UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

EFFECT OF WORKLOAD HISTORY
ON VIGILANCE PERFORMANCE

A Dissertation

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

degree of

Doctor of Philosophy

By

LUZ-EUGENIA COX-FUENZALIDA

Norman, Oklahoma

2000

UMI Number: 9985572

THE EFFECT OF WORKLOAD HISTORY
AND INTROVERSION-EXTRAVERSION ON VIGILANCE PERFORMANCE

A Dissertation APPROVED FOR THE
DEPARTMENT OF PSYCHOLOGY

BY

Dr. Kirby Gilliland

Dr. Larry Toothaker

Dr. Scott Gronlund

Dr. Robert Schlegel

Dr. Cal Stoltenberg

Acknowledgments

I would like to express my appreciation to all of you who supported me through the completion of not only my doctoral dissertation but my graduate training, as well. I would like to thank the following members of my committee Drs. Larry Toothaker, Bob Schlegel, Scott Gronlund, and Cal Stoltenberg. Dr. Toothaker, I am very grateful for your taking the time to discuss methodology, statistical analyses and so much more. Thank you mostly for your never failing sound advice and for your confidence in my ability. Dr. Schlegel, thank you for your careful reading and thorough review of my drafts. Your trouble shooting, constructive criticism and helpful suggestions were invaluable. Dr. Stoltenberg, your comments and questions at meetings brought new perspective. Thank you especially for your continual words of encouragement and support.

In addition, my gratitude is extended to Dr. Scott Mills, Dr. Rhonda Swickert and Jennifer Perry who on many occasions provided both personal support and professional advice. Scott, thank you particularly for the many hours of programming and technical assistance. Rhonda and Jennifer, among other things thank you for always finding the time to listen to presentations and provide sincere and helpful feedback.

My undergraduate research assistants were instrumental in the smooth execution of this project. Particularly helpful was Shannon Brown's personal initiative and Tracy Hudson's attention to detail.

I would also like to thank my daughters, M. Cecilia and A. Gabriela for reminding me of the more important things. Cita, in this year of writing my dissertation you have taught me the art of smelling silk flowers, the value of picking up acorns and the beauty of wearing fifteen barrettes in my hair. Gabi, your calls were welcome breaks during long hours of analyses and writing. I am grateful to you both for introducing balance and providing perspective in my life.

Finally, to the two people in my life without whom this would not have been as enjoyable: my mentor, Dr. Kirby Gilliland and my husband, Dr. Robert Cox. Kirby, here's to the ideal mentor! Thank you for questioning, for listening, for providing perspective, for challenging me with alternative arguments, for helping me set deadlines (for knowing when to move and when to keep them), for teaching me how to question, for knowing and sometimes not (easily) giving up the answers. Above all, thank you for your friendship.

Finally, I believe I am most indebted to my life partner, Bob. I think it only fair to say that without you this would not have been possible. Thank you for believing when I doubted, for willingly stepping in and helping with whatever was needed—being both Mamá and Papá to our children, reading drafts and providing feedback, listening, offering opinions and advice. Thank you for your patience and for sacrificing so that I might succeed.

Table of Contents

List of Tables

List of Figures

Abstract

Effects of sudden transitions in workload level on performance indices were investigated within the newly emerging paradigm of workload history. It has been previously reported that a sudden decrease in workload level results in a significant immediate decrement in performance (Cumming & Croft, 1973; Goldberg & Stewart, 1980; Matthews, 1986). Study 1 consisted of 198 participants who completed a protocol of training, baseline and test trials designed to examine the effects of workload history on performance. Specifically, this study tested for the reported general decrement in performance and investigated the nature of the decrement. Results indicated that a sudden decrease in workload level results in an immediate significant decrement in response time, total errors, etc., while a sudden increase fails to result in the same immediate decrement. This study reports new findings, i.e., either a sudden increase or decrease could lead to a loss in accuracy and a slowing of response time in a longer time course. An explanation of the decrement is offered in terms of a resource adaptation-depletion model. Study 2 tested whether time on task or fatigue might have been responsible for the decrement in performance following sudden transitions in workload (reported in Study 1). No significant results were found; therefore, it appears that the decrement in performance following workload transitions is a result of something inherent in the workload shift rather than an effect of time on task or fatigue.

Effect of Workload History on Vigilance Performance

Most human task performance research has focused on determining task performance capability during a specific time period or exploring the influence of such factors as stress, sleep loss, drugs, individual differences and numerous other variables on task performance. Recent research has suggested that the type and duration of work performed prior to some point in time (that is, workload history) may have a strong influence on work performance following that point in time. However, little is known about the characteristics and dynamics of workload history that are important in influencing subsequent work performance. Yet, workload history could have significant implications for many work environments, particularly those where individuals are confronted with varying levels of workload demand. For example, consider the bus driver who suddenly enters heavy traffic, or the air traffic controller who suddenly has significantly fewer planes to manage within an air space. Both of these examples demonstrate situations where an individual's workload history may influence performance during a subsequent critical period of time in a safety sensitive occupation. It would therefore seem prudent to have a fuller understanding of workload history and its possible influence on later performance.

The purpose of this project was twofold. Study 1 of the present investigation examined a recently reported effect of workload history on performance. For purposes of this study, workload history refers to previous workload activity that might have an effect on subsequent task performance. A common example of this is when a person has been working at a high workload level for some period and then is shifted to a low workload level, or vice versa. Study 1 was structured to test theories of why decrements in performance occur in response to this type of change in workload. Study 2 was designed to determine whether the decrement in performance following a workload transition (in Study 1) was the result of time on task or fatigue. Specifically, Study 2 tested for a time on task or fatigue effect utilizing a workload history paradigm requiring participants to work at a fixed workload level following a training/baseline protocol identical to that of Study 1.

Study 1

Past experimental studies focusing on cognitive performance have typically controlled for changes in workload level (i.e., treating it as a between Ss factor). This has often been done in an effort to control for possible confounding variables or to explore effects at a stable workload level (Matthews, 1986). Although this method has enabled the study of individual responses during fixed workload levels, it has also resulted in little information pertaining to how individuals respond to dynamic workload situations that tend to be more representative of many real-world jobs. In fact, it is not inconceivable that this attention to the experimental control of workload (that has led to many advances in understanding the effects of fixed workload) has inadvertently suppressed the study of broader workload dynamics such as workload history. Nonetheless, a small number of studies have examined the consequences of workload variation on performance. One trend emerging from these studies has been the finding that a general decrement in performance is most likely to occur in situations where there is a decrease in task demand. Each of these studies has attempted to explain the nature of the decrement in a different way.

Workload History

In 1973, Cumming and Croft reported the effects of workload history on an auditory monitoring task. The rate of signal presentation or workload in their study was manipulated in a cyclical fashion. Specifically, participants performed the monitoring task as the difficulty level of the task was systematically raised and lowered in an alternating fashion. Results suggested a significantly greater performance decrement during the decreasing workload phase of the cycles compared with the increasing workload phase of the cycles. Although Cumming and Croft (1973) failed to operationally define performance decrements and/or errors, it appears that they made direct comparisons between performance during increasing and decreasing workload phases of the task.

Cumming and Croft (1973) proposed a complex explanation for this finding. Their explanation began by noting a previously reported relationship between performance effectiveness and stimulus frequency—namely that people tend to respond faster or more accurately to stimuli presented more frequently, compared to stimuli presented less frequently. Cumming and Croft also believed that, during the phases of the repeating cycles in their experiment when workload increased, their participants realized that they could not maintain uniformly high response quality, so they chose to accept higher levels of error as the task difficulty increased. Because it was difficult for their participants to know exactly when the cycles changed and the task began to get easier, the participants continued to accept more errors as they moved into the descending task difficulty phase of the workload cycle (i.e., a period when the task grew increasingly easier). This acceptance of higher error levels (which were functional during a difficult phase of task performance) simply compounded the previously known phenomena that people respond more poorly when signal frequency (i.e., task difficulty) decreases. As a result, Cumming and Croft believed that this inability to match expectancies regarding changes in task difficulty explained why their participants demonstrated a much larger performance decrement when they moved to a low workload level after performing at a high workload level, as compared to when they experienced an increase in task difficulty following a low workload level.

In 1980, Goldberg and Stewart reported a study designed to test whether expectancies were responsible for the decline in task performance when workload shifted from high to lower levels. They employed a visual monitoring task. One condition included a display characteristic that served as a visual cue signaling changes in task demand (i.e., characters appeared on the left of the screen with subsequent characters appearing progressively farther to the right and then back to the left as demand increased and decreased, respectively). The purpose of the cue was to inform participants that either an increase or decrease in task demand was taking place. Even when participants were given this cue to indicate the onset of a reduction in task demand, their performance was still worse when shifting from high to lower workload as compared to shifting from low to higher workload—suggesting that it is not a person's inaccurate expectations that are responsible for the effect. Thus, the study failed to provide support for the explanation based on expectancies provided by Cumming and Croft (1973) concerning performance decrement in the presence of workload variability. Rather, Goldberg and Stewart (1980) suggested that perhaps the decrements following a decrease in work demand resulted from a temporary overload of short-term memory rather than an individual's inaccurate expectancy of task demand.

In 1986, Matthews conducted a set of studies to further explore the effects of workload history on visual monitoring performance. The results of the Matthew's study provided further support for the findings of Cumming and Croft (1973) and Goldberg and Stewart (1980) that higher decrements in performance result when workload decreases as compared to when workload increases. In addition, Matthews attempted to test the theory that the performance decrement was due to a failure of short-term memory. Matthews used a visual monitoring task that did not require the preservation of serial order information (i.e., short-term memory). Specifically, participants were required to monitor matrix displays containing 3, 6, 9 or 12 target/noise strings for the occurrence of target signals. Target signals consisted of arithmetic expressions (e.g., 22 + 12 < 18) that had to be detected from non-arithmetic expressions (e.g., 27 - #6 > Y$). Participants had to first locate and indicate the position of a target signal (if present) by pressing an appropriate key, and then evaluate and indicate whether the expression was true or false by pressing another key. A blank screen was presented immediately after a participant's response. This resulted in 3.5 to 6.0 seconds between stimuli, making unlikely any carryover effect from one stimulus to the next, especially given the 100 to 250 ms limit of visual short-term memory. This methodology generally provided a situation in which memory capacity and demands were not considered significant factors. Matthews included one condition during which a high workload level suddenly decreased to a low level. Another condition included a low workload level that suddenly shifted to a high level. A significant decrement in performance was found up to one minute following a sudden workload reduction, but not following a sudden workload increase. Thus, the data failed to support an explanation of the performance decrement based on the short-term memory overload hypothesis provided by Goldberg and Stewart (1980).

Instead, Matthews proposed that strategic persistence might account for the performance decrement following a decrease in workload. He proposed that participants were able to mobilize effective strategies when workload increased, but then retained these strategies long after the workload level had fallen (i.e., when workload changed from higher to lower levels). For example, when suddenly reduced to a low workload level, participants might have continued the level of effort they applied during the previous period at high workload thereby overworking or overdriving the task. This theoretical model might be easily confused with that of Cumming and Croft (1973). It is important to note that while both may discuss strategies in some way, Cumming and Croft (1973) offer an explanation based on inability to match expectancies. They believed their participants were unable to tell when a decrease in workload was going to occur, so they incorrectly "allowed" themselves errors at the lower level. Matthews attributed the decrement to overworking or overdriving the task when it was reduced to a lower workload level.

In summary, all three studies examined the effects of workload history on performance. These studies clearly demonstrated the ability to document the influence of work history on subsequent work performance. They also have identified a performance phenomenon that appears to occur fairly reliably. Specifically, they identified a general decrement in performance in a situation where an individual moves to a low workload level after working for some period at a higher level. In some way they have each contributed to the evolution of a theory explaining the decrement in performance associated with a transition to a lower workload level. Upon careful examination of these studies, however, it appears that design and/or methodological limitations might have compromised data integrity and consequently cast doubt on the generalizability of their findings.

Design and or Methodological Limitations of Previous Studies

One concern with the preceding studies is that each used relatively small participant groups. This fact might have not only compromised the integrity of the statistical analyses, but also raised a question regarding the representativeness of the findings. For example, Cumming and Croft (1973) employed four male participants, while Goldberg and Stewart (1980) limited their study to 24 female participants. Matthews (1986) employed 80 participants, but they were distributed over eight conditions resulting in only 10 participants per comparison group.

A second concern arises from an insufficient number and/or an absence of practice or training trials. The studies of Cumming and Croft (1973) and Goldberg and Stewart (1980) neglected to provide participants with training sessions. This is of special concern given that a minimum level of training is essential to bring participants to a reasonably stable performance level and consequently minimize contamination of test data from learning effects (Schlegel & Gilliland, 1990). Although Matthews (1986) provided a training session for all participants, the training session was 40 minutes long. The characteristics and parameters of the performance task (including workload level) were identical to the testing session. It is problematic that participants were exposed to 40 minutes of the "treatment effect" during the training session, because the nature and sequencing of the training and testing sessions might have left participants bored or exhausted.

Third, all three studies either failed to collect baseline data or collected what might be considered inappropriate baseline data for comparison with test data. The studies of Cumming and Croft (1973) and Goldberg and Stewart (1980) collected no baselines for comparison with test data. Matthews (1986) employed a between-subjects design to collect baselines, however, it seems questionable whether comparisons between baseline and test data were appropriate. Performance from participants engaging in approximately 40 minutes of training trials involving high to low workload transitions, followed by approximately 40 minutes of test trials involving high to low workload transitions, were compared to baselines collected from other participants engaging in 80 minutes at a fixed low workload level. Consequently, given the non-comparability of the baseline samples, it might not be surprising to find response times during the shift from high to low workload levels that were slower than those for a fixed low workload level.

One purpose of Study 1 in the present research project was to examine the effects of workload history on performance. Specifically, after correcting for the methodological limitations of previous studies, does the decrement in performance following a transition from high to low workload persist? Indeed, until the issues raised in the present study are addressed, the certainty of the negative effect on performance remains questionable.

This study had a second purpose as well. The previous studies attempted to explain the nature of the general decrement in performance following a decrease in workload. Should there be a significant effect, this study was designed to explore the nature of these decrements. Recall that the most recent interpretation of the nature of performance decrements following a decrease in work demand was based on strategic persistence (Matthews, 1986). However, another explanation might be plausible as well.

Consider, for example, a theory as elementary as the general adaptation syndrome (GAS; Selye, 1956; 1978). The GAS would suggest that a decrease in initial performance following high workload levels might be a result of recuperative efforts interfering with task performance despite the low workload level, as opposed to strategic persistence. Thus, something more akin to a resource adaptation-depletion model might better explain the performance phenomenon following decreases in workload.

The present study was structured to provide a potential method for comparing the strategic persistence and resource adaptation-depletion models. Specifically, a task was selected that provided the ability to use error rates to test these two models. If Matthews was correct in viewing the decrement in performance as a function of strategic persistence, that is, the person was maintaining unnecessarily high levels of effort for a low-demand task, then there should be evidence of significant commission errors following the transition from a high to a low level of workload. In other words, if the person was overdriving or overworking the task at the low workload level, this should be manifested in an increase in false alarms. However, if the decrement in performance was due to recuperative effort associated with resource depletion, then there ought to be significant evidence of omission errors. In other words, if the person was trying to recover from the high workload level at the low workload level this should be manifested in an increase in timeouts.

Method

Participants

Participants were 198 undergraduate male and female students from the University of Oklahoma, randomly selected from approximately 450 available participants. The only constraint on the random sampling for this study was that selected participants had to have also taken the Eysenck Personality Inventory (EPI, Form A; Eysenck & Eysenck, 1968) during pre-screening sessions because a representative sample of the normal distribution of introversion-extraversion was needed in a third phase of this investigation. Participants received academic credit or extra credit for their participation as one option for fulfilling the requirements of undergraduate psychology courses.

Materials

A computer-based version of the Bakan Vigilance Task (1959) was employed. This is an auditory vigilance task consisting of a series of digits presented to the participants via earphones. Each training and test trial was three minutes in duration. During each trial, participants were instructed to detect odd-even-odd sequences of digits (the signal, e.g., 7-8-3). Participants were instructed to press a specified key on the computer keyboard when they detected a signal. A total of ten signals were presented in each three-minute period among a string of random digits for the high and low conditions (225 and 90 digits total, respectively). Workload level was manipulated by changing the speed of digit presentation (i.e., high workload of 1 digit every 0.8 sec. versus low workload of 1 digit every 2 seconds). In other words, the number of signals was the same for both high and low workload conditions, but total number of digits and the difficulty level varied due to digit rate presentation during the trial. Following signal presentation, the time out period for participants' responses was 4.8 sec.

Procedure

Following random selection from overall participant lists, participants for this study were contacted by telephone and were invited to participate. Of the participants contacted, only five individuals declined participation. When participants arrived at the laboratory they were seated at individual workstations and asked to complete informed consent forms. Workstation partitions minimized distraction from other participants yet permitted individuals to view the experimenter during instruction administration.

Pilot Study

Prior to the study reported in this manuscript, a pilot study (N = 40 participants) was conducted to determine the amount and type of training that would be needed to assure reasonably stable performance levels from most participants. The training regimen was patterned after that used by Schlegel and Gilliland (1990), who found that thorough instructions followed by five to six, three-minute trials were sufficient to produce reasonably asymptotic performance on a variety of human performance tasks. This pilot study confirmed that a thorough instruction set combined with an initial three-minute (low workload) trial that included feedback to the participant, followed by six, three-minute training trials (three, three-minute trials at high difficulty and three, three-minute trials at low difficulty, counterbalanced in their presentation), was sufficient to attain reasonably stable performance levels. A visual inspection of the pilot data suggested that, following the 18-minute training session, participants understood and were performing the task well, therefore lessening the likelihood of test data contamination from learning effects.

Training and Baseline Data Collection

Participants performed the Bakan Vigilance Task during three phases: training, baseline, and the experimental testing session. Table 1 presents the counterbalanced orderings for all training, baseline, and test trials. Each of the 198 participants was block randomly assigned to one of the eight training/baseline/testing sequences in Table 1.

Each participant was first familiarized with the location and operation of the computer response keys relevant for performing the monitoring task. Participants then engaged in the initial feedback trial followed by an 18-minute training session (three, three-minute trials at high difficulty and three, three-minute trials at low difficulty, counterbalanced in their presentation) to ensure understanding of task instructions and lessen the likelihood of data contamination with learning effects. Training session trials were followed by an 18-minute baseline session (three, three-minute trials at high difficulty and three, three-minute trials at low difficulty in counterbalanced order) to establish baseline data for later comparisons. To minimize fatigue, five-minute rest breaks were given between each series of three training and three baseline trials. During these breaks, participants were required to engage in a low-demand distractor task (i.e., completion of participant demographic survey). The three trials of either training or baseline were each three minutes long and were presented with no discernable break between trials (i.e., appeared to be nine continuous minutes). Following the second series of three baseline trials and prior to the testing session, participants were given a 15-minute break during which they completed the form mentioned previously.

Testing Session

A participant's assignment to the A or B testing conditions was based on the training/baseline/testing protocol sequence to which a participant was originally block randomly assigned (see Table 1). During the test session, participants in the A condition engaged in three, three-minute trials at high task difficulty followed immediately by three, three-minute trials at low task difficulty. In contrast, participants in the B test condition engaged in three, three-minute trials at low task difficulty followed immediately by three,

three-minute trials at high task difficulty. These two test conditions created a situation where participants developed a workload history at one workload level and then moved immediately to a dramatically different workload level. Thus, transitions between workload levels during the A and B test sessions were uninterrupted by rest periods (i.e., the total of six trials that included the shift in workload) was perceived by the participant as 18 minutes of continuous work. Participants were run between the hours of 9:00 a.m. to 4:00 p.m. to control for time-of-day effects (Revelle, Humphreys, Simon & Gilliland, 1980).

Results and Discussion

Study 1 was designed to examine workload history phenomena while addressing a number of design and methodological limitations identified in previous studies. Previous studies have reported a decrement in performance when one moves from a high to low workload level, but these same studies did not appear to control a number of factors that could have influenced their results (e.g., sample size, training, and appropriate comparative baselines). The present study used 198 participants providing 104 and 94 participants in conditions A (sudden increase) and B (sudden decrease), respectively. This provided approximately 25 participants in each trial sequence of the training-baseline-testing protocol that controlled for task-order effects. This study also provided participants with a training session of 18 minutes (including counterbalanced high and low difficulty levels). This training session lessened the likelihood of data contamination. Finally, this study included carefully constructed baseline performance trails that allowed comparisons with test trial data at comparable workload levels.

For purposes of baseline and test data comparisons, median baseline scores were computed for each participant. In other words, from the appropriate set of these baseline trials (low or high, given the participant's assignment to condition A or B, respectively), the median trial was selected for comparison to performance on the three (three-minute) later appropriate test trials. Recall that sequence orderings for all training and baseline trials were counterbalanced in an effort to control for possible task order effects (see Table 1). A (task order X condition) repeated measures analysis of variance was used to test whether task order had a significant effect. No significant main effects for task order or interactions were found for any of the dependent variables. Therefore, task order was not considered a significant factor in this experiment. Then separate repeated measures analyses of variance were conducted for each of the A and B test conditions (including the appropriate median baseline and the three test trials at high or low) for the dependent measures of correct responses, response time, and total errors, with alpha level controlled by Dunn's procedure. Correct responses consisted of the number of responses made within 4.8 seconds following signal presentation [1]. Response time was recorded only for correct responses. Total errors represented the combination of false alarms and time out errors. The repeated measures analyses of variance noted above were significant for correct responses (CR), response time (RT), and total errors (TE) data for both condition A (CR: $F(3,309) = 5.41$, $p = .002$; RT: $F(3,309) = 3.40$, $p = .021$; and TE: $F(3,309) = 8.19$, $p = .0001$) and B (CR: $F(3,279) = 7.10$, $p = .0002$; RT: $F(3,279) = 4.09$, $p = .009$; and TE: $F(3,279) = 6.08$, $p = .0008$).

Significant ANOVA results were further analyzed using complex multiple contrasts among relevant means. Specifically, median baseline scores were contrasted with each of the appropriate three test trials. For example, in condition A (sudden decrease), the median low baseline score was compared with the three test trials at low difficulty. The mean correct responses and mean response time data for each trial, as a function of condition, are presented in Figure 1, while the mean total errors data for each trial, as a function of condition are indicated in Figure 2.

To examine immediate effects of workload history, a participant's performance on the appropriate median baseline trial was compared to the first test trial (test trial 1) following the sudden shift in workload. Examining the second and third trials (test trials 2 & 3) following the sudden shift in workload provided a method for examining longer-term or time course effects of workload history. Results of contrast analyses indicated significant contrasts between baseline and the first test trial for correct responses and total errors following a workload shift for the A (sudden decrease) condition (CR: F (1, 103) = 11.05, p = .001 and TE: $F(1,103) = 15.71$, $p = .0001$), but not for the B (sudden increase) condition (CR: $F(1, 93) = .48$, p = .490 and TE: $F(1,93) = .71$, $p = .402$). No significant differences were found for response time between baselines and the first test trial following a sudden workload shift in conditions A or B. Therefore, at least in terms of correct responses and total errors there does appear to be a relatively immediate decrement in performance in the three-minute period following a sudden shift from high to low workload. In addition, it appears that an immediate shift from low to high workload did not result in an immediate statistically significant decrement in correct responses, total errors, nor response time during this same initial three-minute period. Consequently, after correcting for the design and methodological limitations of previous studies, the results of this study provided evidence for the previously reported decrement in performance immediately following a transition from high to low workload.

The design of this study provided the opportunity to examine this decrement within a broader time course as well. Contrast analyses examining the second and third three-minute trials (trials 2 and 3) following the shift in workload for correct response and total error data revealed the existence of a persistent significant decrement in performance following a shift in workload level for condition A (sudden decrease) (CR: trial 2 = $F(1,103) = 6.24$, $p = .014$; CR: trial 3 = $F(1,103) = 12.50$, $p = .0006$; and TE: trial 2 = $F(1,103) = 8.51$, $p = .004$; TE: trial 3 = $F(1,103) = 20.22$, $p = .0001$). Contrast analyses examining trials 2 and 3 for condition B revealed significant findings only for trial 3, (CR: trial 3 = $F(1, 93) = 14.50$, $p = .0003$; TE: trial 3 = $F(1, 93) = 16.59$, $p = .0001$). The nature of this performance decrement varied somewhat in time course depending on the type of workload history. Although the performance decrement in correct responses and total errors for condition A (sudden decrease) was immediate, it also appears that the performance decrement persisted for the full nine minutes of test trials. The performance decrement effect was also present, but delayed for condition B (sudden increase), emerging in the third testing trial. This is an important finding given that previous research has suggested that only a shift from a higher workload to a lower workload level results in a greater performance decrement as compared to a shift from a lower workload to a higher workload level (Cumming & Croft, 1973; Goldberg & Stewart, 1980; Matthews, 1986). Consequently, these findings also extended the apparent time course of the effect. Previous studies examining the effects of workload history limited their data sampling to much shorter periods of time following the workload shift, ranging from 10-72 seconds

(Cumming & Croft, 1973; Goldberg & Stewart, 1980; Matthews, 1986), which may explain the discrepancy in results between this and other studies. Data from the present study suggested that the direction of the shift is less important than the shift itself, especially when considering longer-term effects (i.e., both high to low workload and low to high workload shifts have significantly longer-term negative effects on performance). Thus, the dynamics of performance decrements associated with workload history may be considerably more complex than originally thought.

Response time was also investigated with contrast analyses. While no immediate (or first test trial, trial 1) effects were found, analyses of later test trials revealed decrements in RT performance during the last test trial, trial 3, for both A ($F(1, 103) = 8.61$, $p = .004$) and B ($F(1, 93) = 5.89$, $p = .017$) test conditions. So, the results of this study confirm that immediate decrements in performance appear to occur after a reduction in workload, but that either an increase or decrease in workload can lead to a loss in accuracy and a slowing of response time. It appears that for this task, however, the RT variable may be less sensitive than the CR and TE variables to the development of the decrement in performance.

In addition to the analyses of CR, RT and TE, recall the two competing theoretical models offered to explain the decrement in performance after a high to low workload shift. Matthews (1986) suggested strategic persistence as an explanation for the decrement in performance. An alternative explanation suggested a resource adaptation-depletion model. It was noted that examining errors of omission or commission might provide a method to test which of these theoretical models best explains the decrement in performance. The absence of a response within the time out period (4.8 sec.) following the presentation of a signal resulted in an error of omission or time out, while errors of commission or false alarms were recorded for responses in the absence of a signal. Errors of commission seem more likely to occur if a person is "overdriving" a task (as strategic persistence would suggest), while errors of omission would be more likely if a person were seeking an opportunity to overcome a resource depletion state.

To examine the nature of the performance decrement following a sudden decrease in workload level (condition A), repeated measures analyses of variance were conducted for errors of omission and errors of commission separately. The analyses identified statistically significant differences for both errors of omission, or time outs ($F(3, 309) = 5.41$, $p = .002$) and errors of commission, or false alarms ($F(3, 309) = 4.34$, $p = .006$). The significant effects were investigated using multiple contrasts among pertinent means. Mean time out and false alarm data for each trial are presented in Figure 3. The comparison of baseline scores and the first test trials provided the basis to investigate the nature of the immediate decrement following the sudden decrease in workload level. These contrasts revealed significant negative effects for both time outs and false alarms. Given the nature of the findings, it is difficult to unambiguously discriminate between the two theories on the basis of these variables. However, it seems that a significant increase in time outs would appear to be incompatible with a theoretical explanation based on strategic persistence. If, in fact, participants were maintaining unnecessarily high levels of effort, it is unlikely that they would commit significantly more errors of omission, or time outs. More importantly, a significant increase in false alarms might not be incompatible with a theoretical explanation based on resource adaptation-depletion. From a resource adaptation-depletion perspective, if participants were responding more slowly to signals, these correct responses made after the time out period would have been incorrectly or misleadingly registered as errors of commission, or false alarms. Unfortunately, the constraints on the data collection method for the present task do not permit a definitive test of this post-hoc hypothesis.

In contrast to Matthew's strategic persistence theory, the resource adaptation-depletion perspective offers another possible explanation for the results of this study—one based on motivation. Specifically, participants might have demonstrated a simple lack of motivation while experiencing a state of depleted resources. This might be responsible for the significant increase in both false alarm and time out errors. In other words, it is possible that loss of motivation might result in random or reduced key pressing producing significantly more errors of both types.

Given that the present study not only found evidence for the short-term or immediate performance decrement effect (following a high to low workload shift—condition A) previously reported in the literature, but also for a longer-term performance decrement effect for both A and B conditions (see Figures 1 & 2), it seemed a logical step to investigate the nature of decrements in a longer time course. Contrast analyses examining trial 2 and trial 3 data revealed significantly more false alarms for trial 3 ($F(1, 103) = 12.40$, $p = .0006$), and timeouts for trial 2 ($F(1, 103) = 6.24$, $p = .014$) and trial 3 ($F(1, 103) = 12.50$, $p = .0006$) in condition A (compared with baseline data).

In addition, separate repeated measures analyses of variance were conducted for errors of commission (false alarms) and errors of omission (time outs) for condition B, to examine longer-term effects. Results were significant for errors of omission, or time outs $F(3,279) = 7.10$, $p = .0002$, but not for errors of commission or false alarms $F(3,279) =$

2.27, $p = .09$. Significant effects were further analyzed with multiple contrasts. Results of contrast analyses indicated a significant contrast between baseline and test trial 3 ($F(1, 93) = 14.50$, $p = .0003$) for time outs. Furthermore, an inspection of the means indicates a trend of fewer time outs immediately following the shift, followed by a gradual increase, resulting in a delayed performance decrement.

The workload history theories that were previously discussed were proposed specifically to explain the nature of reported performance decrements following decreases in workload levels. However, a general theory of workload history should be applicable to workload in general. In other words, previous workload history theories do not provide predictive efficiency regardless of the direction of the workload shift.

An alternative theory, such as one based on the general adaptation syndrome (Selye, 1956; 1978) may have broader predictive or explanatory value. Predictions for condition A based on such a theory have already been noted. Interestingly, such a theory could also address the result found in condition B. This theory would suggest that a modest period of low workload would not seriously deplete resources. Though not statistically significant, the initial increase in performance efficiency following a shift to a higher workload level as seen in condition B, low to high timeouts (see Figure 3, trial 1) might be a result of recruiting the necessary resources to meet the demand of the task. Although resources are being used at a low workload level, significant depletion is unlikely. Consequently, when a shift occurs, the individual is initially able to recruit the resources necessary to meet the demands of the task and perform at the same level or at a higher level. The GAS-based theory would also suggest that as the high workload level continues, the later decrement in performance (see trial 3) might be a result of resource depletion and/or subsequent recuperative efforts. Thus, following the longer period at the high workload level, increased resource depletion becomes evident. However, the strategic persistence model fails to provide a satisfactory explanation for the initial improvement in performance followed by the gradual decrement in the time out data for condition B. In fact, the strategic persistence model would predict significantly more (not less) time outs following the immediate shift from the low to high workload level.

Furthermore, contrast analyses between appropriate baselines and test trials were found to be significant for false alarms (trial 1 and trial 3) and time outs (trial 1, trial 2, and trial 3) in condition A. As previously mentioned, the significant increase in time outs following a sudden decrease in workload (condition A) would appear to be incompatible with the strategic persistence theory, while the significant increase in false alarms might not be incompatible with a theoretical explanation based on resource adaptation-depletion. Again, from a resource adaptation-depletion perspective, if participants were responding more slowly, correct responses after the time out period could have been incorrectly registered as errors of commission, or false alarms. Indeed, it appears that the resource adaptation-depletion model offers a more robust explanation of the data as compared to the strategic persistence theory.

Study 2

Based on the results of Study 1, a logical question would be the extent to which time on task or fatigue might have influenced the results. A second study was conducted to investigate the presence of a time on task or fatigue effect. It was expected that if time on task or fatigue was a significant factor, the effect might be present if not more pronounced over a prolonged period at a sustained level of workload. Consequently, the workload history paradigm employed required participants to work at a fixed workload level following a training/baseline protocol identical to that of Study 1 (see Table 2).

Method

Participants

Participants were 37 undergraduate male and female students from the University of Oklahoma, randomly selected from approximately 450 available participants. Participants received academic credit or extra credit for their participation as one option for fulfilling the requirements of undergraduate psychology courses.

The methods for the second study are essentially the same as those in the first study (i.e., materials, training and baseline data collection), with the exception of the testing session.

In Study 2, the participant's assignment to the C (N = 18) or D (N = 19) testing conditions was based on the training/baseline/testing protocol sequence to which a participant was originally randomly assigned (see Table 2). During the test session, participants in the C condition engaged in six, three-minute trials at high task difficulty. In contrast, participants in the D test condition engaged in six, three-minute trials at low task difficulty. These two test conditions created a situation where participants engaged in a workload history at a single workload level. Transitions between three-minute trials were uninterrupted by rest periods. Thus, the transition times between three-minute trials for the C and D test conditions were no different than the transition time within the series of high or low trials (i.e., in the C condition the six, three-minute test trials at high workload was perceived by the participant as 18 minutes of continuous work).

Results and Discussion

The median baseline scores and test trial data were used in this analysis much like that in Study 1. Separate repeated measures analyses of variance were conducted for correct responses, response time and total errors measures for condition C and D. These analyses yielded no significant time on task effects for correct responses, response time nor total errors data for both condition C (CR: $F(6,102) = .97$, $p = .484$; RT: $F(6,102) = 1.85$, $p = .172$; and TE: $F(6,102) = 1.47$, $p = .267$) and D (CR: $F(6,108) = 2.20$, $p = .110$; RT: $F(6,108) = 1.02$, $p = .461$; and TE: $F(6,108) = .796$, $p = .589$). The correct responses, response time and total errors data for each trial are indicated in Figures 4 and 5. Therefore, time on task or fatigue was not considered a significant factor in this experiment, and is considered unlikely to have been a significant factor in Study 1.

General Discussion

The purpose of this project was twofold. First, in Study 1 the effects of workload history were examined while addressing several design and methodological limitations that posed threats to the internal and external validity of previous studies. The results of Study 1 indicated, at least in terms of correct responses, a relatively immediate significant decrement in performance following a sudden shift from high to low workload, but no immediate significant decrement following a sudden shift from low to high workload. Thus, this study supported the previously reported finding of an immediate decrement in performance following a sudden decrease in workload that does not always exist immediately following a sudden increase in workload level. However, this study is the first to suggest that both conditions A (sudden decrease) and B (sudden increase) result in negative effects in a longer time course. In addition, Study 1 was designed to test two competing theories developed to explain the nature of responses to these changes in workload level. The findings made it difficult to discriminate definitively between the two theories, but post-hoc interpretations for both immediate and longer-term effects favored a resource adaptation-depletion theory.

In addition, Study 2 investigated the extent to which time on task or fatigue might have been responsible for the decrement in performance following sudden transitions in workload (reported in Study 1). This study yielded no significant findings for participants working at neither high nor low workload level over a prolonged period of time. Thus, it appears that the decrement in performance following workload transitions might in fact be a consequence of something inherent in the shift rather than an effect of fatigue.

Lastly, a third study explored the relationship between personality and workload history. Specifically, it examined the degree to which manipulating workload history might be effective in exploring the personality trait of introversion-extraversion. This third study is a re-analysis of the data from Study 1 based on the introversion-extraversion dimension (see Appendix for a more extensive discussion).

Nonetheless, it is becoming increasingly apparent that a simple explanation addressing responses to a change or changes in workload level fails to address the complexities involved in the more dynamic effect of workload history on performance. Improvements in methodology (i.e., adjustments in task difficulty), and changes in types of tasks need to be considered in future studies before we can come to a better understanding of the effect of workload history on performance. In addition, further exploration of the relationship between personality and workload history might have considerable promise for expanding our understanding not only of personality, but also the dynamic influences of workload history on performance.

References

Bakan, P. (1959). Extraversion-introversion and improvement in an auditory vigilance task. British Journal of Psychology, 50, 325-332.

Broadbent, D. E. (1965). A reformulation of the Yerkes-Dodson law. British Journal of Mathematical & Statistical Psychology, 18(2), 145-157.

Cumming, R. W., & Croft, P. G. (1973). Human information processing under varying task demand. Ergonomics, 16, 581-586.

Cox, D. R. (1957). A note on grouping. Journal of the American Statistical Association, 52, 543-547.

Davies, D., Hockey, J., & Taylor, A. (1969). Varied auditory stimulation, temperature differences, and vigilance performance. British Journal of Psychology, 60, 453-457.

Eysenck, H. J. (1967). The biological basis of personality. Springfield, Illinois: Thomas.

Eysenck, H. J. (Ed.). (1981). A model for personality. New York: Springer-Verlag.

Eysenck, H. J., & Eysenck, S. B. G. (1968). The Eysenck Personality Inventory. San Diego: Educational and Industrial Testing Services.

Eysenck, H. J., & Eysenck, M. (1985). Personality and individual differences. New York: Plenum Press.

Franks, C. (1957). Personality factors and the rate of conditioning. British Journal of Psychology, 46, 119-126.

Gale, A. (1973). The psychophysiology of individual differences: Studies of extraversion and the EEG. In P. Kline (Ed.), New approaches in psychological measurement. London: Wiley.

Geen, R. G. (1976). Personality, the skein of behavior. St. Louis: Mosby.

Gilliland, K. (1980). The interactive effect of introversion-extraversion with caffeine-induced arousal on verbal performance. Journal of Research in Personality, 14, 482-492.

Goldberg, D. R., & Stewart, M. R. (1980). Memory overload or expectancy effect? 'Hysteresis' revisited. Ergonomics, 23, 1173-1178.

Howarth, E., & Eysenck, H. J. (1968). Personality differences in serial learning under distraction. Journal of Experimental Research in Personality, 3, 114-116.

Keister, M. E., & McLaughlin, R. J. (1972). Vigilance performance related to extraversion-introversion and caffeine. Journal of Experimental Research in Personality, 6, 5-11.

Koelega, H., S. (1992). Extraversion and vigilance performance: 30 years of inconsistencies. Psychological Bulletin, 112 (2), 239-258.

Krupski, A., Raskin, D., & Bakan, P. (1971). Physiological and personality correlates of commission errors in an auditory vigilance task. Psychophysiology, 8, 304-311.

Malmo, R. B. (1957). Anxiety and behavioral arousal. Psychological Review, 64, 276-287.

Matthews, M. L. (1986). The influence of visual workload history on visual performance. Human Factors, 28 (6), 623-632.

Matthews, G., & Gilliland, K. (1999). The personality theories of H.J. Eysenck and J. A. Gray: A comparative review. Personality and Individual Differences, 26, 583-626.

Osborne, J. W. (1972). Short- and long-term memory as a function of individual differences in arousal. Perceptual and Motor Skills, 34, 587-593.

Revelle, W., Amaral, P., & Turriff, S. (1976). Introversion/extraversion, time stress, and caffeine: Effect on verbal performance. Science, 192, 149-150.

Revelle, W., Humphreys, M., Simon, L., & Gilliland, K. (1980). The interactive effect of personality, time of day, and caffeine: A test of the arousal model. Journal of Experimental Psychology: General, 109, 1-26.

Schlegel, R. E., & Gilliland, K. (1990). Evaluation of the Criterion Task Set-Part I CTS Performance and SWAT Data-Baseline Conditions (U) (Tech. Report AAMRL-TR-90-007). Wright Patterson, Norman, Oklahoma.

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

Selye, H. (1978). The stress of life (Rev. ed.). New York: McGraw-Hill.

Smith, A., P. & Maben, A. (1993). Effects of sleep deprivation, lunch, and personality on performance, mood, and cardiovascular function. Physiology and Behavior, 54 (5), 967-972.

Stelmack, R. (1981). The psychophysiology of extraversion and neuroticism. In H. Eysenck (Ed.) A model for personality (pp. 38-60). New York: Springer.

Stelmack, R. M., & Wilson, K. G. (1982). Extraversion and the effects of frequency and intensity on the brainstem evoked response. Personality and Individual Differences, 3, 373-380.

Thackray, R., Jones, K., & Touchstone, R. (1974). Personality and physiological correlates of performance decrement on a monotonous task requiring sustained attention. British Journal of Psychology, 65, 351-358.

Wilson, G. (1978). Introversion-extroversion. In H. London & J. E. Exner (Eds.) Dimensions of personality (pp. 217-261). New York: John Wiley & Sons.

Footnote

[1]Because the time-out period was 4.8 seconds and the stimulus presentation rate was .8 or 2.0 seconds (high - low, respectively) it might have been possible for participants to respond to an odd-even-odd signal after additional digits were presented. Inspection of RT data confirmed that no correct response by any participant was made slower than 2.6 seconds.

Table 1

Counterbalanced Orderings for Training, Baseline and Test Trials for Conditions A and B

Training and Baseline – Conditions A and B

Training-Low	Training-High	Baseline-Low	Baseline-High
Training-High	Training-Low	Baseline-Low	Baseline–High
Training-Low	Training-High	Baseline–High	Baseline-Low
Training-High	Training-Low	Baseline–High	Baseline-Low

Test Condition A
(Sudden Decrease)

Test-High	Test-Low
Test-High	Test-Low
Test-High	Test-Low
Test-High	Test-Low

Test Condition B
(Sudden Increase)

Test-Low	Test-High
Test-Low	Test-High
Test-Low	Test-High
Test-Low	Test-High

Table 2

Counterbalanced Orderings for Training, Baseline and Test Trials for Conditions C and D

Training and Baseline – Conditions C and D

Training-Low	Training-High	Baseline-Low	Baseline-High
Training-High	Training-Low	Baseline-Low	Baseline–High
Training-Low	Training-High	Baseline–High	Baseline-Low
Training-High	Training-Low	Baseline–High	Baseline-Low

Test Condition C
(High Workload)

Test-High	Test-High
Test-High	Test-High
Test-High	Test-High
Test-High	Test-High

Test Condition D
(Low Workload)

Test-Low	Test-Low
Test-Low	Test-Low
Test-Low	Test-Low
Test-Low	Test-Low

Figure Captions

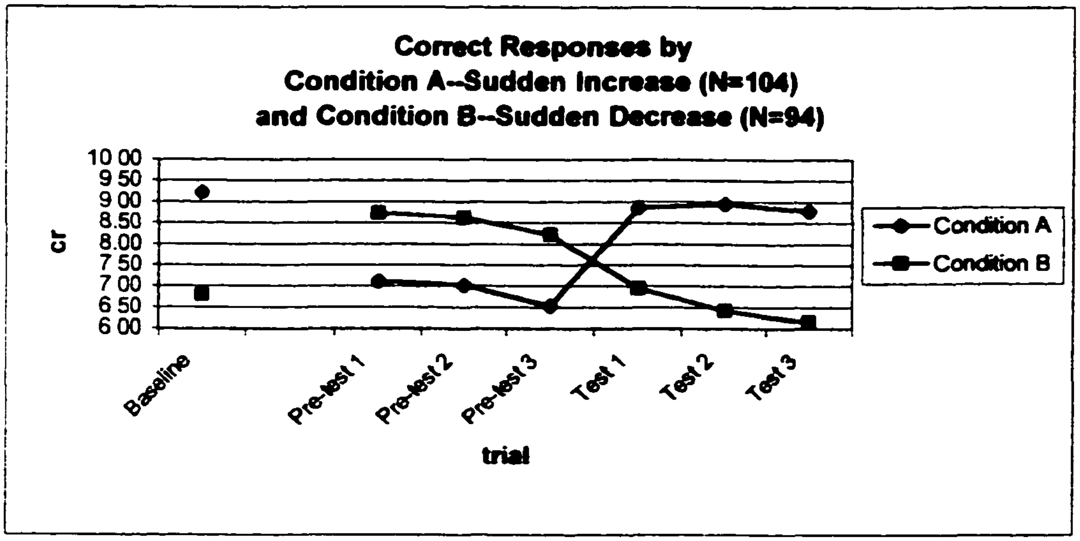
Figure 1. Correct responses and response time as a function of trial for condition A (sudden decrease) and condition B (sudden increase).

Figure 2. Total errors as a function of trial for condition A (sudden decrease) and condition B (sudden increase).

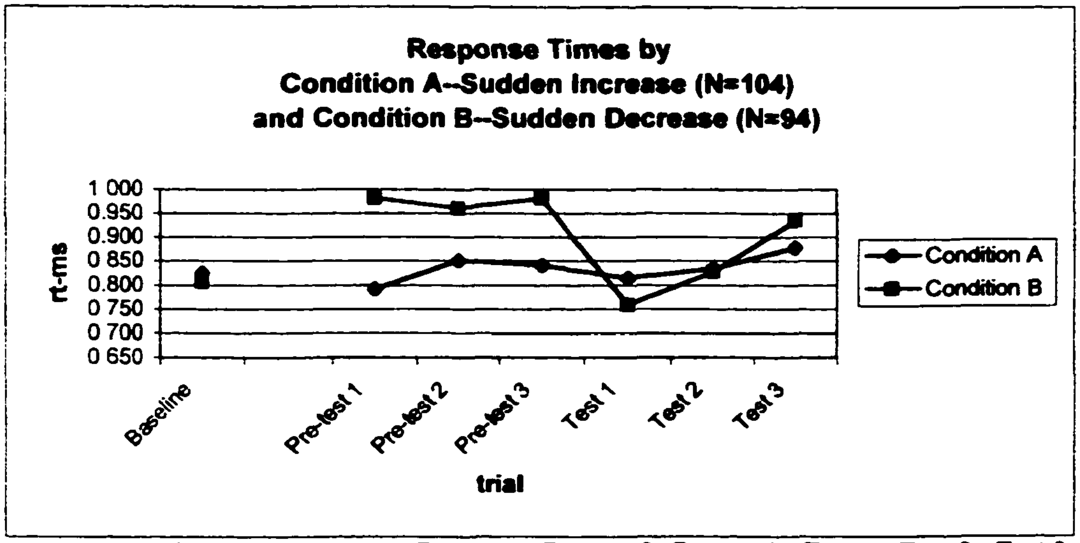
Figure 3. False alarms and time outs as a function of trial for condition A (sudden decrease) and condition B (sudden increase).

Figure 4. Correct responses and response times as a function of trial for condition C (high workload) and condition D (low workload).

Figure 5. Total errors as a function of trial for condition C (high workload) and condition D (low workload).

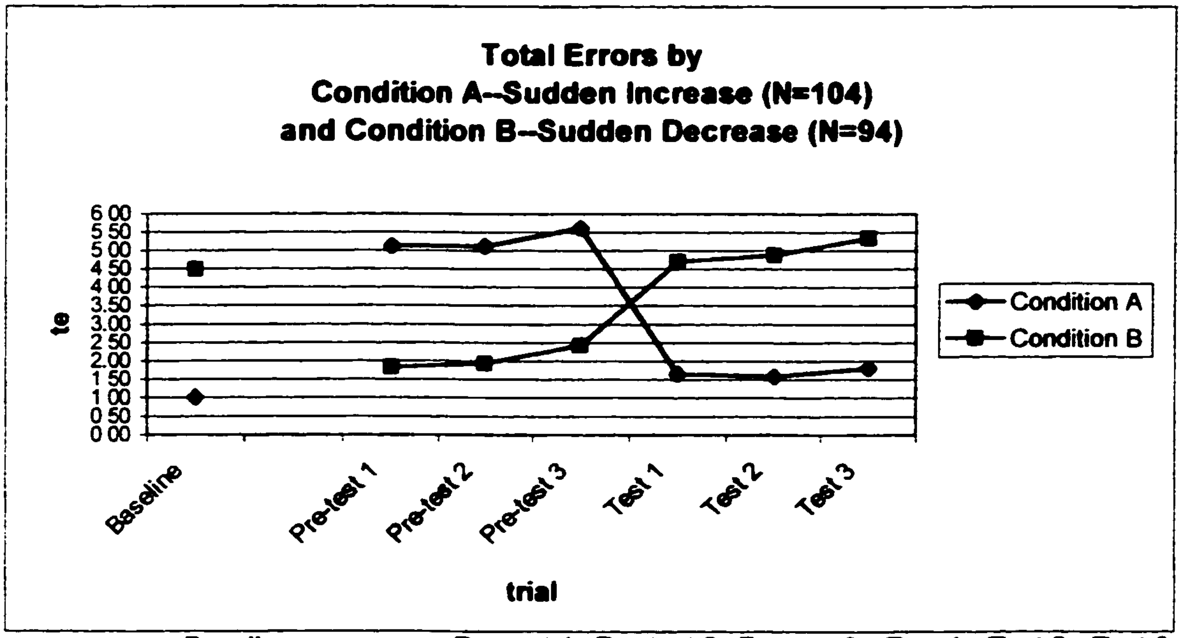


	Baseline	Pre-test 1	Pre-test 2	Pre-test 3	Test 1	Test 2	Test 3
Condition A	9.21	7.10	7.00	6.51	**8.84**	**8.94**	**8.77**
Condition B	6.78	8.71	8.60	8.20	6.93	6.40	**6.14**



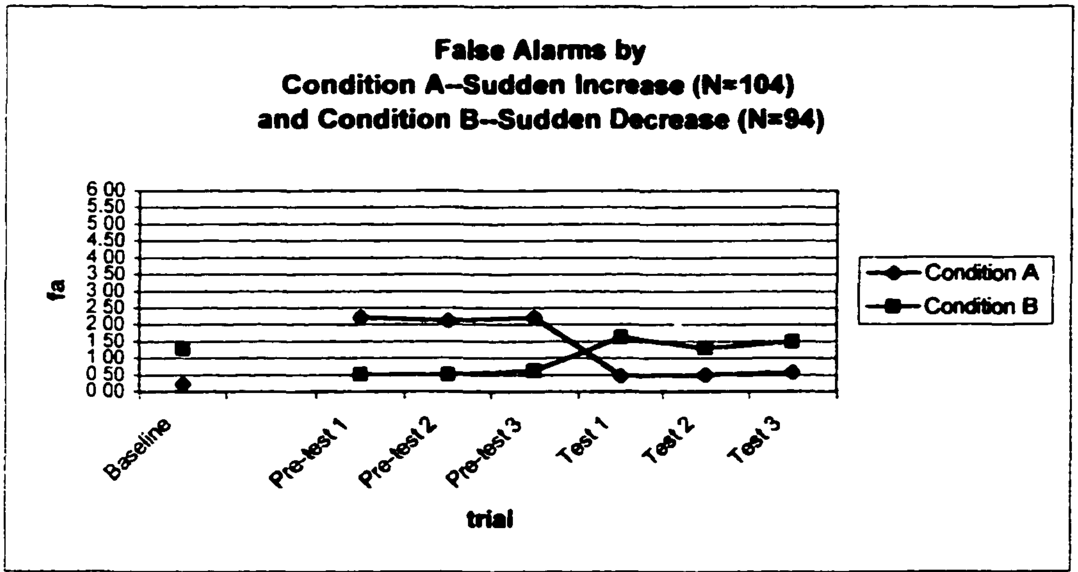
	Baseline	Pre-test 1	Pre-test 2	Pre-test 3	Test 1	Test 2	Test 3
Condition A	0.824	0.790	0.850	0.840	0.814	0.833	**0.878**
Condition B	0.805	0.980	0.960	0.980	0.758	0.827	**0.934**

Underlined bold values signify that the mean for trial is significantly different from baseline data ($p<.017$).

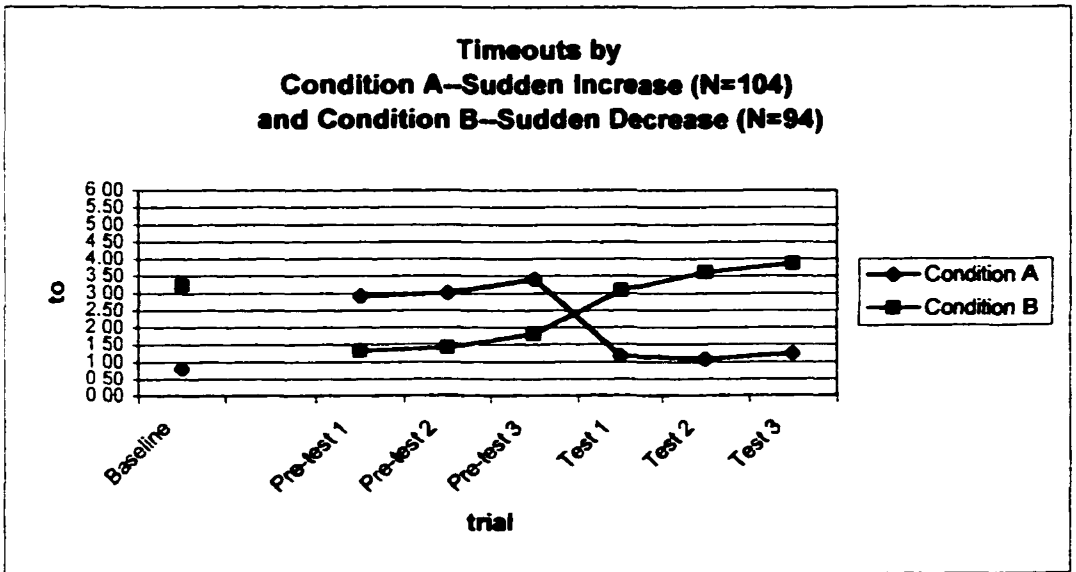


	Baseline	Pre-test 1	Pre-test 2	Pre-test 3	Test 1	Test 2	Test 3
Condition A	1.00	5.10	5.10	5.60	**<u>1.63</u>**	**<u>1.56</u>**	**<u>1.79</u>**
Condition B	4.47	1.80	1.90	2.40	4.69	4.87	**<u>5.33</u>**

<u>Underlined bold values signify that the mean for trial is significantly different from baseline data ($p<.017$).</u>

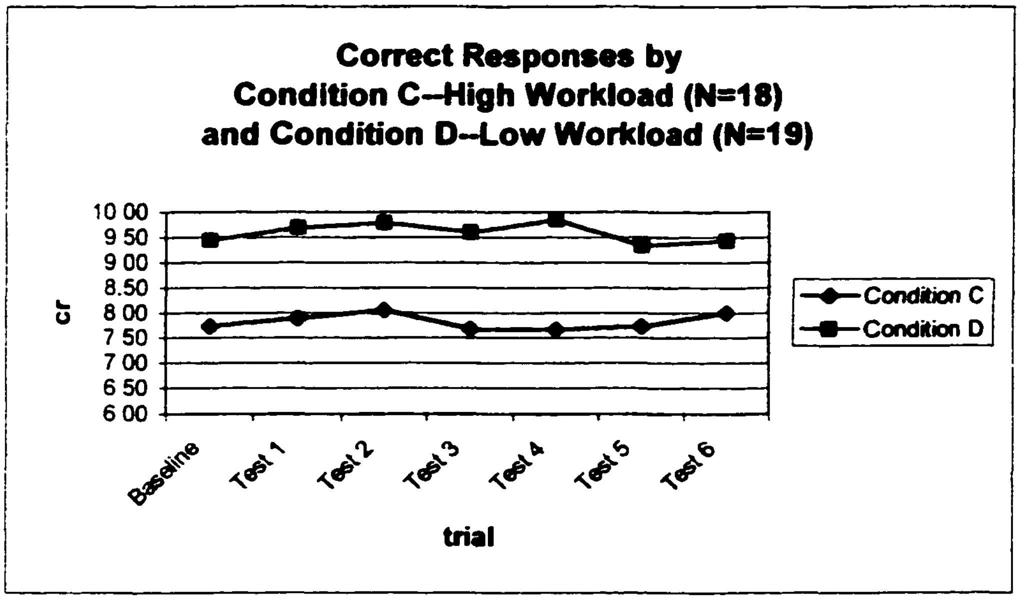


	Baseline	Pre-test 1	Pre-test 2	Pre-test 3	Test 1	Test 2	Test 3
Condition A	0.21	2.20	2.10	2.20	**0.46**	0.50	**0.56**
Condition B	1.24	0.50	0.50	0.60	1.61	1.28	1.47

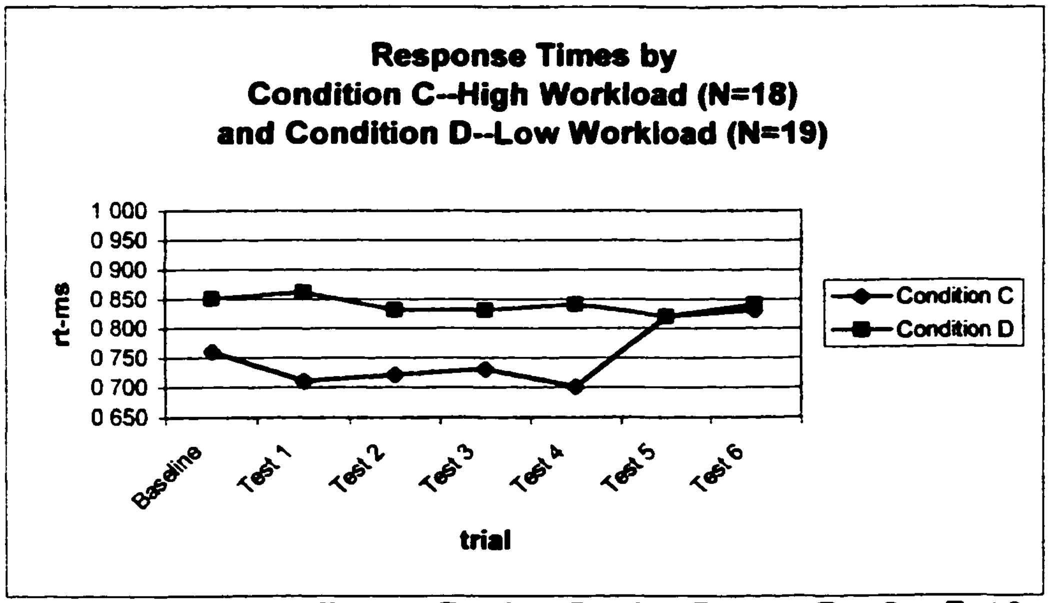


	Baseline	Pre-test 1	Pre-test 2	Pre-test 3	Test 1	Test 2	Test 3
Condition A	0.79	2.90	3.00	3.40	**1.16**	**1.06**	**1.23**
Condition B	3.22	1.30	1.40	1.80	3.07	3.60	**3.86**

Underlined bold values signify that the mean for trial is significantly different from baseline data ($p<.017$).

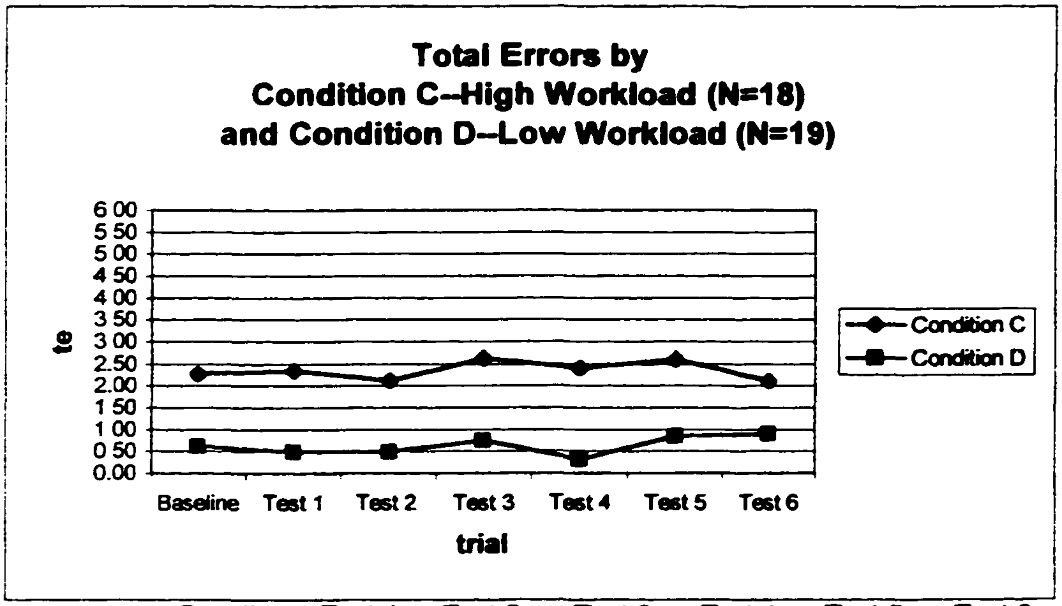


	Baseline	Test 1	Test 2	Test 3	Test 4	Test 5	Test 6
Condition C	7.72	7.88	8.05	7.67	7.66	7.72	8.00
Condition D	9.42	9.68	9.78	9.58	9.84	9.32	9.42



	Baseline	Test 1	Test 2	Test 3	Test 4	Test 5	Test 6
Condition C	0.760	0.710	0.720	0.730	0.700	0.820	0.830
Condition D	0.850	0.860	0.830	0.830	0.840	0.820	0.840

* No significant findings (means for trials are not significantly different from baseline data ($p>.05$)).



	Baseline	Test 1	Test 2	Test 3	Test 4	Test 5	Test 6
Condition C	2.27	2.33	2.11	2.61	2.40	2.60	2.11
Condition D	0.60	0.47	0.47	0.73	0.30	0.84	0.89

* No significant findings (means for trials are not significantly different from baseline data (p>.05)).

Appendix

Effect of Workload History and Introversion-Extraversion on Vigilance Performance

This study was concerned with the degree to which manipulating workload history might be effective in exploring the personality trait of introversion-extraversion. This third study is a re-analysis of the data from Study 1 based on the introversion-extraversion dimension.

The most prevalent theory of introversion-extraversion assumes that behavioral differences between introverts and extraverts are due to physiological arousal differences mediated by the ascending reticular activating system (Eysenck, 1967). According to this theory, introverts were believed to possess higher levels of ascending reticular activating system activity than extraverts. Several studies of introversion-extraversion have employed psychophysiological measures such as: electrodermal activity, electroencephalographic activity (EEG), and evoked potential recordings (see Geen, 1976, Stelmack, 1981, and Gale, 1973, for reviews of psychophysiologically-based research on introversion-extraversion). There has been some inconsistency among findings in the psychophysiological literature because of differences in the way many of the studies were conducted (Gale, 1973). The encouraging finding, however, was the considerable uniformity in research results among studies that carefully controlled experimental variables, such as maintaining moderate task difficulty (Gale, 1973; Stelmack, 1981). Indeed, considerable psychophysiologically-based research has generally been supportive of introversion-extraversion theory (Eysenck & Eysenck, 1985; Gale, 1973; Matthews & Gilliland, 1999; Stelmack & Wilson, 1982).

The differential levels of arousal between introverts and extraverts are also believed to form the basis for the differences observed between these groups in their responses to the environment and on human performance tasks. For example, classical conditioning studies were among the first empirical tests of Eysenck's introversion-extraversion theory. These studies explored the manner in which introversion-extraversion arousal differences mediated responses to environmental stimuli. They established a link between introversion-extraversion, arousal, and conditioning, and demonstrated how the higher levels of arousal associated with introversion facilitate conditioning (Eysenck, 1981; Franks, 1957; Wilson, 1978). Specifically, the higher arousal of introverts is believed to enable them to develop conditioned responses faster and better than extraverts (Eysenck, 1981).

Another property of Eysenck's introversion-extraversion theory is its applicability to performance tasks where levels of difficulty (or arousal) are concerned. Past research using performance tasks has enabled researchers to use well-known relationships, such as the inverted-U relationship between stimulus intensity and performance (Broadbent, 1965; Malmo, 1957), to explore the hypothesized differences between introverts and extraverts. For example, studies have identified significant differences between introverts and extraverts on cognitive performance, particularly memory. Eysenck's theory predicted that introverts (who are more highly aroused than extraverts) should perform better on long-term memory tasks and poorer on short-term memory tasks, as compared to extraverts (Eysenck & Eysenck, 1985). Research has been generally supportive of this hypothesized relationship (Howarth & Eysenck, 1968; Osborne, 1972; Wilson, 1978). Also, verbal performance studies have been employed together with drug-induced arousal to investigate differences between introverts and extraverts, and these findings have also been supportive of Eysenck's introversion-extraversion arousal-based theory (Gilliland, 1980; Revelle, Amaral & Turriff, 1976).

Important for the purpose of this study is the finding that vigilance studies have proven particularly useful for investigating performance differences between introverts and extraverts. Although, there has been some variability in the types of vigilance tasks employed, the objective is essentially one in which the participant is expected to detect inconspicuous auditory or visual signals over a relatively prolonged period of time. Although there are important variables that can produce different research outcomes (e.g., strength of stimulus, environmental stressors, rate of signals, etc.), Eysenck's theory would predict superiority of introverts on most vigilance tasks based on introverts' heightened arousal state. The theory suggests that introverts' higher level of arousal enables them to sustain a heightened level of vigilant attention especially when the task is not particularly stimulating. The prediction that introverts characteristically perform better on vigilance tasks than extraverts has been established with some regularity (Davies, Hockey & Taylor, 1969; Eysenck, 1967, 1981; Eysenck & Eysenck, 1985; Koelega, 1992; Krupski, Raskin & Bakan, 1971; Smith & Maben, 1993; Thackray, Jones & Touchstone, 1974).

While the use of workload history has not been applied to the study of introversion-extraversion, sustained vigilance studies that investigated performance differences between introverts and extraverts might indirectly shed light on the viability of workload history as a method for investigating introversion-extraversion. For example, Keister and McLaughlin (1972) conducted a vigilance study investigating signal detection in each of three consecutive 16-minute periods. Analyses showed a significant relationship between introversion-extraversion and signals detected per period. The significant relationship resulted from discrepant scores in the third 16-minute period. As predicted, there were no differences between introverts and extraverts during the first part of the task. However, toward the end of the task, the performance of extraverts dropped markedly, resulting in only 42% signal detection, as compared to 58% for the introverts. Keister and McLaughlin (1972) offered two explanations for these results. One explanation is based on the superior conditionability and less proneness to extinction of introverts as compared to extraverts. In 1959, Bakan proposed evidence for the reinforcing properties of signal detection during a vigilance task. He suggested that extraverts extinguish more rapidly than introverts during a task in which the presentation of relatively infrequent signals serve as reinforcement for task attention. Inevitably, the appropriate attentive response decreases while inappropriate responses due to attention to irrelevant stimuli and drowsiness continue to reduce the strength of task appropriate attention. Introverts have generally been found to show less extinction and condition more efficiently than extraverts (Eysenck, 1967). The second possible explanation is based on the hypothesized differences in cortical inhibition between introverts and extraverts. In essence, it is proposed that the task redundancy inherent in vigilance tasks generates cortical inhibition, which leads to increased involuntary rest pauses resulting in missed signals. Recall that extraverts are believed to possess strong inhibitory processes that develop quickly but, dissipate slowly as compared to introverts (Eysenck & Eysenck, 1985). Thus, extraverts would exhibit more involuntary rest pauses and consequently miss more signals. Both of these explanations are conceivable and most importantly both are reflective of the hypothesized differential levels of cortical arousal and the accretion and dissipation of inhibition characteristic of introverts and extraverts (Keister & McLaughlin, 1972). Although there were no explicit manipulations of task difficulty, clearly introverts and extraverts responded differently across the successive workload intervals.

An important point might be drawn from this study that serves to demonstrate the utility of manipulating workload history for the study of introversion-extraversion. First, performance differences between introverts and extraverts on the vigilance task, which extended over a 48-minute period, showed that introverts and extraverts responded to their workload environment in unique ways. Based on this finding, one could predict that changes in workload history (which often serve to directly manipulate arousal and may be more representative of many real-world work environments than fixed workload levels) might also result in differential responses for introverts and extraverts. Thus, the experimental marriage between the research domains of introversion-extraversion and workload history might serve to further illuminate the nature of performance differences between introverts and extraverts by providing a unique and more exacting examination of the arousal-related processes that are believed to be responsible for performance differences.

This re-analysis of Study 1 data assesses the relationship between workload history, introversion-extraversion, and performance on an auditory vigilance task. It provides a test of whether the dynamics of workload history can be used to explore the personality trait of introversion-extraversion. Because extraverts are known to possess lower levels of arousal in comparison to introverts, it was hypothesized that changes in workload levels, specifically a decrease in task demand following a period of high workload (condition A of Study 1), might very well result in extraverts having a larger performance decrement relative to introverts. It is difficult to derive a prediction from arousal theory about the effects of an increase in task demand following a period of low workload (condition B of Study 1). In fact, the situation created by condition B of Study 1 (i.e., prolonged period of low stimulation) is well known in the introversion-extraversion literature for producing highly unpredictable (often paradoxical) results (Gale, 1973). Therefore no prediction was made regarding introversion-extraversion performance for condition B.

Method

The methods for the second study are essentially the same as those in the first study, with one exception. In Study 3, the 198 participants of Study 1 that had completed the Eysenck Personality Inventory (EPI, Form A; Eysenck & Eysenck, 1968) were screened to identify the upper and lower 27% of the introversion-extraversion distribution. Participants in the upper 27% were classified as extraverts, and participants in the lower 27% were classified as introverts (Cox, 1957). In addition, to minimize the potential arousing influence of the limbic system (see Eysenck, 1967) all participants were within one standard deviation of the mean EPI neuroticism score for the original sample. The EPI was used to assess the dimensions of extraversion and neuroticism. Reliability estimates for the EPI extraversion scale range from 0.80 to 0.97 and reliability estimates for the EPI neuroticism scale range from 0.81 to 0.91 (see Eysenck & Eysenck, 1968).

Results and Discussion

The median baseline scores and test trial data from Study 1 for introverts and extraverts were used in an individual differences analysis much like that in Study 1.

Separate repeated measures analyses of variance, including an introversion-extraversion factor, were conducted for correct responses, response time and total errors measures for condition A. These analyses yielded no significant effects for personality for correct responses, response time nor total errors data. The correct responses, response time and total errors data by personality group for each trial is indicated in Figures A6 and A7.

It was predicted that changes in workload levels, specifically a decrease in task demand following a prolonged period of high workload (condition A- sudden decrease), would result in extraverts having a larger performance decrement relative to introverts. Results revealed no significant findings confirming this prediction. However, the trends in the data, though not significant, were encouraging as they were consistent with the prediction and previous research findings based on Eysenck's theory. Namely, the performance of extraverts was inferior to that of introverts for CR, RT and TE for condition A. The sudden shift from high to low workload resulted in a non-significant negative trend for both introverts and extraverts. However, why the decline was not significantly worse for extraverts compared to introverts is not clear.

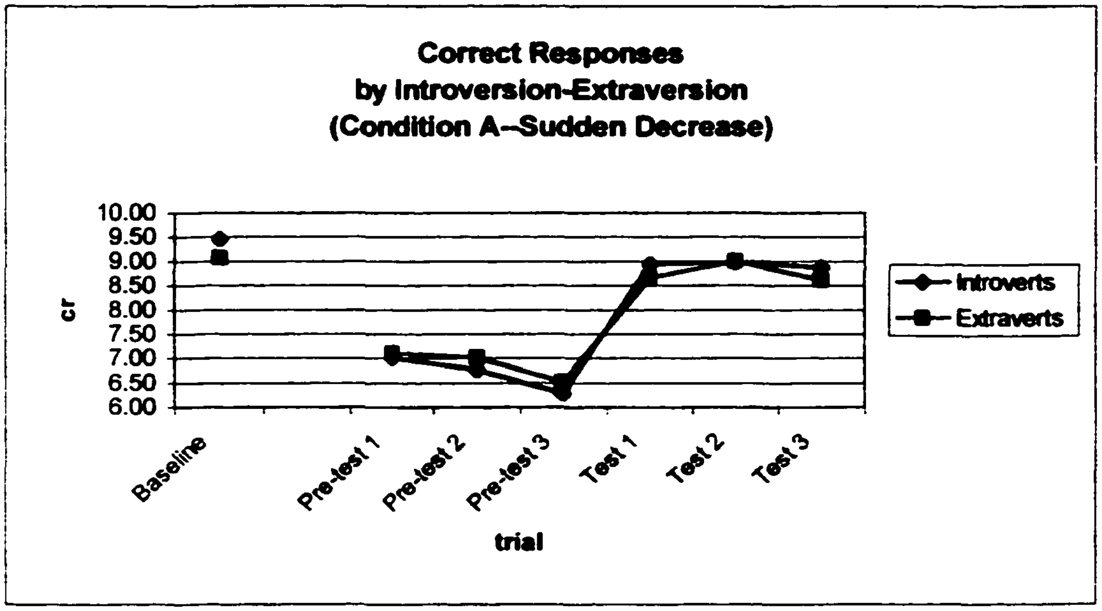
One explanation might be that the difficulty levels (though disparate enough for the general population in Study 1) were not disparate enough to cause significant performance differences between the personality groups. In other words, a lack of greater differences in workload levels may have rendered the task insensitive to the personality effects. However, in determining the parameters for the high and low workload levels of the vigilance task used in this study, pilot studies were conducted in which high workload was increased to a point of participant failure. The rate of signal presentation chosen for the high workload level in this study was slightly below this level, at a point that seemed to provide data consistent with what was considered capable but clearly pressured performance for all participants. Given that pilot data suggested participants experienced significant problems at higher difficulty levels (rendering the data useless in most cases), an explanation suggesting the high workload level in this study was not difficult enough seems unlikely. Nevertheless, future studies could easily explore this by slightly increasing the difference between high and low task difficulty (i.e., making the high workload level of the task harder than the level used in this study).

A second explanation might involve a lack of general task sensitivity. This is improbable given that previous studies using the Bakan Vigilance Task have successfully identified performance differences between introverts and extraverts (Davies, Hockey & Taylor, 1969; Keister & McLaughlin, 1972). However, while the Bakan Task has shown introversion-extraversion differences in previous studies, within the context of this workload history manipulation an effect was not optimized. Consequently, it is possible that using this task within a workload history manipulation may not be effective or sensitive to introversion-extraversion differences. Although no personality effects were found, the effect of time was found significant as in Study 1—that is, the same negative effects on dependent variables as found in Study 1 were found for the combined introversion and extraversion participant groups in Study 3, but no significant differences between the introversion-extraversion groups were found. Nonetheless, this study was the first to suggest that workload history might provide a systematic method for controlling environmental variables that mediate the relationship between introversion-extraversion and human performance capabilities in such a manner useful for testing introversion-extraversion theory. Clearly, however, more studies are necessary before any conclusions can be drawn regarding the utility of the workload history paradigm in exploring the personality trait of introversion-extraversion.

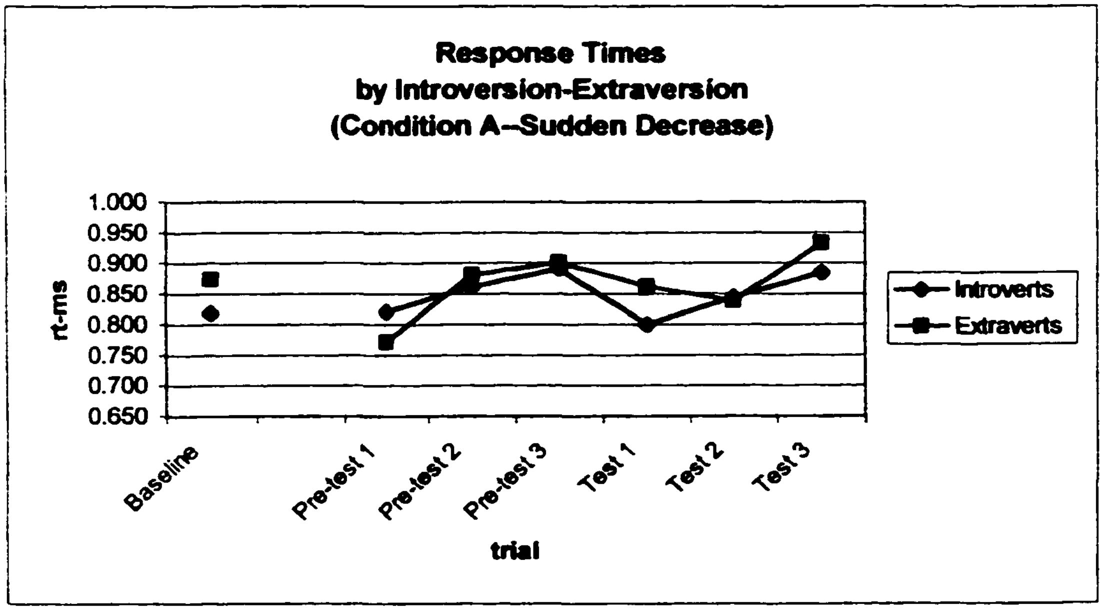
Figure Captions

Figure A6. Correct responses, response time, and total errors as a function of personality and trial for condition A (sudden decrease).

Figure A7. Total errors as a function of personality and trial for condition A (sudden decrease).

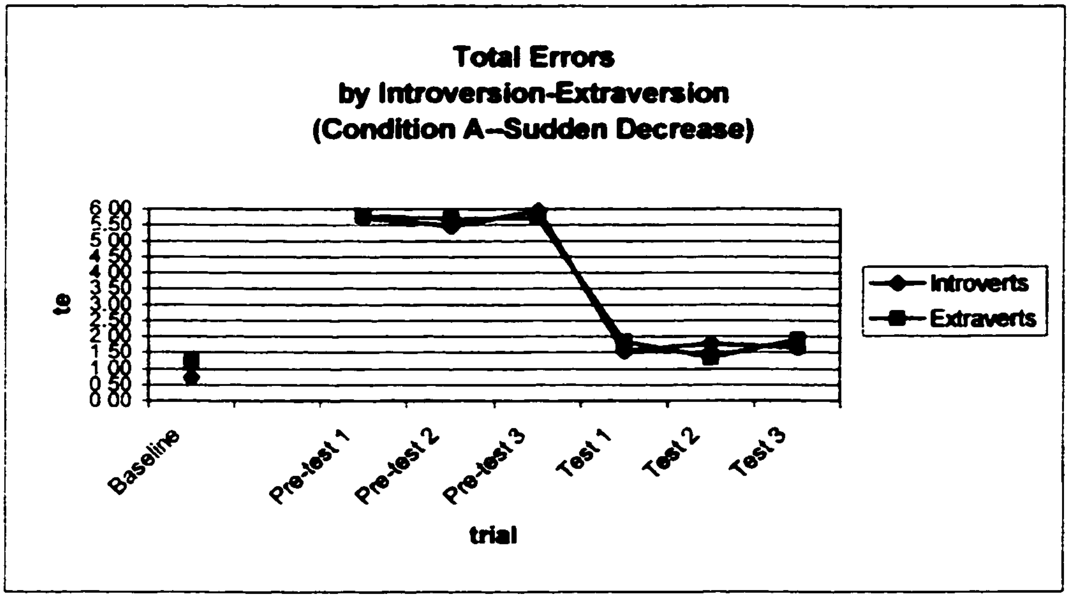


	Baseline	Pre-test 1	Pre-test 2	Pre-test 3	Test 1	Test 2	Test 3
Introverts	9.45	7.00	6.75	6.28	8.93	8.97	8.86
Extraverts	9.06	7 09	7 00	6.51	8.64	9.00	8.61



	Baseline	Pre-test 1	Pre-test 2	Pre-test 3	Test 1	Test 2	Test 3
Introverts	0.818	0.820	0.860	0.890	0.799	0.844	0.884
Extraverts	0.872	0 770	0.880	0.900	0.861	0.837	0.932

* No significant findings (means for trials are not significantly different from baseline data ($p>.05$)).



	Baseline	Pre-test 1	Pre-test 2	Pre-test 3	Test 1	Test 2	Test 3
Introverts	0.72	5.70	5.45	5.92	1.52	1.76	1.62
Extraverts	1.21	5.77	5.67	5.71	1.82	1.33	1.88

* No significant findings (means for trials are not significantly different from baseline data ($p>.05$))