26	1	2.26	< 1
A X B	8.56	1	8.56	2.32
Error-Between	132.95	36	3.69	
Within Subjects				
Time-on-Task (R)	222.87	3	74.29	99.92 *
A X R	1.72	3	.57	< 1
B X R	2.92	3	.97	1.31
A X B X R	.57	3	.19	< 1
Error-Within	80.30	108	.74	
Total	452.46	159		

* $p < .001$

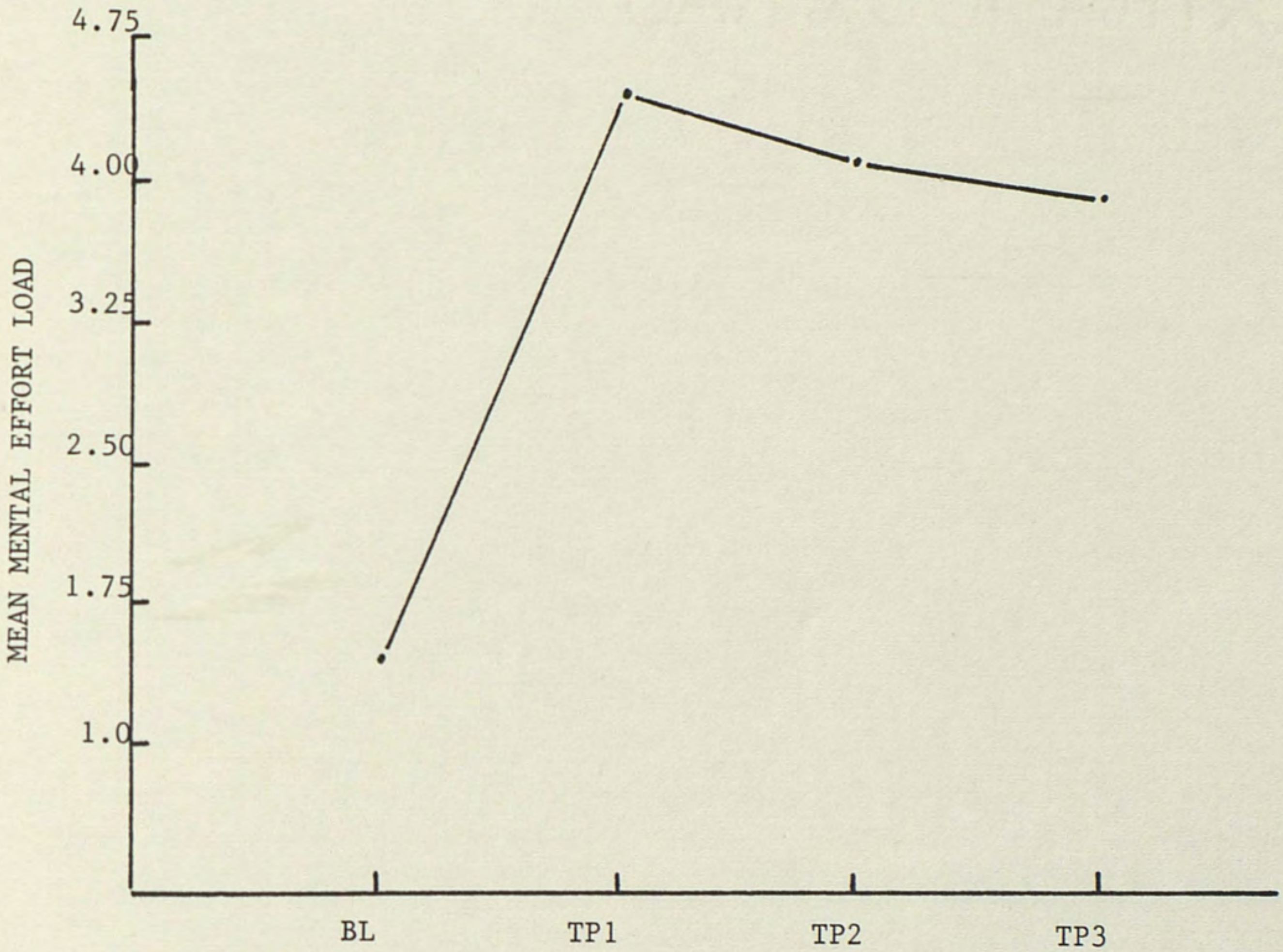

Figure 3. Mean Mental Effort Load Dimension From the WMS Across Baseline and the Three Task Periods (four levels)

TABLE 5

SUMMARY ANALYSIS OF VARIANCE WITH REPEATED MEASURES FOR THE EFFECT OF ASSIGNED GOAL STRATEGY ON ELECTROMYOPOTENTIAL

Source of Variation	Sum of Squares	df	Mean Squares	F
Between Subjects				
Time Strategy (A)	3.08	1	3.08	< 1
Accuracy Strategy (B)	.11	1	.11	< 1
A X B	16.57	1	16.57	1.27
Error-Between	470.59	36	13.07	
Within Subjects				
Time-on-Task (R)	162.81	5	32.56	27.89 *
A X R	4.76	5	.95	< 1
B X R	5.30	5	1.06	< 1
A X B X R	4.46	5	.89	< 1
Error-Within	210.11	180	1.17	
Total	877.79	239		

* $p < .001$

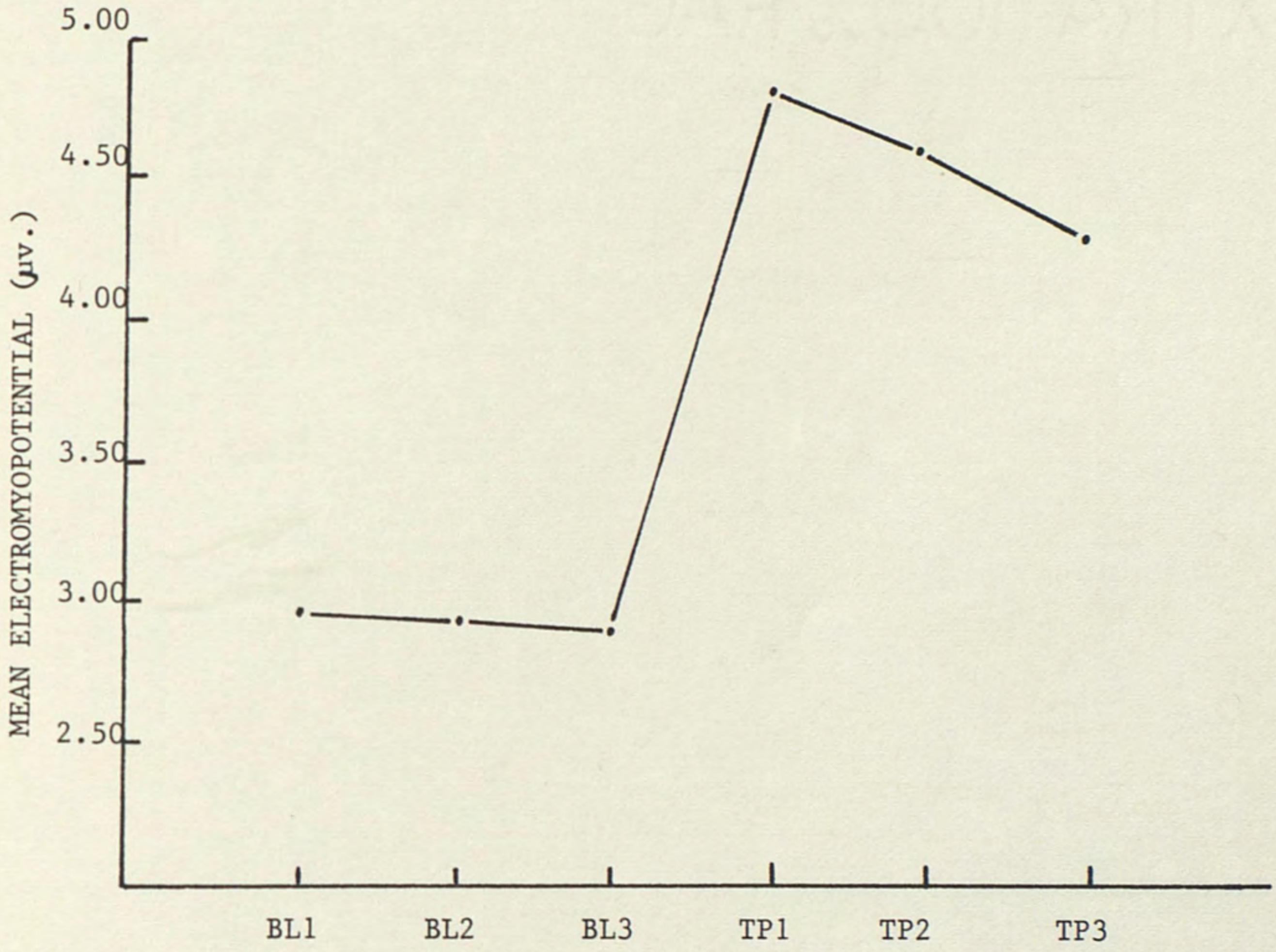

Figure 4. Mean Electromyopotential Across the Three Baseline and Three Task Periods (six levels)

Tables 6-9 show the intercorrelations between dependent variables during baseline and during the three task periods. There appeared no correlations during baseline; however during the first task period EMG and time load were related (r= -.4184, $p < .01$) as was GPS load and mental effort load (r= .4718, $p < .01$). During task period 2, EMG and time load remained correlated (r= -.3408, $p < .05$); however, the comparison of this coefficient with the one from task period 1 revealed no difference. In addition, the significant correlation between GPS load and mental effort load that was present during task period 1 failed to appear during task period 2. Finally, task period 3 witnessed no intercorrelations.

TABLE 6

CORRELATION MATRIX OF EMG, MENTAL EFFORT LOAD, AND GENERAL PSYCHOLOGICAL STRESS LOAD DURING BASELINE

	EMG	MENEFF	PSTRESS
EMG	1.00	--	--
MENEFF	.0461	1.00	--
PSTRESS	-.1711	.2408	1.00

Note. EMG= electromyography, MENEFF= mental effort load, and PSTRESS= general psychological stress load.

TABLE 7

CORRELATION MATRIX OF EMG, TIME LOAD, MENTAL EFFORT LOAD AND GENERAL PSYCHOLOGICAL STRESS LOAD DURING TASK PERIOD 1

	EMG	TIME	MENEFF	PSTRESS
EMG	1.00	--	--	--
TIME	-.4184 *	1.00	--	--
MENEFF	-.0705	.1026	1.00	--
PSTRESS	-.1259	.0764	.4718 *	1.00

* $p < .01$

Note. EMG= electromyography, TIME= time load, MENEFF= mental effort load, and PSTRESS= general psychological stress load.

TABLE 8

CORRELATION MATRIX OF EMG, TIME LOAD, MENTAL EFFORT LOAD, AND GENERAL PSYCHOLOGICAL STRESS LOAD DURING TASK PERIOD 2

	EMG	TIME	MENEFF	PSTRESS
EMG	1.00	--	--	--
TIME	-.3408 *	1.00	--	--
MENEFF	.0240	-.0930	1.00	--
PSTRESS	.0706	-.0291	.2260	1.00

* $p < .05$

Note. EMG= electromyography, TIME= time load, MENEFF= mental effort load, and PSTRESS= general psychological stress load.

TABLE 9

CORRELATION MATRIX OF EMG, TIME LOAD, MENTAL EFFORT LOAD, AND GENERAL PSYCHOLOGICAL STRESS LOAD DURING TASK PERIOD 3

	EMG	TIME	MENEFF	PSTRESS
EMG	1.00	--	--	--
TIME	-.2167	1.00	--	--
MENEFF	-.0060	-.0267	1.00	--
PSTRESS	.0475	-.1797	.0873	1.00

Note. EMG= electromyography, TIME= time load, MENEFF= mental effort load, and PSTRESS= general psychological stress load.

DISCUSSION

For the purpose of this research, the WMS dimension reliabilities were deemed satisfactory even though two of the three dimensions had internal consistencies of less than .70; however, Nunnally (1978) has argued that .50 and .60 reliabilities will suffice for exploratory research. The lower reliabilities of the mental effort dimension could be attributed to the fact that this dimension only had two items. Also, since the GPS dimension consisted of many items which may or may not have been task specific, this could explain a lower reliability coefficient. Ironically, stress as a research topic shares this same problem.

Based on the results of this study there are no significant workload differences (either physiological or subjective) that are functions of assigned task performance strategies. This can be attributed to at least three phenomena. First, the strategies, even though they seemed appropriate for the task, were still general in nature. The subjects were not given any specific or target goal such as: "...answer each question within 3 seconds..."; or "...achieve a 97 percent correct response rate..." By using this more specific method, goals may then produce amounts of workload in proportion to the perceived requirements of the task. This is an area of future research. Second, the lack of variablity between groups could possibly be attributed to the task itself. That is, it just was not difficult enough (for long enough) to allow differences to appear. A third explanation for this finding is that subjects were given no feedback as to their performance. This was primarily due to limitations in the software. Therefore, subjects may have abandoned their assigned strategy and just tried to do their best simply because they received no cues as to their level of performance.

A second hypothesis stated that across a repeated measure such as time-on-task, all workload indices will show practice or habituation effects. This conjecture was supported in three out of four cases. EMG, time load, and mental effort load decreased in each successive task period. Although time load showed a significant linear decrease across the task, EMG and mental effort load graphically (see Figures 3 and 4) showed decreases during task, but linear trend analysis produced $p = .06$ and $p = .05$ respectively. General psychological stress load, however, decreased from baseline to task period 1 then increased in a significant linear fashion across the remaining two task periods. Subjects evidently viewed baseline as producing more non-specific stress or anxiety then when the requirements and nature of the task were introduced and subjects were asked to perform the task. This finding was not expected. The value of a repeated measure design becomes apparent because there would have appeared no difference between baseline and the final task period rating on GPS load if this dimension was rated only after the task was completed.

The use of EMG as an physiological index of workload must be tempered with a few caveats. First, there are relatively unlimited muscle sites on which to record EMG. The question centers on the frontalis muscle and its ability to reflect global muscle tension or muscle activity. Basmajian (1979) states that electrode placement on the frontalis may provide an index of muscle activity that also includes "...repeated swallowing, breathing, movements of the jaw, tongue, lips, eyelids and eyeballs rather than real myopotential originating from the frontalis muscle" (p. 152). In any case, it would be wise to use the frontalis muscle rather than a muscle site that could generate erroneous readings simply because of task requirements that involve hand and arm movement (e.g., forearm extensor). Connally, Nelesen, Dieter, and Uliano (1983) discussed this when they presented a laboratory model for EMG research that included non-involved distal muscle sites as the recording sites of choice.

With regard to correlations between physiological and subjective workload variables, the value of a repeated measure design again becomes apparent. If the intercorrelations between dependent variables (used for this study) were examined after the task there would have appeared no relationships; however, that would not have accurately represented the pattern of correlations as the task progressed.

EMG and time load were inversely related during the first two task periods. As EMG decreased, subjective feelings of time load increased. This relationship is odd and future research is needed before any serious implications are developed. It could very well be that feelings of being pressed for time cause increased concentration and attention which produce physiological responses that involve, at least in part, reduced skeletal muscle activity during early stages of novel task performance.

Feelings of general psychological stress and task-specific mental effort were significantly related only during the first task period. It is believed that this is the point at which workload is at its greatest.

This discussion concludes with a developmental note. The three dimensions that were used as subjective indices of workload are not etched in stone. Patterns of intercorrelations between items suggest that the dimensions could possibly be collapsed to include only stress load as a byproduct of task performance, and resource load which requires human resources such as attention and effort. Factor analytic research is needed in this area. Also, in an effort to keep the subjective workload instruments non-intrusive, the number of items must be kept to a minimum. This is also a valid rationale to accept lower reliability coefficients.

CONCLUSION

Workload is a dynamic evolving concept which intuitively is related to the requirements of the task. In a repetitive decision making task, workload is at its greatest during the onset of the task and then generally decreases across time. Other more active psychomotor tasks will most likely produce different patterns. It is not enough to ask how much workload a task produces; rather, at which point(s) during task acquisition is workload at a high enough level to warrant additional training or practice. Using this methodology could identify the relative contributions of specific situational components of the task to overall workload assessment. Future research should utilize this type of methodology and multivariate analysis is the design of choice when there is more than one dependent variable involved. This present study lacked a significant subject pool on which to perform this kind of analysis. This design could possibly show us the contributions of subjective and physiological variability to overall workload. Finally, since workload (and stress) appear very individual-specific, the use of subjective techniques continues to remain at the forefront in workload assessment.

APPENDIX A

INFORMED CONSENT FORM

UNIVERSITY OF CENTRAL FLORIA
DEPARTMENT OF PSYCHOLOGY

INFORMED CONSENT

The purpose of this study is to observe two kinds of responses that occur when people perform a cognitive task. The first type of response is physiological. These data will be gathered by an EMG machine which records the electrical activity of muscles. In this case, EMG electrodes will be attached to your forehead. The second type of response is subjective and will require you to complete a short paper-and-pencil rating form at four different times during the session. The task is computer-based and as such requires you to interact with an Apple computer.

- I understand that I am free to withdraw my consent and terminate my participation at any time, without penalty.
- I understand that I am free to withhold any answers to specific items or questions.
- I understand that any data or answers to questions will remain confidential with regard to my identity.

Your signature below acknowledges that you have read and understand the above and are willing to participate in this study.

SIGNATURE ______________
DATE ______________
AGE _______
SEX _______
SS# ____________

APPENDIX B

TASK INSTRUCTIONS

SCREEN #1

COMPARING VISUAL AND SEMANTIC
INFORMATION

THIS TASK REQUIRES THAT YOU EXAMINE A PICTURE MADE UP OF TWO FIGURES:

*

+

THE TOP FIGURE (*) IS CALLED A "STAR", AND THE BOTTOM FIGURE (+) IS CALLED A PLUS. THERE ARE SEVERAL WAYS TO INTERPRET THIS PICTURE; FOR EXAMPLE, "THE STAR IS ABOVE THE PLUS", "THE PLUS IS BELOW THE STAR", "THE PLUS IS NOT ABOVE THE STAR", ETC.

PRESS (N)EXT TO SEE MORE

SCREEN #2

FIRST, THE TASK REQUIRES YOU TO DETERMINE THE MEANING OF THE DISPLAYED PICTURE. NEXT, A SENTENCE SIMILAR TO THE ONES IN THE PREVIOUS PARAGRAPH WILL APPEAR, AND YOU MUST DECIDE IF THE SENTENCE IS TRUE OR FALSE.

IF THE SENTENCE IS TRUE, INDICATE THIS BY PRESSING THE "T" KEY. IF THE SENTENCE IS FALSE, INDICATE THIS BY PRESSING "F". DO NOT PRESS "RETURN". YOU WILL NOT BE TOLD IF YOUR ANSWERS ARE RIGHT OR WRONG.

THERE WILL BE FIVE PRACTICE TRIALS. YOU WILL THEN BE GIVEN AN OPTION OF REVIEWING THE INSTRUCTIONS OR CONTINUING ON TO THE TASK.

(N)EXT TO PRACTICE OR (L)AST TO REVIEW

APPENDIX C

SAMPLE TASK FRAMES

SCREEN #1

+
*

The * is called a star
The + is called a plus

When you understand the picture,
Press the space bar

SCREEN #2

+
*

The star is not below the plus

Press 'T' or 'F' ONLY

APPENDIX D

WORKLOAD MEASUREMENT SCALE FOR BASELINE

S#____
Condition ___
Sect. Code _BL_

DIRECTIONS: Carefully read and answer each question. Choose the number which best represents your feelings by circling the appropriate number.

---

1. The amount of attention that this part of the task required was:

```
+----------+----------+----------+----------+----------+----------+
1          2          3          4          5          6          7
*                                *                                *
not very                      moderate                        extreme
much                          amount of                       amount of
attention                     attention                       attention
```

2. The effort required to perform this part of the task was:

```
+----------+----------+----------+----------+----------+----------+
1          2          3          4          5          6          7
*                                *                                *
not very                      moderate                        extreme
much effort                   amount of                       amount of
                              effort                          effort
```

3. To what extent did you understand the nature of the task:

```
+----------+----------+----------+----------+----------+----------+
1          2          3          4          5          6          7
*                                *                                *
very clearly                 moderately                       not at
understood                   understood                         all
```

4. To what extent were you afraid you would fail at performing this part of the task:

```
+----------+----------+----------+----------+----------+----------+
1          2          3          4          5          6          7
*                                *                                *
not at all                   moderately                     extremely
afraid                         afraid                         afraid
```

5. To what extent did you feel "anxious" or "uptight" during this part of the task:

```
+----------+----------+----------+----------+----------+----------+
1          2          3          4          5          6          7
*                                *                                *
not at                       moderately                     extremely
all                            anxious                        anxious
anxious
```

6. To what extent did you feel this part of the task was boring:

```
+----------+----------+----------+----------+----------+----------+
1          2          3          4          5          6          7
*                                *                                *
not at                       moderately                     extremely
all                            boring                          boring
boring
```

7. To what extent did you feel fatiqued or tired during this part of the task:

```
+----------+----------+----------+----------+----------+----------+
1          2          3          4          5          6          7
*                                *                                *
not at                       moderately                     extremely
all                           fatiqued                       fatiqued
fatiqued
```

8. To what extent do you feel comfortable interacting with computers:

```
+---------+---------+---------+---------+---------+---------+
1         2         3         4         5         6         7
*                             *                             *
extremely                  moderately                    not at all
comfortable                comfortable                   comfortable
```

APPENDIX E

WORKLOAD MEASUREMENT SCALE FOR TASK PERIODS

S#____
Condition ___
Sect. Code ___

DIRECTIONS: Carefully read and answer each question. Choose the number which best represents your feelings by circling the appropriate number.

---

1. I felt the overall time available to perform this part of the task was:

```
+---------+---------+---------+---------+---------+---------+
1         2         3         4         5         6         7
*                             *                             *
extremely                     adequate                      not at
adequate                                                    all
                                                            adequate
```

2. I felt that the time between individual presentations was:

```
+---------+---------+---------+---------+---------+---------+
1         2         3         4         5         6         7
*                             *                             *
extremely                     adequate                      not at
adequate                                                    all
                                                            adequate
```

3. Regardless of the time available, the amount of attention that this part of the task required was:

```
+---------+---------+---------+---------+---------+---------+
1         2         3         4         5         6         7
*                             *                             *
not very                      moderate                      extreme
much                          amount of                     amount of
attention                     attention                     attention
```

4. Regardless of the time available, the effort required to perform this part of the task was:

```
+---------+---------+---------+---------+---------+---------+
1         2         3         4         5         6         7
*                             *                             *
not very                      moderate                      extreme
much effort                   amount of                     amount of
                              effort                        effort
```

5. To what extent did you understand the nature of the task:

```
+---------+---------+---------+---------+---------+---------+
1         2         3         4         5         6         7
*                             *                             *
very clearly                  moderately                    not at
understood                    understood                    all
```

6. To what extent were you afraid you would fail at performing this part of the task:

```
+---------+---------+---------+---------+---------+---------+
1         2         3         4         5         6         7
*                             *                             *
not at all                    moderately                    extremely
afraid                        afraid                        afraid
```

7. To what extent did you feel "anxious" or "uptight" during this part of the task:

```
+---------+---------+---------+---------+---------+---------+
1         2         3         4         5         6         7
*                             *                             *
not at                        moderately                    extremely
all                           anxious                       anxious
anxious
```

8. To what extent did you feel theis part of the task was boring:

```
+----------+----------+----------+----------+----------+----------+
1          2          3          4          5          6          7
*                                *                                *
not at                       moderately                      extremely
all                            boring                          boring
boring
```

9. To what extent did you feel fatiqued or tired during this part of the task:

```
+----------+----------+----------+----------+----------+----------+
1          2          3          4          5          6          7
*                                *                                *
not at                        moderately                     extremely
all                            fatiqued                       fatiqued
fatiqued
```

10. To what extent do you feel comfortable interacting with computers:

```
+----------+----------+----------+----------+----------+----------+
1          2          3          4          5          6          7
*                                *                                *
extremely                     moderately                    not at all
comfortable                   comfortable                   comfortable
```

APPENDIX F

STRATEGY BY GROUP

Group 1- Time/Accuracy Strategy

"Please read and answer each question as quickly and as accurately as possible. The computer will be recording your responses and how long it took you to make them."

Group 2- Time Strategy

"Please read and answer each question as quickly as possible. The computer will be timing your responses."

Group 3- Accuracy Strategy

"Please take as much time as you need; answer each question as accurately as possible. The computer will be recording your responses."

Group 4- No Strategy

"Please read and answer each question."

REFERENCES

Alluisi, E. A. (1982). Stress and stressors, commonplace and otherwise. In E. A. Allusi & E. A. Fleishman (Eds.), Human performance and productivity: Vol. 3: Stress and performance effectiveness (pp. 1-10). Hillsdale, NJ: Lawrence Erlbaum Associates.

Alluisi, M. J., DeGroot, S. G., & Alluisi, E. A. (Eds.). (1984). Proceedings of the Human Factors Society 28th Annual Meeting. (Vols. 1,2). Santa Monica, CA: Human Factors Society.

Anastasi, A. (1979). Fields of applied psychology (2nd ed.). New York: McGraw-Hill.

Averill, J. R. (1973). Personal control over aversive stimuli and its relationship to stress. Psychological Bulletin, 80, 286-303.

Basmajian, J. V. (1979). Facts vs. myths in EMG biofeedback. In Biofeedback and self-control 1977/78: An Aldine annual on the regulation of bodily processes and consciousness (pp. 151-153). Chicago: Aldine.

Borg, G. (1978a). Subjective aspects of physical and mental load. Ergonomics, 21, 215-220.

Borg, G. (1978b). Subjective effort in relation to physical performance and working capacity. In H. Pick (Ed.). Psychology from research to practice (pp. 331-361). New York: Plenum Press.

Borg, G. (1970). Perceived exertion as an indicator of somatic stress. Scandinavian Journal of Rehabilitation Medicine, 2, 92-98.

Borg, G., Bratfisch, O., & Dornic, S. (1971a). Perceived difficulty of a visual search task (Report No. 16). Stockholm: Institute of Applied Psychology.

Borg, G., Bratfisch, O., & Dornic, S. (1971b). On the problem of perceived difficulty. Scandinavian Journal of Psychology, 12, 249-260.

Borg, G., & Noble, B. J. (1974). Perceived exertion. In J. Wilmore (Ed.), Exercise and sport science reviews (pp. 131-153). New York: Academic Press.

Casali, J. G., & Wierwille, W. W. (1983). A comparison of rating scale, secondary task, physiological, and primary-task workload estimation techniques in a simulated flight task emphasizing communication load. Human Factors, 25, 623-641.

Chiles, W. D. (1982). Workload, tasks, and situational factors as modifiers of complex human performance. In E. A. Alluisi & E. A. Fleishman (Eds.), Human performance and productivity: Vol. 3: Stress and performance effectiveness (pp. 11-56). Hillsdale, NJ: Lawrence Erlbaum Associates.

Connally, R. E., Nelesen, R. A., Dieter, J. N., & Uliano, K. C. (1983). Three recurring electromyographic biofeedback research problems and a laboratory model. Perceptual and Motor Skills, 57, 1059-1069.

Cooper, G. E. (1957). Understanding and interpreting pilot opinion. Aeronautical Engineering Review, 16, 47-52.

Cooper, G. E., & Harper, R. P. (1969). The use of pilot ratings in the evaluation of aircraft handling characteristics (Report No. TN-D-5153). Washington, DC: National Aeronautics and Space Administration.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. Psychometrica, 16, 297-334.

Eggemeier, F. T. (1984). Workload metrics for system evaluation. Unpublished manuscript.

Eggemeier, F. T. (1980). Some current issues in workload measurement. Proceedings of the 24th Human Factors Society Annual Meeting. 669-673.

Eggemeier, F. T., Crabtree, M. S., Zingg, J. J., Reid, G. B., & Shingledecker, C. A. (1982). Subjective workload assessment in a memory update task. Proceedings of the 26th Human Factors Society Meeting, 643-647.

Eggleston, R. G., & Quinn, T. J. (October, 1984). A preliminary evaluation of a projective workload assessment procedure. Paper presented at the 28th Human Factors Society Annual Meeting, San Antonio, TX.

Fisk, A. D., Derrick, W. L., & Schneider, W. (1983). The assessment of workload: Dual task methodology. Proceedings of the 27th Human Factors Society Annual Meeting, 229-233.

Gartner, W. P., & Murphy, M. R. (1976). Pilot workload and fatigue: A critical survey of concepts and assessment techniques (Report No. NASA-TN-D-8365). Washington, DC: National Aeronautics and Space Administration.

Gopher, D., & Braune, R. (1984). On the psychophysics of workload: Why bother with subjective measures? Human Factors, 26, 519-532.

Hacker, W., Plath, H. E., Richter, P., & Zimmer, K. (1978). Internal representations of task structures and mental load of work: Approaches and methods of assessment. Ergonomics, 21, 194-197.

Hays, W. L. (1981). Statistics (3rd ed.). New York: Holt, Rinehart, and Winston.

Hicks, T., & Wierwille, W. W. (1979). Comparison of five mental workload assessment procedures in a moving-base simulator. Human Factors, 21, 129-144.

Hogan, R., & Hogan, J. C. (1982). Subjective correlates of stress and human performance. In E. A. Alluisi & E. A. Fleishman (Eds.), Human performance and productivity: Vol. 3: Stress and performance effectiveness (pp. 141-163). Hillsdale, NJ: Lawrence Erlbaum Associates.

Johanssen, G., Moray, N., Pew, R., Rasmussen, J., Sanders, A., & Wickens, C. D. (1979). Final report of the experimental psychology group. In N. Moray (Ed.), Mental workload: Its theory and measurement. New York: Plenum Press.

Knowles, W. B. (1963). Operator loading tasks. Human Factors, 5, 155-161.

Kramer, A. F., Wickens, C. D.. & Donchin, E, (1983). An analysis of the processing requirements of a complex perceptual motor task. Human Factors, 25, 597-621.

Lazarus, R. S., Deese, J., & Osler, S. F. (1952). The effects of psychological stress upon performance. Psychological Bulletin, 49, 293-317.

Leplat, J. (1978). Factors determining workload. Ergonomics, 21, 143-149.

Lester, J. T. (1979). An annotated bibliography of DoD-funded reports concerning pshchological stress: 1958-1978. Boston, MA: Office of Naval Research.

Likert, R. A. (1932). A technique for the measurement of attitudes. Archives of Psychology, 140, 44-53.

Lippold, O. (1967). Electromyography. In P. Venables & I. Martin (Eds.), A manual of psychophysiological methods (pp. 245-299). New York: Wiley.

Locke, E. A., & Bryan, J. F. (1969). Cognitive aspects of psychomotor performance: The effects of performance goals on level of performance. Journal of Applied Psychology, 50, 286-291.

Locke, E. A., Shaw, K. N., Saari, L. M., & Latham, G. P. (1981). Goal setting and task performance: 1969-1980. Psychological Bulletin, 90, 125-152.

McDonnell, J. D. (1969). An application of measurement methods to improve the quantitative nature of pilot rating scales. IEEE Transactions on Man-Machine Systems, 10, 81-92.

McGrath, J. E. (Ed.). (1970). Social and psychological factors in stress. New York: Holt, Rinehart, and Winston.

Moray, N. (1982). Subjective mental workload. Human Factors, 24, 143-149.

Moray, N. (Ed.). (1979). Mental workload: Its theory and measurement. New York: Plenum Press.

North, R. A., & Graffunder, K. (1979). Evaluation of a pilot workload metric for simulated VTOL landing tasks. Proceedings of the Human Factors 23rd Annual Meeting, Boston, MA.

Notestine, J. (1983). Subjective workload assessment in a probability monitoring task and the effect of delayed ratings. Unpublished master's thesis, Wright State University, Dayton, Ohio.

Nunnally, J. C. (1978). Psychometric theory (2nd ed.). New York: McGraw- Hill.

Ogden, G. D., Levine, J. W., & Eisner, E. J. (1979). Measurement of workload by secondary tasks. Human Factors, 21, 660-673.

Pervin, L. A. (1968). Performance and satisfaction as a function of individual-environment fits. Psychological Bulletin, 69, 56-68.

Pew, R. W. (1979). Secondary tasks and workload measurement. In N. Moray (Ed.), Mental workload: Its theory and measurement (pp. 23-28). New York: Plenum Press.

Rasmussen, J. (1979). Reflection on the concept of operator workload. In N. Moray (Ed.), Mental Workload: Its theory and measurement (pp. 29- 40). New York: Plenum Press.

Rehmann, J. T., Stein, E. S., & Rosenberg, B. L. (1983). Subjective pilot workload assessment. Human Factors, 25, 297-307.

Reid, G. B., Shingledecker, C. A., & Eggemeier, F. T. (1981). Application of conjoint measurement to workload scale development. Proceedings of the 25th Human Factors Society Meeting, 522-526.

Reid, G. B., Shingledecker, C. A., Nygren, T. E., & Eggemeier, F. T. (1981). Development of multidimensional subjective measures of workload. Proceedings of the 1981 IEEE International Conference on Cybernetics and Society, 403-406.

Rolfe, J. M. (1976). The measurement of human response in man-machine vehicle control situations. In T.Sheridan & G.Johannsen (Eds.), Monitoring behavior and supervisory control. New York: Plenum Press.

Roll, N. (1981). The effect of machine pacing of simulated inspection tasks on physiological and psychological stress reactions. Unpublished master's thesis, University of Central Florida, Orlando.

Selye, H. (1956). The stress of life. New York: McGraw-Hill.

Senders, J. W. (1970). The estimation of operator workload in complex systems. In K. B. DeGreen (Ed.), Systems psychology. New York: McGraw-Hill.

Stamford, B. A. (1976). Validity and reliability of subjective ratings of perceived exertion during work. Ergonomics, 19, 53-60.

Trotter, P. (1980). Comparing visual and semantic information [Computer program]. Iowa City, IA: Conduit. (National Science Foundation No. 77-2036; Fund for the Improvement of Post-Secondary Education Grant No. G007905216).

Uliano, K. C., & Carey, J. R. (1983). MBERT: A BASIC program to alleviate experimenter effect in biofeedback research. Perceptual and Motor Skills, 58, 206.

Wardle, M. G. (1978). A psychological approach to estimating endurance in performing physically demanding work. Human Factors, 20, 745- 747.

Wierwille, W. W. (1979). Physiological measures of aircrew mental workload. Human Factors, 21, 575-593.

Wierwille, W. W., & Casali, J.G. (1983). A validated rating scale for global mental workload measurement applications. Proceedings of the 27th Human Factors Society Meeting, 129-133.

Wierwille, W. W., & Connor, S. A. (1983). Evaluation of 20 workload measures using a psychomotor task in a moving-base aircraft simulator. Human Factors, 25, 1-16.

Wierwille, W. W. & Williges, R. C. (1978). Survey and analysis of operator workload assessment techniques (Report No. S-78-101). Blacksburg, VA: Systematics, Inc.

Wilkens, W. L. (1982). Psychophysiological correlates of stress and human performance. In E. A. Alluisi & E. A. Fleishman (Eds.). Human performance and productivity: Vol. 3: Stress and performance effectiveness (pp. 57-90). Lawrence Erlbaum Associates.

Williges, R. C., & Wierwille, W. W. (1979). Behavioral measures of aircrew mental workload. Human Factors, 21, 549-574.